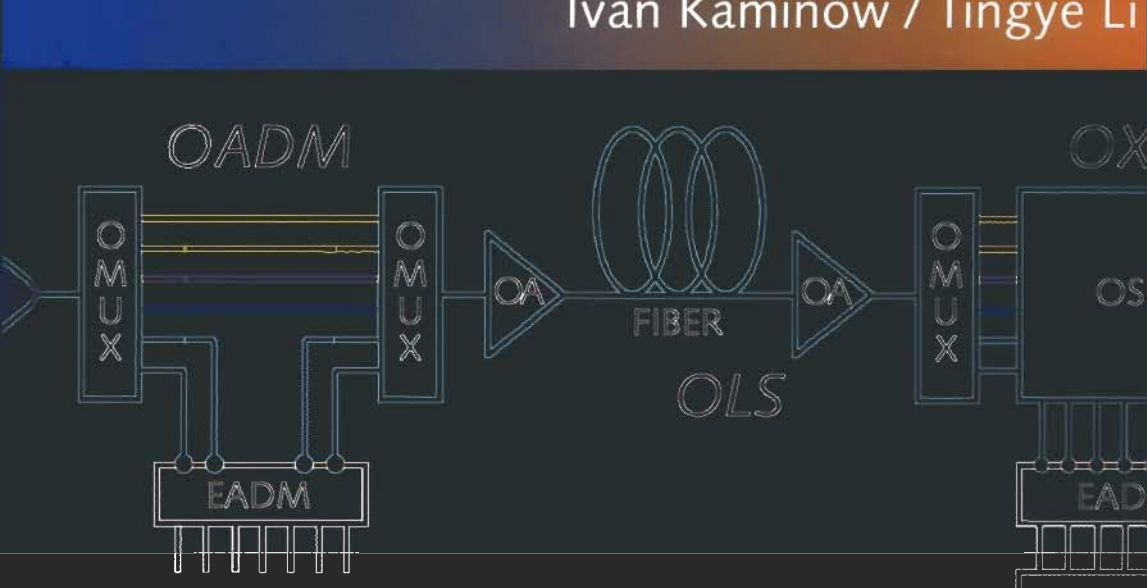
OPTICAL FIBER TELECOMMUNICATIONS IVA

COMPONENTS

Edited by
Ivan Kaminow / Tingye Li



OPTICAL FIBER TELECOMMUNICATIONS IV A COMPONENTS

OPTICAL FIBER TELECOMMUNICATIONS IV A COMPONENTS

Edited by

IVAN P. KAMINOW
Bell Laboratories (retired)
Kaminow Lightwave Technology
Holmdel, New Jersey

TINGYE LI
AT&T Labs (retired)
Boulder, Colorado

ACADEMIC PRESS

An Elsevier Science Imprint

San Diego San Francisco New York
Boston London Sydney Tokyo

This book is printed on acid-free paper. ♾

Academic Press
An Elsevier Science Imprint
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
http://www.academicpress.com

Academic Press
Harcourt Place, 32 Jamestown Road, London NW1 7BY, UK
http://www.academicpress.com

Library of Congress Control Number: 2001098830

International Standard Book Number: 0-12-395172-0

PRINTED IN CHINA
02 03 04 05 06 07 RDC 9 8 7 6 5 4 3 2 1

For Florence and Edith, with love

Contents

Contributors

D. A. Ackerman (A:587), Agere Systems, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Daniel Y. Al-Salameh (A:295), JDS Uniphase Corporation, 100 Willowbrook Road, Bldg. 1, Freehold, New Jersey 07728-2879

Rick Barry (B:198), Sycamore Networks, 10 Elizabeth Drive, Chelmsford, Massachusetts 01824-4111

Polina Bayvel (B:611), Optical Networks Group, Department of Electronic and Electrical Engineering, University College London (UCL), Torrington Place, London WC1E 7JE, United Kingdom

Neal S. Bergano (B:154), Tyco Telecommunications, 250 Industrial Way West, Eatontown, New Jersey 07724-2206

Lee L. Blyler (A:17), OFS Fitel, LLC, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Raymond K. Boncek (A:17), OFS Fitel, LLC, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Michael Cahill (B:198), Sycamore Networks, 10 Elizabeth Drive, Chelmsford, Massachusetts 01824-4111

Gary M. Carter (B:305), Computer Science and Electrical Engineering Department, TRC-201A, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250 and Laboratory for Physical Sciences, College Park, Maryland

Connie J. Chang-Hasnain (A:666), Department of Electrical Engineering and Computer Science, University of California, Berkeley, California 94720 and Bandwidth 9 Inc., 46410 Fremont Boulevard, Fremont, California 94538

Young-Kai Chen (A:784), Lucent Technologies, High Speed Electronics Research, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Xin Cheng (B:329), Sorrento Networks Inc., 9990 Mesa Rim Drive, San Diego, California 92121-2930

Dominique Chiaroni (A:732), Alcatel Research & Innovation, Route de Nozay, F-91461 Marcoussis cedex, France

Kerry G. Coffman (B:17), AT&T Labs-Research, A5-1D03, 200 Laurel Avenue South, Middletown, New Jersey 07748

Jan Conradi (B:862), Director of Strategy, Corning Optical Communications, Corning Incorporated, MP-HQ-W1-43, One River Front Plaza, Corning, New York 14831

Santanu K. Das (A:17), OFS Fitel, LLC, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Emmanuel Desurvire (A:732), Alcatel Technical Academy, Villarceaux, F-91625 Nozay cedex, France

David J. DiGiovanni (A:17), OFS Fitel, LLC, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Christopher R. Doerr (A:405), Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, New Jersey 07733

Adam Ellison (A:80), Corning, Inc., SP-FR-05, Corning, New York 14831

L. E. Eng (A:587), Agere Systems, Room 2F-204, 9999 Hamilton Blvd., Breinigsville, Pennsylvania 18031-9304

Turan Erdogan (A:477), Semrock, Inc., 3625 Buffalo Road, Rochester, New York 14624

René-Jean Essiambre (B:232), Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, New Jersey 07733

Costas N. Georghiades (B:902), Texas A&M University, Electrical Engineering Department, 237 Wisenbaker, College Station, Texas 77843-3128

Nasir Ghani (B:329), Sorrento Networks Inc., 9990 Mesa Rim Drive, San Diego, California 92121-2930

Steven E. Golowich (A:17), Bell Laboratories, Lucent Technologies, Room 2C-357, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Christoph S. Harder (A:563), Nortel Networks Optical Components, Binzstrasse 17, CH-8045 Zürich, Switzerland

Edward Harstead (B:438), Bell Laboratories, Lucent Technologies, 101 Crawford Corners Road, Holmdel, New Jersey 07733

Bogdan Hoanca (B:642), Phaethon Communications, Inc., Fremont, California 96538

J. E. Johnson (A:587), Agere Systems, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Robert M. Jopson (B:725), Crawford Hill Laboratory, Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, New Jersey 07733

Ivan P. Kaminow (A:1, B:1), Bell Laboratories (retired), Kaminow Lightwave Technology, 12 Stonehenge Drive, Holmdel, New Jersey 07733

Bryon L. Kasper (A:784), Agere Systems, Advanced Development Group, 4920 Rivergrade Road, Irwindale, California 91706-1404

William L. Kath (B:305), Computer Science and Electrical Engineering Department, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250 and Applied Mathematics Department, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208-3125

L. J. P. Ketelsen (A:587), Agere Systems, 600 Mountain Avenue, Murray Hill, New Jersey 07974

P. A. Kiely (A:587), Agere Systems, 9999 Hamilton Blvd., Breinigsville, Pennsylvania 18031-9304

Robert Killey (B:611), Optical Networks Group, Department of Electronic and Electrical Engineering, University College London (UCL), Torrington Place, London WC1E 7JE, United Kingdom

Herwig Kogelnik (B:725), Crawford Hill Laboratory, Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, New Jersey 07733

Steven K. Korotky (A:295), Bell Laboratories, Lucent Technologies, Room H0 3C-351, 101 Crawfords Corner Road, Holmdel, New Jersey 07733-1900

P. Vijay Kumar (B:902), Communication Science Institute, Department of Electrical Engineering – Systems, University of Southern California, 3740 McClintock Avenue, EEB500, Los Angeles, California 90089-2565 and Scintera Networks, Inc., San Diego, California

Cedric F. Lam (B:514), AT&T Labs-Research, 200 Laurel Avenue South, Middletown, New Jersey 07748

Bruno Lavigne (A:732), Alcatel CIT/ Research & Innovation, Route de Nozay, F-91461 Marcoussis cedex, France

Olivier Leclerc (A:732), Alcatel Research & Innovation, Route de Nozay, F-91460 Marcoussis cedex, France

David S. Levy (A:295), Bell Laboratories, Lucent Technologies, Room H0 3B-506, 101 Crawfords Corner Road, Holmdel, New Jersey 07733-3030

Arthur J. Lowery (B:564), VPIsystems Inc., Design Center Group, 17-27 Cotham Road, Kew, Melbourne 3101, Australia

Xiaolin Lu (B:404), Morning Forest, LLC, 8804 S. Blue Mountain Place, Highlands Ranch, Colorado 80126

Hsiao-Feng Lu (B:902), Communication Science Institute, Department of Electrical Engineering – Systems, University of Southern California, 3740 McClintock Avenue, EEB500, Los Angeles, California 90089-2565

Amaresh Mahapatra (A:258), Linden Corp., 10 Northbriar Road, Acton, Massachusetts 01720

T. G. B. Mason (A:587), Agere Systems, 9999 Hamilton Blvd., Breinigsville, Pennsylvania 18031-9304

Curtis R. Menyuk (B:305), Computer Science and Electrical Engineering Department, TRC-201A, University of Maryland Baltimore County 1000 Hilltop Circle, Baltimore, Maryland 21250 and PhotonEx Corporation, 200 MetroWest Technology Park, Maynard, Massachusetts 01754

Benny Mikkelsen (B:232), Mintera Corporation, 847 Rogers Street, One Lowell Research Center, Lowell, Massachusetts 01852

John Minelly (A:80), Corning, Inc., SP-AR-02-01, Corning, New York 14831

Osamu Mizuhara (A:784), Agere Systems, Optical Systems Research, 9999 Hamilton Blvd., Breinigsville, Pennsylvania 18031

Stefan Mohrdiek (A:563), Nortel Networks Optical Components, Binzstrasse 17, CH-8045 Zürich, Switzerland

Ruo-Mei Mu (B:305), Tyco Telecommunications, 250 Industrial Way West, Eatontown, New Jersey 07724-2206

Edmond J. Murphy (A:258), JDS Uniphase, 1985 Blue Hills Avenue Ext., Windsor, Connecticut 06095

Timothy O. Murphy (A:295), Bell Laboratories, Lucent Technologies, Room H0 3D-516, 101 Crawfords Corner Road, Holmdel, New Jersey 07733-3030

Lynn E. Nelson (B:725), OFS Fitel, Holmdel, New Jersey 07733

Andrew M. Odlyzko (B:17), University of Minnesota Digital Technology Center, 1200 Washington Avenue S., Minneapolis, Minnesota 55415

Jin-Yi Pan (B:329), Sorrento Networks Inc., 9990 Mesa Rim Drive, San Diego, California 92121-2930

Sunita H. Patel (A:295), Bell Laboratories, Lucent Technologies, Room H0 3D-502, 101 Crawfords Corner Road, Holmdel, New Jersey 07733-3030

Graeme Pendock (B:198), Sycamore Networks, 10 Elizabeth Drive, Chelmsford, Massachusetts 01824-4111

Jinendra Ranka (B:198), Sycamore Networks, 10 Elizabeth Drive, Chelmsford, Massachusetts 01824-4111

Gregory Raybon (B:232), Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, New Jersey 07733

Gaylord W. Richards (A:295), Bell Laboratories, Lucent Technologies, Room 6L-219, 2000 Naperville Road, Naperville, Illinois 60566-7033

Karsten Rottwitt (A:213), Ørsted Laboratory, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, Copenhagen dk 2100, Denmark

Bertold E. Schmidt (A:563), Nortel Networks Optical Components, Binzstrasse 17, Ch-8045 Zürich, Switzerland

Oleh Sniezko (B:404), Oleh-Lightcom, Highlands Ranch, Colorado 80126

Leo H. Spiekman (A:699), Genoa Corporation, Lodewijkstraat 1A, 5652 AC Eindhoven, The Netherlands

Atul K. Srivastava (A:174), Onetta Inc., 1195 Borregas Avenue, Sunnyvale, California 94089

Andrew J. Stentz (A:213), Photuris, Inc., 20 Corporate Place South, Piscataway, New Jersey 08809

John Strand (B:57), AT&T Laboratories, Lightwave Networks Research Department, Room A5-106, 200 Laurel Avenue, Middletown, New Jersey 07748

Thomas A. Strassser (A:477), Photuris Inc., 20 Corporate Place South, Piscataway, New Jersey 08854

Yan Sun (A:174), Onetta Inc., 1195 Borregas Avenue, Sunnyvale, California 94089

Eric S. Tentarelli (A:295), Bell Laboratories, Lucent Technologies, Room H0 3B-530, 101 Crawfords Corner Road, Holmdel, New Jersey 07733-3030

Pieter H. van Heyningen (B:438), Lucent Technologies NL, P.O. Box 18, Huizen 1270AA, The Netherlands

Giorgio M. Vitetta (B:965), University of Modena and Reggio Emilia, Department of Information Engineering, Via Vignolese 905, Modena 41100, Italy

W. White (A:17), OFS Fitel, LLC, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Alan E. Willner (B:642), University of Southern California, Los Angeles, California 90089-2565

Moe Z. Win (B:902, B:965), AT&T Labs-Research, Room A5-1D01, 200 Laurel Avenue South, Middletown, New Jersey 07748-1914

Jack H. Winters (B:965), AT&T Labs-Research, Room 4-147, 100 Schulz Drive, Middletown, New Jersey 07748-1914

Martin Zirngibl (A:374), Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, New Jersey 07733-0400

John Zyskind (B:198), Sycamore Networks, 10 Elizabeth Drive, Chelmsford, Massachusetts 01824-4111

Chapter 1 | Overview

Ivan P. Kaminow

Bell Laboratories (retired), Kaminow Lightwave Technology, Holmdel, New Jersey

Introduction

Modern lightwave communications had its origin in the first demonstrations of the laser in 1960. Most of the early lightwave R&D was pursued by established telecommunications company labs (AT&T, NTT, and the British Post Office among them). By 1979, enough progress had been made in lightwave technology to warrant a book, *Optical Fiber Telecommunications (OFT)*, edited by S. E. Miller and A. G. Chynoweth, summarizing the state of the art. Two sequels have appeared: in 1988, *OFT II*, edited by S. E. Miller and I. P. Kaminow, and in 1997, *OFT III (A & B)*, edited by I. P. Kaminow and T. L. Koch. The rapid changes in the field now call for a fourth set of books. *OFT IV (A & B)*.

This chapter briefly summarizes the previous books and chronicles the remarkably changing climates associated with each period of their publication. The main purpose, however, is to summarize the chapters in *OFT IV* in order to give the reader an overview.

History

While many excellent books on lightwave communications have been published, this series has developed a special character, with a reputation for comprehensiveness and authority, because of its unique history. *Optical Fiber Telecommunications* was published in 1979, at the dawn of the revolution in lightwave telecommunications. It was a stand-alone work that aimed to collect all available information on lightwave research. Miller was Director of the Lightwave Systems Research Laboratory and, together with Rudi Kompfner, the Associate Executive Director, guided the system research at the Crawford Hill Laboratory of AT&T Bell Laboratories; Chynoweth was an Executive Director in the Murray Hill Laboratory, leading the optical fiber research. Many groups were active at other laboratories in the United States, Europe, and Japan. *OFT*, however, was written exclusively by Bell Laboratories authors, who nevertheless aimed to incorporate global results.

OPTICAL FIBER TELECOMMUNICATIONS.
VOLUME IVA

ISBN: 0-12-395172-0

Miller and Chynoweth had little trouble finding suitable chapter authors at Bell Labs to cover practically all the relevant aspects of the field at that time.

Looking back at that volume, it is interesting that the topics selected are still quite basic. Most of the chapters cover the theory, materials, measurement techniques, and properties of fibers and cables (for the most part, multimode fibers). Only one chapter covers optical sources, mainly multimode AlGaAs lasers operating in the 800- to 900-nm band. The remaining chapters cover direct and external modulation techniques, photodetectors and receiver design, and system design and applications. Still, the basic elements of the present day systems are discussed: low-loss vapor-phase silica fiber and double-heterostructure lasers.

Although system trials were initiated around 1979, it required several more years before a commercially attractive lightwave telecommunications system was installed in the United States. The AT&T Northeast Corridor System, operating between New York and Washington, DC, began service in January 1983, operating at a wavelength of 820 nm and a bit rate of 45 Mb/s in multimode fiber. Lightwave systems were upgraded in 1984 to 1310 nm and 417 or 560 Mb/s in single-mode fiber in the United States as well as in Europe and Japan.

The year 1984 also saw the Bell System broken up by the court-imposed "Modified Final Judgment" that separated the Bell operating companies into seven regional companies and left AT&T as the long distance carrier as well as a telephone equipment vendor. Bell Laboratories remained with AT&T, and Bellcore was formed to serve as the R&D lab for all seven regional Bell operating companies (RBOCs). The breakup spurred a rise in diversity and competition in the communications business. The combination of technical advances in computers and communications, growing government deregulation, and apparent new business opportunities all served to raise expectations.

Tremendous technical progress was made during the next few years, and the choice of lightwave over copper coaxial cable or microwave relay for most long-haul transmission systems was assured. The goal of research was to improve performance, such as bitrate and repeater spacing, and to find other applications beyond point-to-point long haul telephone transmission. A completely new book, *Optical Fiber Telecommunications II*, was published in 1988 to summarize the lightwave R&D advances at the time. To broaden the coverage, non-Bell Laboratories authors from Bellcore (now Telcordia), Corning, Nippon Electric Corporation, and several universities were represented among the contributors. Although research results are described in *OFT II*, the emphasis is much stronger on commercial applications than in the previous volume.

The initial chapters of *OFT II* cover fibers, cables, and connectors, dealing with both single- and multimode fiber. Topics include vapor-phase methods for fabricating low-loss fiber operating at 1310 and 1550 nm, understanding chromatic dispersion and nonlinear effects, and designing polarization-maintaining fiber. Another large group of chapters deals with

a range of systems for loop, intercity, interoffice, and undersea applications. A research-oriented chapter deals with coherent systems and another with possible local area network designs, including a comparison of time-division multiplexing (TDM) and wavelength division multiplexing (WDM) to efficiently utilize the fiber bandwidth. Several chapters cover practical subsystem components, such as receivers and transmitters and their reliability. Other chapters cover photonic devices, such as lasers, photodiodes, modulators, and integrated electronic and integrated optic circuits that make up the subsystems. In particular, epitaxial growth methods for InGaAsP materials suitable for 1310 and 1550 nm applications and the design of high-speed single-mode lasers are discussed in these chapters.

By 1995, it was clear that the time had arrived to plan for a new volume to address recent research advances and the maturing of lightwave systems. The contrast with the research and business climates of 1979 was dramatic. Sophisticated system experiments were being performed utilizing the commercial and research components developed for a proven multibillion-dollar global lightwave industry. For example, 10,000-km lengths of high-performance fiber were assembled in several laboratories around the world for nonreturn-to-zero (NRZ), soliton, and WDM transmission demonstrations. Worldwide regulatory relief stimulated the competition in both the service and hardware ends of the telecommunications business. The success in the long-haul market and the availability of relatively inexpensive components led to a wider quest for other lightwave applications in cable television and local access network markets. The development of the diode-pumped, erbium-doped fiber amplifier (EDFA) played a crucial role in enhancing the feasibility and performance of long-distance and WDM applications. By the time of publication of *OFT III* in 1997, incumbent telephone companies no longer dominated the industry. New companies were offering components and systems and other startups were providing regional, exchange, and Internet services.

In 1996, AT&T voluntarily separated its long distance service and telephone equipment businesses to better meet the competition. The former kept the AT&T name, and the latter took on the name Lucent Technologies. Bell Labs remained with Lucent, and AT&T Labs was formed. Bellcore was put up for sale, as the consolidating and competing RBOCs found they did not need a joint lab.

Because of a wealth of new information, *OFT III* was divided into two books, *A* and *B*, covering systems and components, respectively. Many topics of the previous volumes, such as fibers, cables, and laser sources, are updated. But a much larger list of topics covers fields not previously included. In *A*, for example, transceiver design, EDFAs, laser sources, optical fiber components, planar (silica on silicon) integrated circuits, lithium niobate devices, and photonic switching are reviewed. And in *B*, SONET (synchronous optical network) standards, fiber and cable design, fiber nonlinearities, polarization effects, solitons, terrestrial and undersea systems, high bitrate transmission, analog

cable systems, passive optical networks (PONs), and multiaccess networks are covered.

Throughout the two books, erbium amplifiers and WDM are common themes. It is difficult to overstate the impact these two technologies have had in both generating and supporting the telecommunications revolution that coincided with their commercial introduction. The EDFA was first reported in about 1987 by researchers at Southampton University in the UK and at AT&T Bell Labs. In 1990, driven by the prospect of vast savings offered by WDM transmission using EDFAs, Bell Labs began to develop long-haul WDM systems. By 1996, AT&T and Alcatel had installed the first transatlantic cable with an EDFA chain and a single 5 Gb/s optical channel. AT&T installed the first commercial terrestrial WDM system employing EDFAs in 1995. Massive deployment of WDM worldwide soon followed. WDM has made the exponential traffic growth spurred by the coincident introduction of the Internet browser economically feasible. If increased TDM bitrates and multiple fibers were the only alternative, the enthusiastic users and investors in the Internet would have been priced out of the market.

Optical Fiber Telecommunications IV

BACKGROUND

There was considerable excitement in the lightwave research community during the 1970s and early 1980s as wonderful new ideas emerged at a rapid pace. The monopoly telephone system providers, however, were less enthusiastic. They were accustomed to moving at their own deliberate pace, designing equipment to install in their own systems, which were expected to have a long economic life. The long-range planners projected annual telephone voice traffic growth in the United States at about 5–10%, based on population and business growth.

Recent years, on the other hand, have seen mind-numbing changes in the communication business—especially for people brought up in the telephone environment. The Internet browser spawned a tremendous growth in data traffic, which in turn encouraged visions of tremendous revenue growth. Meanwhile, advances in WDM technology and its wide deployment synergistically supported the Internet traffic and enthusiasm. As a result, entrepreneurs invested billions of dollars in many companies vying for the same slice of pie. The frenzy reached a peak in the spring of 2000 and then rapidly melted down as investors realized that the increased network capacity had already outstripped demand. As of October 2001, the lightwave community is waiting for a recovery from the current industry collapse.

Nevertheless, the technical advances achieved during these last five years will continue to impact telecommunications for years to come. Thus, we are proud to present a comprehensive and forward-looking account of these accomplishments.

Survey of *OFT IV A* and *B*

Advances in optical network architectures have followed component innovations. For example, the low loss fiber and double heterostructure laser enabled the first lightwave system generation; and the EDFA has enabled the WDM generation. Novel components (such as tunable lasers, MEMS switches, and planar waveguide devices) are making possible more sophisticated optical networks. At the same time, practical network implementations uncover the need for added device functionality and very low cost points. For example, 40 Gb/s systems need dynamic dispersion and PMD compensation to overcome system impairments.

We have divided *OFT IV* into two books: book *A* comprises the component chapters and book *B* the system and system impairment chapters.

BOOK A: COMPONENTS

Design of Optical Fibers for Communications Systems (Chapter 2)

Optical fiber has been a key element in each of the previous volumes of *OFT*. The present chapter by DiGiovanni, Boncek, Golowich, Das, Blyler, and White reflects a maturation of the field: fiber performance must now go beyond simple low attenuation and must exhibit critical characteristics to support the high speeds and long routes on terrestrial and undersea systems. At the same time, fiber for the metropolitan and access markets must meet demanding price points.

The chapter reviews the design criteria for a variety of fibers of current commercial interest. For the traditional long-haul market, impairments such as dispersion slope and polarization mode dispersion (PMD) that were negligible in earlier systems are now limiting factors. If improved fiber design is unable to overcome these limits, new components will be required to solve the problem. These issues are addressed again from different points of view in later systems and components chapters in *OFT IV A* and *B*.

The present chapter also reviews a variety of new low-cost fiber designs for emerging metropolitan and access markets. Further down the network chain, the design of multimode glass and plastic fiber for the highly cost-sensitive local area network market are also explored. Finally, current research on hollow core and photonic bandgap fiber structures is summarized.

New Materials for Optical Amplifiers (Chapter 3)

In addition to transport, fiber plays an important role as an amplifying medium. Aluminum-doped silica has been the only important commercial host and erbium the major amplifying dopant. Happily, erbium is soluble in Al-silica and provides gain at the attenuation minimum for silica transmission fiber. Still, researchers are exploring other means for satisfying demands

for wider bandwidth in the 1550 nm region as well as in bands that might be supported by other rare-earth ions, which have low efficiency in silica hosts.

Ellison and Minelly review research on new fiber materials, including fluorides, alumina-doped silica, antimony silicates, and tellurite. They also report on extended band erbium-doped fiber amplifiers (EDFAs), thulium-doped fiber amplifiers, and 980 nm ytterbium fiber lasers for pumping EDFAs.

Advances in Erbium-Doped Fiber Amplifiers (Chapter 4)

The development of practical EDFAs has ushered in a generation of dense WDM (DWDM) optical networks. These systems go beyond single frequency or even multifrequency point-to-point links to dynamic networks that can be reconfigured by add/drop multiplexers or optical cross-connects to meet varying system demands. Such networks place new requirements on the EDFAs: they must maintain flatness over many links, and they must recover from sudden drops or adds of channels. And economics drives designs that provide more channels and denser spacing of channels.

Srivastava and Sun summarize recent advances in EDFA design and means for coping with the challenges mentioned above. In particular, they treat long wave L-band amplifiers, which have more than doubled the conventional C-band to 84 nm. They also treat combinations of EDFA and Raman amplification, and dynamic control of gain flatness.

Raman Amplification in Lightwave Communication Systems (Chapter 5)

Raman amplification in fibers has been an intellectual curiosity for nearly 30 years; the large pump powers and long lengths required made Raman amplifiers seem impractical. The advent of the EDFA appeared to drive a stake into the heart of Raman amplifiers. Now, however, Raman amplifiers are rising along with the needs of submarine and ultralong-haul systems. More powerful practical diode pumps have become available; and the ability to provide gain at any wavelength and with low effective noise figure is now recognized as essential for these systems.

Rottwitt and Stentz review the advances in distributed and lumped Raman amplifiers with emphasis on noise performance and recent system experiments.

Electrooptic Modulators (Chapter 6)

Modulators put the payload on the optical carrier and have been a focus of attention from the beginning. Direct modulation of the laser current is often the cheapest solution where laser linewidth and chirp are not important. However, for high performance systems, external modulators are needed. Modulators based on the electrooptic effect have proven most versatile in meeting performance requirements, although cost may be a constraint.

Titanium-diffused lithium niobate has been the natural choice of material, in that no commercial substitutes have emerged in nearly 30 years. However, integrated semiconductor electroabsorption modulators are now offering strong competition on the cost and performance fronts.

Mahapatra and Murphy briefly compare electroabsorption-modulated lasers (EMLs) and electrooptic modulators. They then focus on titanium-diffused lithium niobate modulators for lightwave systems. They cover fabrication methods, component design, system requirements, and modulator performance. Mach-Zehnder modulators are capable of speeds in excess of 40 Gb/s and have the ability to control chirp from positive through zero to negative values for various system requirements. Finally, the authors survey research on polymer electrooptic modulators, which offer the prospect of lower cost and novel uses.

Optical Switching in Transport Networks: Applications, Requirements, Architectures, Technologies, and Solutions (Chapter 7)

Early DWDM optical line systems provided simple point-to-point links between electronic end terminals without allowing access to the intermediate wavelength channels. Today's systems carry over 100 channels per fiber and new technologies allow intermediate routing of wavelengths at add/drop multiplexers and optical cross-connects. These new capabilities allow "optical layer networking," an architecture with great flexibility and intelligence.

Al-Salameh, Korotky, Levy, Murphy, Patel, Richards, and Tentarelli explore the use of optical switching in modern networking architectures. After reviewing principles of networking, they consider in detail various aspects of the topic. The performance and requirements for an optical cross connect (OXC) for opaque (with an electronic interface and/or electronic switch fabric) and transparent (all-optical) technologies are compared. Also, the applications of the OXC in areas such as provisioning, protection, and restoration are reviewed. Note that an OXC has all-optical ports but may have internal electronics at the interfaces and switch fabric.

Finally, several demonstration OXCs are studied, including small optical switch fabrics, wavelength-selective OXCs, and large strictly nonblocking cross connects employing microelectromechanical system (MEMS) technology. These switches are expected to be needed soon at core network nodes with 1000×1000 ports.

Applications for Optical Switch Fabrics (Chapter 8)

Whereas the previous chapter looked at OXCs from the point of view of the network designer, Zirngibl focuses on the physical design of OXCs with capacities greater than 1 Tb/s. He considers various design options including MEMS switch fabrics, transparent and opaque variants, and nonwavelength-blocking

configurations. He finds that transport in the backplane for very large capacity (bitrate × port number) requires optics in the interconnects and switch fabric.

He goes beyond the cross-connect application, which is a slowly reconfigurable circuit switch, to consider the possibility of a high-capacity packet switch, which, although schematically similar to an OXC, must switch in times short relative to a packet length. Again the backplane problem dictates an optical fabric and interconnects. He proposes tunable lasers in conjunction with a waveguide grating router as the fast optical switch fabric.

Planar Lightwave Devices for WDM (Chapter 9)

The notion of integrated optical circuits, in analogy with integrated electronic circuits, has been in the air for over 30 years, but the vision of large-scale integration has never materialized. Nevertheless, the concept of small-scale planar waveguide circuits has paid off handsomely. Optical waveguiding provides efficient interactions in lasers and modulators, and novel functionality in waveguide grating routers and Bragg gratings. These elements are often linked together with waveguides.

Doerr updates recent progress in the design of planar waveguides, starting with waveguide propagation analysis and the design of the star coupler and waveguide grating router (or arrayed waveguide grating). He goes on to describe a large number of innovative planar devices such as the dynamic gain equalizer, wavelength selective cross connect, wavelength add/drop, dynamic dispersion compensator, and the multifrequency laser. Finally, he compares various waveguide materials: silica, lithium niobate, semiconductor, and polymer.

Fiber Grating Devices in High-Performance Optical Communication Systems (Chapter 10)

The fiber Bragg grating is ideally suited to lightwave systems because of the ease of integrating it into the fiber structure. The technology for economically fabricating gratings has developed over a relatively short period, and these devices have found a number of applications to which they are uniquely suited. For example, they are used to stabilize lasers, to provide gain flattening in EDFAs, and to separate closely spaced WDM channels in add/drops.

Strasser and Erdogan review the materials aspects of the major approaches to fiber grating fabrication. Then they treat the properties of fiber gratings analytically. Finally, they review the device properties and applications of fiber gratings.

Pump Laser Diodes (Chapter 11)

Although EDFAs were known as early as 1986, it was not until a high-power 1480 nm semiconductor pump laser was demonstrated that people took notice.

Earlier, expensive and bulky argon ion lasers provided the pump power. Later, 980 nm pump lasers were shown to be effective. Recent interest in Raman amplifiers has also generated a new interest in 1400 nm pumps. Ironically, the first 1480 nm pump diode that gave life to EDFAs was developed for a Raman amplifier application.

Schmidt, Mohrdiek, and Harder review the design and performance of 980 and 1480 nm pump lasers. They go on to compare devices at the two wavelengths, and discuss pump reliability and diode packaging.

Telecommunication Lasers (Chapter 12)

Semiconductor diode lasers have undergone years of refinement to satisfy the demands of a wide range of telecommunication systems. Long-haul terrestrial and undersea systems demand reliability, speed, and low chirp; short-reach systems demand low cost; and analog cable TV systems demand high power and linearity.

Ackerman, Eng, Johnson, Ketelsen, Kiely, and Mason survey the design and performance of these and other lasers. They also discuss electroabsorption modulated lasers (EMLs) at speeds up to 40 Gb/s and a wide variety of tunable lasers.

VCSELs for Metro Communications (Chapter 13)

Vertical cavity surface emitting lasers (VCSELs) are employed as low-cost sources in local area networks at 850 nm. Their cost advantage stems from the ease of coupling to fiber and the ability to do wafer-scale testing to eliminate bad devices. Recent advances have permitted the design of efficient long wavelength diodes in the 1300–1600 nm range.

Chang-Hasnain describes the design of VCSELs in the 1310 and 1550 nm bands for application in the metropolitan market, where cost is key. She also describes tunable designs that promise to reduce the cost of sparing lasers.

Semiconductor Optical Amplifiers (Chapter 14)

The semiconductor gain element has been known from the beginning, but it was fraught with difficulties as a practical transmission line amplifier: it was difficult to reduce reflections, and its short time constant led to unacceptable nonlinear effects. The advent of the EDFA practically wiped out interest in the semiconductor optical amplifier (SOA) as a gain element. However, new applications based on its fast response time have revived interest in SOAs.

Spiekman reviews recent work to overcome the limitations on SOAs for amplification in single-frequency and WDM systems. The applications of main interest, however, are in optical signal processing, where SOAs are used in wavelength conversion, optical time division multiplexing, optical phase

conjugation, and all-optical regeneration. The latter topic is covered in detail in the following chapter.

All-Optical Regeneration: Principles and WDM Implementation (Chapter 15)

A basic component in long-haul lightwave systems is the electronic regenerator. It has three functions: reamplifying, reshaping, and retiming the optical pulses. The EDFA is a 1R regenerator; regenerators without retiming are 2R; but a full-scale repeater is a 3R regenerator. A separate 3R electronic regenerator is required for each WDM channel after a fixed system span. As the bitrate increases, these regenerators become more expensive and physically more difficult to realize. The goal of ultralong-haul systems is to eliminate or minimize the need for electronic regenerators (see Chapter 5 in Volume B).

Leclerc, Lavigne, Chiaroni, and Desurvire describe another approach, the all-optical 3R regenerator. They describe a variety of techniques that have been demonstrated for both single channel and WDM regenerators. They argue that at some bitrates, say 40 Gb/s, the optical and electronic alternatives may be equally difficult and expensive to realize, but at higher rates the all-optical version may dominate.

High Bitrate Transmitters, Receivers, and Electronics (Chapter 16)

In high-speed lightwave systems, the optical components usually steal the spotlight. However, the high bitrate electronics in the terminals are often the limiting components.

Kasper, Mizuhara, and Chen review the design of practical high bitrate (10 and 40 Gb/s) receivers, transmitters, and electronic circuits in three separate sections. The first section reviews the performance of various detectors, analyzes receiver sensitivity, and considers system impairments. The second section covers directly and externally modulated transmitters and modulation formats like return-to-zero (RZ) and chirped RZ (CRZ). The final section covers the electronic circuit elements found in the transmitters and receivers, including broadband amplifiers, clock and data recovery circuits, and multiplexers.

BOOK B: SYSTEMS AND IMPAIRMENTS

Growth of the Internet (Chapter 2)

The explosion in the telecommunications marketplace is usually attributed to the exponential growth of the Internet, which began its rise with the introduction of the Netscape browser in 1996. Voice traffic continues to grow steadily, but data traffic is said to have already matched or overtaken it. A lot of self-serving myth and hyperbole surround these fuzzy statistics. Certainly claims of doubling data traffic every three months helped to sustain the market frenzy.

On the other hand, the fact that revenues from voice traffic still far exceed revenues from data was not widely circulated.

Coffman and Odlyzko have been studying the actual growth of Internet traffic for several years by gathering quantitative data from service providers and other reliable sources. The availability of data has been shrinking as the Internet has become more commercial and fragmented. Still, they find that, while there may have been early bursts of three-month doubling, the overall sustained rate is an annual doubling. An annual doubling is a very powerful growth rate; and, if it continues, it will not be long before the demand catches up with the network capacity. Yet, with prices dropping at a comparable rate, faster traffic growth may be required for strong revenue growth.

Optical Network Architecture Evolution (Chapter 3)

The telephone network architecture has evolved over more than a century to provide highly reliable voice connections to a global network of hundreds of millions of telephones served by different providers. Data networks, on the other hand, have developed in a more ad hoc fashion with the goal of connecting a few terminals with a range of needs at the lowest price in the shortest time. Reliability, while important, is not the prime concern.

Strand gives a tutorial review of the Optical Transport Network employed by telephone service providers for intercity applications. He discusses the techniques used to satisfy the traditional requirements for reliability, restoration, and interoperability. He includes a refresher on SONET (SDH). He discusses architectural changes brought on by optical fiber in the physical layer and the use of optical layer cross connects. Topics include all-optical domains, protection switching, rings, the transport control plane, and business trends.

Undersea Communication Systems (Chapter 4)

The oceans provide a unique environment for long-haul communication systems. Unlike terrestrial systems, each design starts with a clean slate; there are no legacy cables, repeater huts, or rights-of-way in place and few international standards to limit the design. Moreover, there are extreme economic constraints and technological challenges. For these reasons, submarine systems designers have been the first to risk adopting new and untried technologies, leading the way for the terrestrial ultralong-haul system designers (see Chapter 5).

Following a brief historical introduction, Bergano gives a tutorial review of some of the technologies that promise to enable capacities of 2 Tb/s on a single fiber over transoceanic spans. The technologies include the chirped RZ (CRZ) modulation format, which is compared briefly with NRZ, RZ, and dispersion-managed solitons (see Chapters 5, 6, and 7 for more on this topic). He also discusses measures of system performance (the Q-factor), forward

error correcting (FEC) codes (see Chapters 5 and 17), long-haul system design, and future trends.

High Capacity, Ultralong-Haul Transmission (Chapter 5)

The major hardware expense for long-haul terrestrial systems is in electronic terminals, repeaters, and line cards. Since WDM systems permit traffic with various destinations to be bundled on individual wavelengths, great savings can be realized if the unrepeatered reach can be extended to 2000–5000 km, allowing traffic to pass through nodes without optical-to-electrical (O/E) conversion. As noted in connection with Chapter 4, some of the technology pioneered in undersea systems can be adapted in terrestrial systems but with the added complexities of legacy systems and standards. On the other hand, the terrestrial systems can add the flexibility of optical networking by employing optical routing in add/drops and OXCs (see Chapters 7 and 8) at intermediate points.

Zyskind, Barry, Pendock, and Cahill review the technologies needed to design ultralong-haul (ULH) systems. The technologies include EDFAs and distributed Raman amplification, novel modulation formats, FEC, and gain flattening. They also treat transmission impairments (see later chapters in this book) such as the characteristics of fibers and compensators needed to deal with chromatic dispersion and PMD. Finally, they discuss the advantages of optical networking in the efficient distribution of data using IP (Internet Protocol) directly on wavelengths with meshes rather than SONET rings.

Pseudo-Linear Transmission of High-Speed TDM Signals: 40 and 160 Gb/s (Chapter 6)

A reduction in the cost and complexity of electronic and optoelectronic components can be realized by an increase in channel bitrate, as well as by the ULH techniques mentioned in Chapter 5. The higher bitrates, 40 and 160 Gb/s, present their own challenges, among them the fact that the required energy per bit leads to power levels that produce nonlinear pulse distortions. Newly discovered techniques of pseudo-linear transmission offer a means for dealing with the problem. They involve a complex optimization of modulation format, dispersion mapping, and nonlinearity. Pseudo-linear transmission occupies a space somewhere between dispersion-mapped linear transmission and nonlinear soliton transmission (see Chapter 7).

Essiambre, Raybon, and Mikkelsen first present an extensive analysis of pseudo-linear transmission and then review TDM transmission experiments at 40 and 160 Gb/s.

Dispersion Managed Solitons and Chirped RZ: What Is the Difference? (Chapter 7)

Menyuk, Carter, Kath, and Mu trace the evolution of soliton transmission to its present incarnation as Dispersion Managed Soliton (DMS) transmission

and the evolution of NRZ transmission to its present incarnation as CRZ transmission. Both approaches depend on an optimization of modulation format, dispersion mapping, and nonlinearity, defined as pseudo-linear transmission in Chapter 6 and here as "quasi-linear" transmission. The authors show how both DMS and CRZ exhibit aspects of linear transmission despite their dependence on the nonlinear Kerr effect. Remarkably, they argue that, despite widely disparate starting points and independent reasoning, the two approaches unwittingly converge in the same place.

Still, on their way to convergence, DMS and CRZ pulses exhibit different characteristics that suit them to different applications: For example, CRZ produces pulses that merge in transit along a wide undersea span and reform only at the receiver ashore, while DMS produces pulses that reform periodically, thereby permitting access at intermediate add/drops.

Metropolitan Optical Networks (Chapter 8)

For many years the long-haul domain has been the happy hunting ground for lightwave systems, since the cost of expensive hardware can be shared among many users. Now that component costs are moderating, the focus is on the metropolitan domain where costs cannot be spread as widely. Metropolitan regions generally span ranges of 10 to 100 km and provide the interface with access networks (see Chapters 9, 10, and 11). SONET/SDH rings, installed to serve voice traffic, dominate metropolitan networks today.

Ghani, Pan, and Chen trace the developing access users, such as Internet service providers, local area networks, and storage area networks. They discuss a number of WDM metropolitan applications to better serve them, based on optical networking via optical rings, optical add/drops, and OXCs. They also consider IP over wavelengths to replace SONET. Finally, they discuss possible economical migration paths from the present architecture to the optical metropolitan networks.

The Evolution of Cable TV Networks (Chapter 9)

Coaxial analog cable TV networks were substantially upgraded in the 1990s by the introduction of linear lasers and single-mode fiber. Hybrid Fiber Coax (HFC) systems were able to deliver in excess of 80 channels of analog video plus a wide band suitable for digital broadcast and interactive services over a distance of 60 km. Currently high-speed Internet access and voice-over-IP telephony have become available, making HFC part of the telecommunications access network.

Lu and Sniezko outline past, present, and future HFC architectures. In particular, the mini fiber node (mFN) architecture provides added capacity for two-way digital as well as analog broadcast services. They consider a number of mFN variants based on advances in RF, lightwave, and DSP (digital signal processor) technologies that promise to provide better performance at lower cost.

Optical Access Networks (Chapter 10)

The access portion of the telephone network, connecting the central office to the residence, is called the "loop." By 1990 half the new loops in the United States were served by digital loop carrier (DLC), a fiber several miles long from the central office to a remote terminal in a neighborhood that connects to about 100 homes with analog signals over twisted pairs. Despite much anticipation, fiber hasn't gotten much closer to residences since. The reason is that none of the approaches proposed so far is competitive with existing technology for the applications people will buy.

Harstead and van Heyningen survey numerous proposals for Fiber-in-the-Loop (FITL) and Fiber-to-the-X (FTTX), where X = Curb, Home, Desktop, etc. They consider the applications and costs of these systems. Considerable creativity and thought have been devoted to fiber in the access network, but the economics still do not work because the costs cannot be divided among a sufficient number of users. An access technology that is successful is Digital Subscriber Line (DSL) for providing high-speed Internet over twisted pairs in the loop. DSL is reviewed in an Appendix.

Beyond Gigabit: Development and Application of High-Speed Ethernet Technology (Chapter 11)

Ethernet is a simple protocol for sharing a local area network (LAN). Most of the data on the Internet start as Ethernet packets generated by desktop computers and system servers. Because of their ubiquity, Ethernet line cards are cheap and easy to install. Many people now see Ethernet as the universal protocol for optical packet networks. Its speed has already increased to 1000 Mb/s, and 10 Gb/s is on the way.

Lam describes the Ethernet system in detail from protocols to hardware, including 10 Gb/s Ethernet. He shows applications in LANs, campus, metropolitan, and long distance networks.

Photonic Simulation Tools (Chapter 12)

In the old days, new devices or systems were sketched on a pad, a prototype was put together in the lab, and its performance tested. In the present climate, physical complexity and the expense and time required rule out this brute-force approach, at least in the early design phase. Instead, individual groups have developed their own computer simulators to test numerous variations in a short time with little laboratory expense. Now, several commercial vendors offer general-purpose simulators for optical device and system development.

Lowery relates the history of lightwave simulators and explains how they work and what they can do. The user operates from a graphic user interface (GUI) to select elements from a library and combine them. The simulated device or system can then be run and measured as in the lab to determine

attributes like the eye-diagram or bit-error-rate. In the end, a physical prototype is required because of limits on computation speed among other reasons.

THE PRECEDING CHAPTERS HAVE DEALT WITH SYSTEM DESIGN; THE REMAINING CHAPTERS DEAL WITH SYSTEM IMPAIRMENTS AND METHODS FOR MITIGATING THEM

Nonlinear Optical Effects in WDM Systems (Chapter 13)

Nonlinear effects have been mentioned in different contexts in several of the earlier system chapters. The Kerr effect is an intrinsic property of glass that causes a change in refractive index proportional to the optical power.

Bayvel and Killey give a comprehensive review of intensity-dependent behavior based on the Kerr effect. They cover such topics as self-phase modulation, cross-phase modulation, four-wave mixing, and distortions in NRZ and RZ systems.

Fixed and Tunable Management of Fiber Chromatic Dispersion (Chapter 14)

Chromatic dispersion is a linear effect and as such can be compensated by adding the complementary dispersion before any significant nonlinearities intervene. Nonlinearities do intervene in many of the systems previously discussed so that periodic dispersion mapping is required to manage them.

Willner and Hoanca present a thorough taxonomy of techniques for compensating dispersion in transmission fiber. They cover fixed compensation by fibers and gratings, as well as tunable compensation by gratings and other novel devices. They also catalog the reasons for incorporating dynamic as well as fixed compensation in systems.

Polarization Mode Dispersion (Chapter 15)

Polarization mode dispersion (PMD), like chromatic dispersion, is a linear effect that can be compensated in principle. However, fluctuations in the polarization mode and fiber birefringence produced by the environment lead to a dispersion that varies statistically with time and frequency. The statistical nature makes PMD difficult to measure and compensate for. Nevertheless, it is an impairment that can kill a system, particularly when the bitrate is large (>10 Gb/s) or the fiber has poor PMD performance.

Nelson, Jopson, and Kogelnik offer an exhaustive survey of PMD covering the basic concepts, measurement techniques, PMD measurement, PMD statistics for first- and higher orders, PMD simulation and emulation, system impairments, and mitigation methods. Both optical and electrical PMD compensation (see Chapter 18) are considered.

Bandwidth Efficient Formats for Digital Fiber Transmission Systems (Chapter 16)

Early lightwave systems employed NRZ modulation; newer long-haul systems are using RZ and chirped RZ to obtain better performance. One goal of system designers is to increase spectral efficiency by reducing the RF spectrum required to transmit a given bitrate.

Conradi examines a number of modulation formats well known to radio engineers to see if lightwave systems might benefit from their application. He reviews the theory and DWDM experiments for such formats as M-ary ASK, duo-binary, and optical single-sideband. He also examines RZ formats combined with various types of phase modulation, some of which are related to discussions of CRZ in the previous Chapters 4–7.

Error-Control Coding Techniques and Applications (Chapter 17)

Error-correcting codes are widely used in electronics, e.g., in compact disc players, to radically improve system performance at modest cost. Similar forward error correcting codes (FEC) are used in undersea systems (see Chapter 4) and are planned for ULH systems (Chapter 5).

Win, Georghiades, Kumar, and Lu give a tutorial introduction to coding theory and discuss its application to lightwave systems. They conclude with a critical survey of recent literature on FEC applications in lightwave systems, where FEC provides substantial system gains.

Equalization Techniques for Mitigating Transmission Impairments (Chapter 18)

Chapters 14 and 15 describe optical means for compensating the linear impairments caused by chromatic dispersion and PMD. Chapters 16 and 17 describe two electronic means for reducing errors by novel modulation formats and by FEC. This chapter discusses a third electronic means for improving performance using equalizer circuits in the receiving terminal, which in principle can be added to upgrade an existing system. Equalization is widely used in telephony and other electronic applications. It is now on the verge of application in lightwave systems.

Win, Vitetta, and Winters point out the challenges encountered in lightwave applications and survey the mathematical techniques that can be employed to mitigate many of the impairments mentioned in previous chapters. They also describe some of the recent experimental implementations of equalizers. Additional discussion of PMD equalizers can be found in Chapter 15.

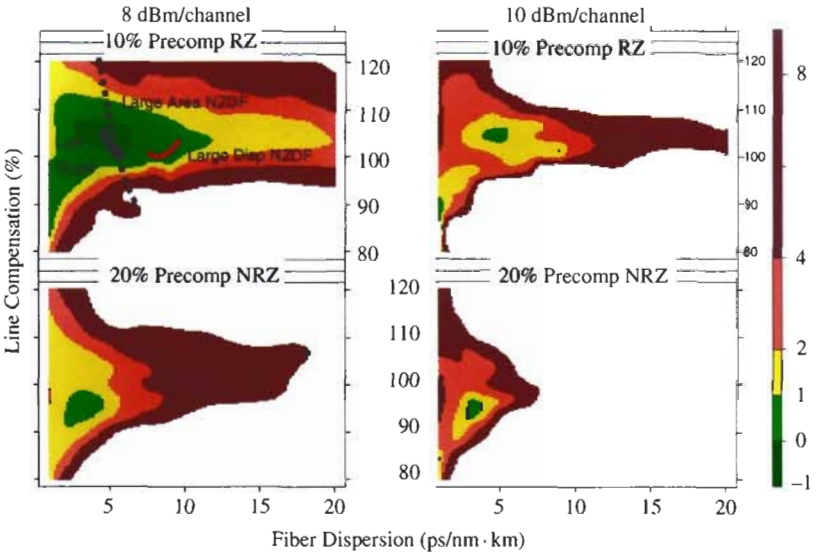


Plate 1 Simulation results of power penalty contours (indicated in dB by bar on right) for percent per-span compensation versus fiber dispersion for a 40 Gbps DWDM system for two-channel power and precompensation levels. 5 channels, 200 GHz spaced NRZ, 6 × 80 km system. Curves indicate dispersion and compensation ratios across C-band channels (from 1530 to 1565 nm) using a high-slope DCF for large-area (dashed), low-slope (solid), and large-dispersion (solid orange) NZDF. Reprinted with permission from Ref. 24.

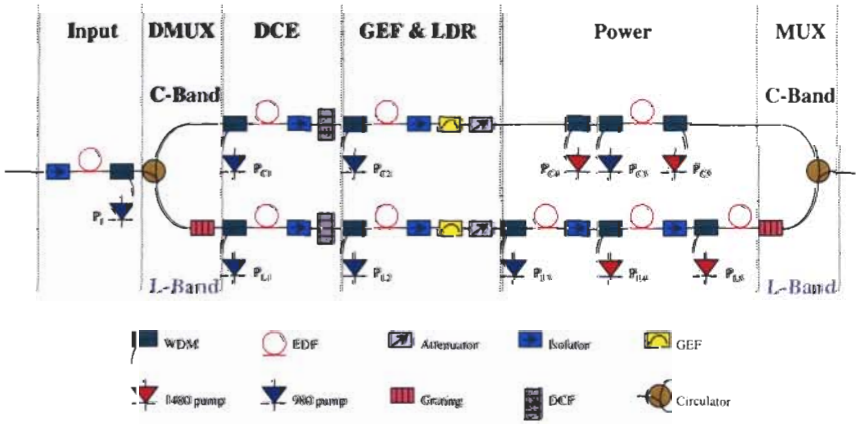


Plate 2 Schematic of ultrawideband amplifier.

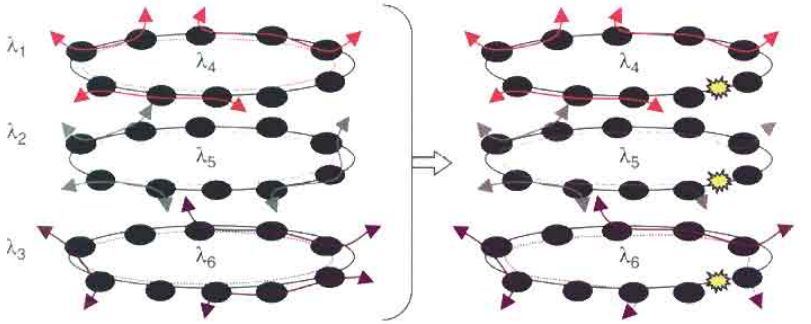


Plate 3 Left: A DWDM fiber optic network utilizing shared optical protection. Each node contains MUX/DEMUX equipment that supports service channel communication over a WDM channel. Each WDM service channel is reused around the ring and has a dedicated WDM protection channel associated with it. Right: Reconfiguration of the ring after the protection event occurs. Note that the interrupted channels (*green and blue*) are re-routed to the corresponding protection channel (*green and blue*, dashed lines). The other channel (*pink*) is not affected by the event.

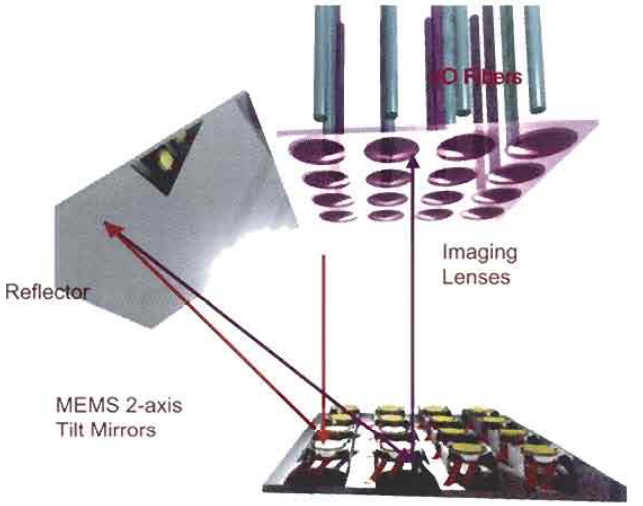


Plate 4 A picture of a 3D MEMS optical cross-connect. Light is collimated across the switch fabric via a lens array. The use of the reflector reduces the package size and the number of chips needed to perform the switching function.

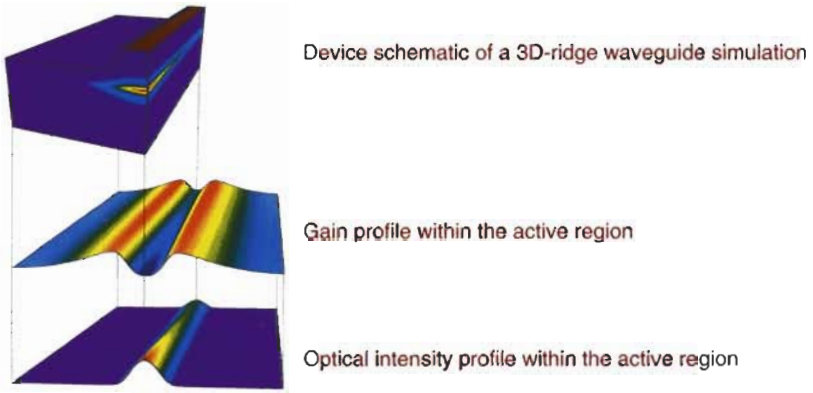


Plate 5 Three-dimensional simulation of a ridge waveguide laser structure incorporating an asymmetric facet coating. Together with the intensity profile along the active region, the gain profile influenced by lateral and longitudinal spatial hole burning is shown.

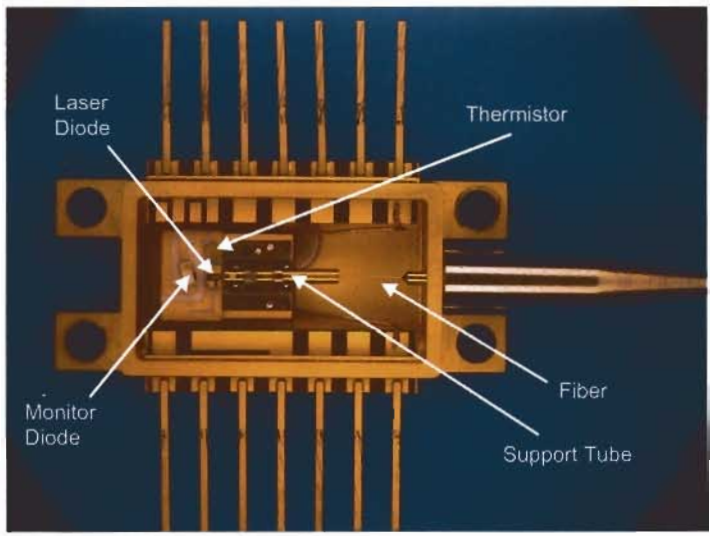


Plate 6 Butterfly package with housing area dimensions (without pins) of $13 \times 30\,\text{mm}^2$. (Courtesy of Nortel Networks Optical Components, Zürich, Switzerland.)

Chapter 2 | Design of Optical Fibers for Communications Systems

David J. DiGiovanni, Santanu K. Das, Lee L. Blyler, W. White, and Raymond K. Boncek

OFS Fitel, LLC, Murray Hill, New Jersey

Steven E. Golowich

Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey

1. Introduction

The optical communications industry has seen phenomenal growth over the last few years, spurring a significant commercial market in optical components and systems. This growth has extended across all application spaces, from transoceanic and transcontinental distances to regional networks to campus and building wiring. The explosion in demand for bandwidth has been fueled by the impending ubiquity of the Internet as more information is handled electronically, as more homes go online, and as more business is transacted over the web. The implication of this growth, however, goes beyond simply increasing the amount of information that can be transmitted between two points. Transport of data over the Internet presents fundamentally different traffic patterns than voice traffic, which dominated telecommunications until around 1998. Voice traffic typically remains within the local or metropolitan calling region where it was generated. In addition, since voice bandwidth requires a data rate of only 64 kbps, terabit transmission over long distances was thought unnecessary.

Now, data generated on the Internet takes many formats, such as audio or video clips and large computer files. This type of data is just as likely to travel ultralong distances[1] as to be dropped locally. Thus, the need for high-capacity transmission over long distances has grown as fast as the Internet. In turn, the explosive growth in the optical backbone has created a bottleneck at the edge of the long distance network. This pushes the bandwidth requirements into shorter-reach applications.

The need for connectivity is driving optics closer to the end user. As the limits of optical technology are approached, requirements for fiber and optical components are becoming specialized for particular applications. For example, current transmission fibers are suboptimal for next-generation long-haul networks which will transmit information at terabit-per-second speeds over thousands of kilometers. Meanwhile, the desire to use low-cost

OPTICAL FIBER TELECOMMUNICATIONS.
VOLUME IVA

ISBN: 0-12-395172-0

components for shorter-reach architectures has driven the development of very-wide-bandwidth fibers. New fiber designs have been evolving to address characteristics such as chromatic dispersion, optical nonlinearities, system cost, and optical bandwidth that are specific to particular applications and architectures. Indeed, many of the major fiber and cable suppliers throughout the world have begun to market application-specific fiber designs. This chapter discusses the requirements of applications spanning the spectrum from transcontinental transport to home wiring and demonstrates the benefits of optimizing fiber design for specific requirements.

Section 2 covers the historically most important markets for optical communication: long-haul and undersea. These were the markets for which optical fiber communication was first developed. Section 3 covers the emerging metropolitan and access markets, which are expected to be the next high-growth arena over the next 5 years, while Section 4 covers development of multimode fiber for local area networks. Plastic optical fiber offers the potential of low-cost installation and is expected to make inroads in traditionally copper-based infrastructures such as in-building wiring and optical interconnection. This new fiber is discussed in Section 5. Section 6 covers an exciting new area, photonic bandgap structures, which offer the potential of guiding light through a hollow-core fiber, removing the traditional impediments like nonlinearity and optical attenuation which define current communications architectures.

2. Long Haul and Undersea Systems

2.1. INTRODUCTION

The first applications of fiber optic communication were to carry aggregated voice traffic between major metropolitan areas, such as the trunk lines from Washington, DC to Boston. In the United States, typical distances between major switch centers are on the order of 1600 km, while in Europe, these distances are typically 400 km. However, with the advent of all-optical or photonic switching located at these centers, the transmission distances without electronic regeneration could reach well into the thousands of kilometers in both cases, with the application space for these systems spilling over into the metro and regional networks. Such ultralong distances have historically been reserved for point-to-point undersea fiber systems where transoceanic distances are typically 10,000 km and 4000 km for Trans-Pacific and Trans-Atlantic routes, respectively. As these distances are approached in terrestrial applications, it is not unreasonable to think of using similar system solutions for land applications.

Both long-haul and undersea systems depend heavily on dense wavelength-division multiplexed (DWDM) signals to achieve high-capacity transport.

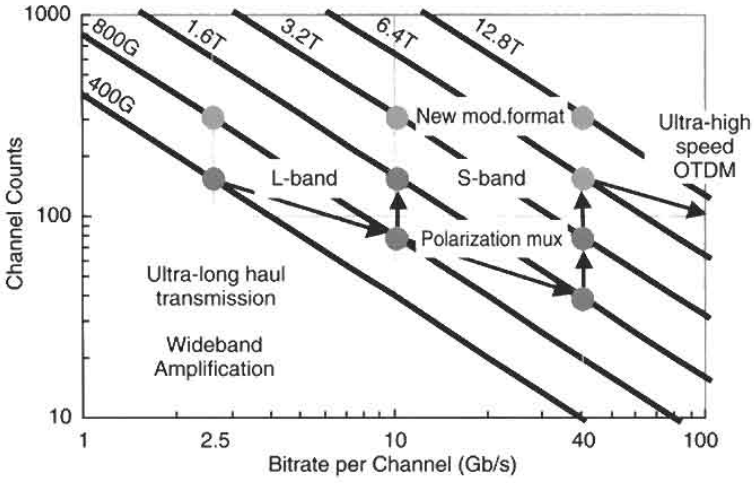


Fig. 1 Advances in system capacity and enabling technologies.

DWDM is a very effective means of sharing transmission costs when fiber and other common components, such as optical amplifiers, dominate the overall system cost. As shown in Fig. 1, the aggregate capacity of a single optical fiber can be increased by either increasing the bitrate or by increasing the number of wavelength channels using DWDM. The former requires development of new high-speed electronics, while DWDM allows fiber and amplifier costs to be shared among all channels, driving down the total system cost per channel. Since information must still be coded onto the wavelength channels, today's long-haul systems combine time-division multiplexing (TDM) with DWDM, taking advantage of high-speed TDM advances to further reduce the system cost per bit per channel.

Advances in high speed electronics have allowed data rates to increase from OC-12 to OC-48 to OC-192 (0.622 Gbps to 2.48 Gbps to 10 Gbps).[2] However, increasing data rates beyond 10 and 40 Gbps could have a diminishing return in terms of cost as one migrates from the realm of electronic components for multiplexing, demultiplexing, modulation, and detection to that of optical processing techniques currently envisioned for 160 Gbps and beyond.[3]

Current long-haul system development efforts have focused on wideband DWDM and ultralong transport.[4,5] These systems are enabled by new modulation formats,[6,7] wideband amplification,[8] wideband dispersion compensation[9–11] and the use of forward error correction coding.[1,4] Taken as a whole, these systems will deliver the lowest cost per transmitted bit over the longest distances. Optical fiber is an integral component of the entire system. The fiber's parameters have a significant impact on both cost and performance and influence the choice of most other components, such as amplifiers and compensators. In fact, the use of wideband DWDM over ultralong distances

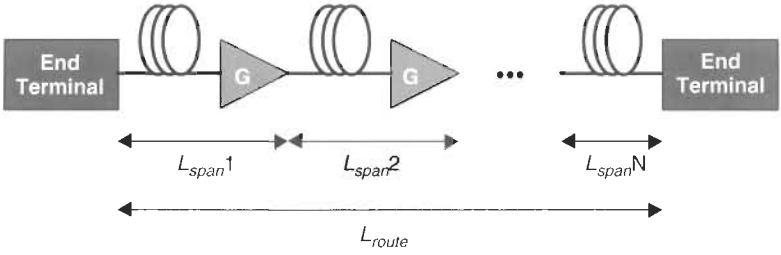


Fig. 2 Simple optical transmission system made up of N concatenated spans of fiber and line amplifiers.

has elevated the fiber requirements in terms of dispersion management, nonlinear performance, distributed gain, spectral loss, and polarization mode dispersion (PMD).

2.2. *TYPICAL LONG-HAUL LINK*

A simple long-haul fiber system is shown in Fig. 2. Here, L_{route} is the total distance between the two end terminals which contain the signal transmitters, wavelength multiplexers and receivers, and L_{span} is the distance between amplifier sites. In a system with unit gain, the amplifier gain G is set to exactly compensate for the loss of the span. The span loss is the sum of the attenuation of the outside plant fiber and losses due to splices and various components used to manage signal properties and integrity, such as dynamic gain equalizing filters (DGEFs), dispersion compensation modules (DCMs), optical add/drop multiplexers (OADMs), and connectors at the amplifier site.

Fiber spans cannot be concatenated indefinitely, since the amplifiers add noise to the system and since power penalties associated with chromatic or polarization mode dispersion and nonlinear crosstalk accumulate with length. A compromise between span length and route distance has to be made in order to optimize a system in terms of both cost and performance. This effect can be understood by investigating the link optical signal-to-noise ratio (OSNR), approximated by

$$\begin{aligned} \mathrm{OSNR[dB]} = {} & 58\,[\mathrm{dBm}] + P_{ch}\,[\mathrm{dBm}] - F\,[\mathrm{dB}] \\ & - L_{span}\,[\mathrm{dB}] - 10\log_{10}(N_{span}) \end{aligned} \qquad \text{Eq. 1}$$

where P_{ch} is the launched power per channel into a span in dBm, assuming unit gain for the span, F is the effective noise figure of the span, L_{span} is the span loss in dB, and N is the total number of spans for the route. There are many ways to optimize the performance of a route based on this equation. Typically, a given receiver OSNR is required to achieve error-free operation for a certain data rate; for example, one might require an OSNR of 23 dB for DWDM system transmitting at OC-192 (~10 Gbps). Then for a given P_{ch}

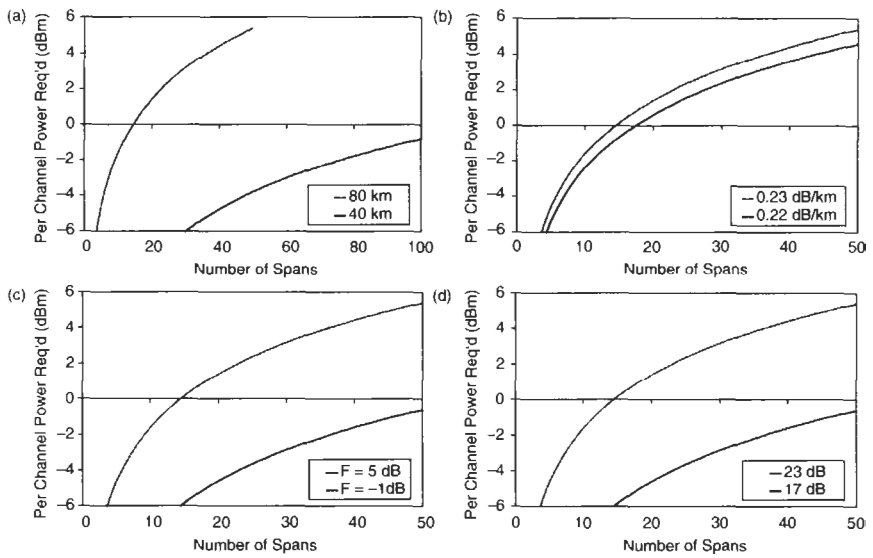


Fig. 3 Effect of (a) span distance, (b) fiber loss, (c) amplifier noise figure, and (d) required OSNR on per-channel power requirements for a 4000 km route link with 80 km spans. $F = 5$ dB for (a) and (b).

and F, longer route distances can be attained using shorter spans (Fig. 3a), lower fiber loss (Fig. 3b), improved noise figure (Fig. 3c), or lower OSNR requirements (Fig. 3d).

Figure 3a illustrates that for a 4000 km route, reducing the span length from 80 to 40 km reduces the per-channel power requirement by over 6 dB. This lower optical power per channel helps reduce penalties due to nonlinear distortion or could be used to increase the total number of channels in power-limited systems, such as those encountered when operating optical line amplifiers in gain saturation. The drawback of reduced amplifier spacing is increased system cost for the same route distance. For the example of Fig. 3a, twice as many optical line amplifiers would be required for the 40 km span case compared to the 80 km span case.

Figure 3b indicates that a reduction of 0.01 dB/km in the cabled fiber span loss results in a lowering of the required optical power per channel by nearly 1 dB for 80 km spans. Alternatively, if the system is limited to 0 dBm of per channel power (for example, because of nonlinear distortion), then a route with 0.22 dB/km average loss would allow 3 more spans than a route with a 0.23 dB/km average loss. Figure 3c illustrates the impact of a reduction in F from 5 dB to -1 dB. Here a significant reduction in required per-channel power can be attained using distributed Raman amplification.[12] Similarly, a reduction in the OSNR required by the receiver from 23 to 17 dB would have the same effect. Such reduction is achievable using FEC.[13,14]

An alternative to lowering F, L, and OSNR is to increase P_{ch}. For long distance transport, however, severe limitations arise due to optical nonlinearities such as self-phase modulation (SPM), cross-phase modulation (XPM) and four-wave mixing (FWM) which generate signal distortion and crosstalk that cannot be undone at the receiver. These effects result in power penalties that decrease the OSNR at the receiver. Thus, P_{ch} is limited by the tolerance of the received signal OSNR to nonlinear crosstalk degradations. As such, there is a power penalty budget typically built into OSNR for a given P_{ch}. Crosstalk and distortion from optical nonlinearites are not the only sources of OSNR degradation. Chromatic and polarization mode dispersion (PMD) generate intersymbol interference, resulting in system power penalties. Dispersion can be managed in most cases using discrete compensators, but at the expense of increased system cost.

As will be seen, each of the impairments illustrated in Fig. 3 are strongly affected by transmission fiber design. This suggests that fiber parameters such as dispersion D (ps/nm · km), dispersion slope S (ps/nm^2 · km), mode field effective area A_{eff} (μm^2), and spectral attenuation α (dB/km) can be optimized for specific system architectures. The impact of these parameters on long-haul, wideband DWDM system performance and cost is reviewed below.

2.3. *SPAN LOSS*

Fiber attenuation is spectrally dependent, showing a minimum around 1550 nm. Systems developed in the 1980s focused on this region, and it is a happy coincidence of nature that one of the most efficient optical amplification mediums, the erbium-doped optical amplifier, operates in this region.

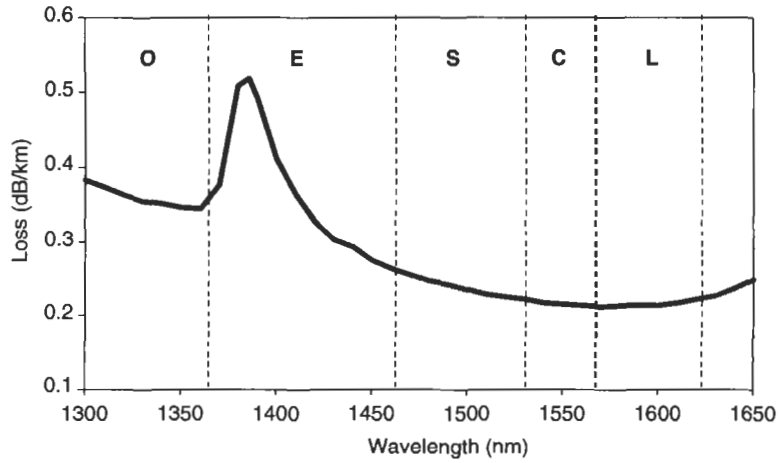


Fig. 4 Typical optical fiber cable loss spectrum indicating lowest-loss regions associated with C- and L-transmission bands.

Thus, long-haul systems were developed first in the C- (1530 to 1565 nm) and then the L-band (1565 to 1625 nm) regions as indicated in Fig. 4. However, span loss is affected by more than just the intrinsic losses associated with the fiber. Cabling induces losses through bend-related effects and fiber designs with greater bend sensitivity will show higher added loss when placed inside of a cable, especially as temperatures decrease below 0°C. As such, cable specifications related to maximum individual fiber loss as opposed to maximum average fiber loss are critical in insuring lowest-loss spans. Packing design, such as ribbon or loose-tube structures, affect cabled fiber density and significantly impacts cable-induced losses.

Fusion splicing of fibers in cable segments adds additional loss. Since cables are typically supplied in 2 to 6 km segments, an 80 km span could have up to 40 splice points. Splice loss can range from 0.02 dB for individual fibers to as high as 0.06 dB for mass fusion spliced fiber ribbons.[15] Even at 0.02 dB per splice, the accumulated loss approaches 1 dB per span. Finally, the outside plant is connected to the inline amplifier sites through a set of optical jumpers with mechanical connectors, adding another few tenths of a dB to the loss budget.[16]

2.4. *WIDEBAND DISPERSION COMPENSATION*

Signal distortion and crosstalk that accumulate during transmission influence the required OSNR at the receiver. Figure 5 compares the nominal dispersion of several commercial ITU G.655 or nonzero dispersion-shifted fiber (NZDF) types. In terms of system performance, too low a level of dispersion in the DWDM band of interest is not adequate to prevent nonlinear crosstalk such as four-wave mixing.

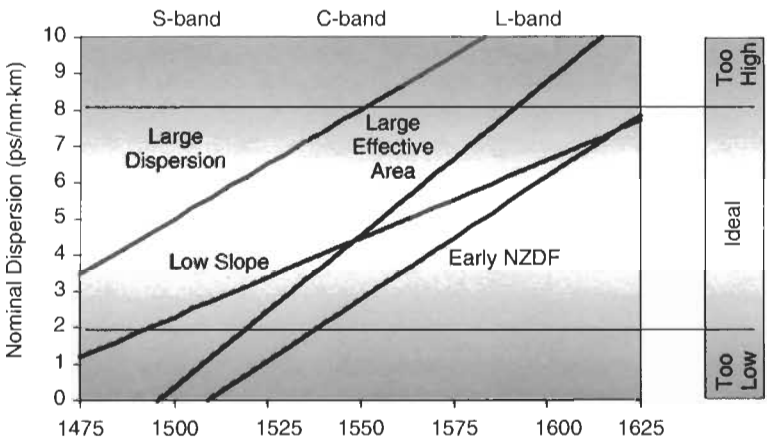


Fig. 5 Comparison of the spectral dispersion of several nonzero dispersion-shifted (G.655) fiber types.

Alternatively, too high a level of dispersion results in intersymbol interference (ISI) as the individual bits broaden to the point of overlapping into adjacent time slots. ISI attributed to chromatic dispersion can be corrected using a dispersion-compensating module (DCM). However, these devices are not able to exactly compensate dispersion across a very wide band of channels, resulting in residual dispersion which limits either the system reach or the number of channels. Also, DCMs add to the loss budget. Typical dispersion-compensating fiber has attenuation of about 0.5 dB/km, and although the required length (and therefore the insertion loss) is proportional to the dispersion of the transmission fiber, compensation of 100 km of standard single-mode fiber entails an additional loss of up to 10 dB. Thus, too high a level of dispersion in the transmission fiber results in higher DCM cost and added amplifier gain needed to overcome the insertion loss of the DCM.

Figure 6 illustrates the residual, or uncompensated, dispersion for two types of NZDF (low-slope and large effective area) paired with two types of commercially available DCMs (standard DCF and extra-high-slope DCF). As seen here, the low-slope NZDF allows a wider wavelength band of compensation compared to the large-area NZDF.

The allowable parameter space for residual dispersion tightens as the bitrate increases from 10 Gbps toward 40 Gbps since the tolerance to dispersion at the receiver decreases as $1/B^2$ where B is the bitrate.[17] Thus, while a 10 Gbps system can tolerate 1000 ps/nm of residual dispersion, a 40 Gbps system would require ~60 ps/nm of residual dispersion for the same level of OSNR or power penalty.

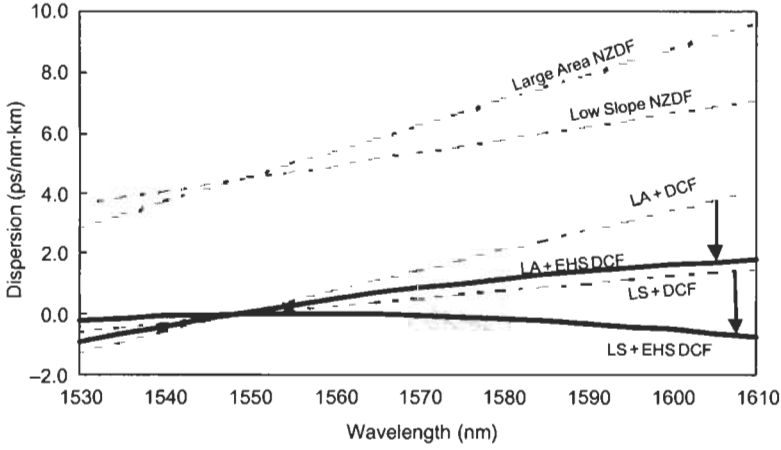


Fig. 6 Improvement of residual dispersion of low-slope (LS) and large effective area (LA) NZDF fiber after compensation with typical commercially available extra-high-slope DCF (RDS ~0.009) compared to standard DCF (RDS ~0.003). The shaded region indicates the operating window of a 100 km span at 40 Gbps.

Table 1 **Approximate Relative Dispersion Slope (nm^{-1}) Values of Various G.655 and G.652 Fiber Types**

	S-band 1500 nm	*C-band 1550 nm*	*L-band 1590 nm*
Early NZDF	NA	0.024	0.012
Low slope NZDF	0.019	0.009	0.007
Large area NZDF	NA	0.018	0.011
Large dispersion NZDF	0.012	0.008	0.006
G.652 SSMF	0.005	0.004	0.003

For complete compensation across a wide bandwidth, both the dispersion D and dispersion slope S of the DCF must scale with those of the transmission fiber. In other words, the ratio of these values, the relative dispersion slope RDS where RDS $= D/S$, should be equal.

$$\text{RDS}_{\text{transmission}} = \text{RDS}_{\text{DCF}}$$

As shown in Fig. 6, RDS is an indicator of the relative ease of creating a DCM for wideband WDM operation. Thus, the lower the RDS, the better for broadband dispersion compensation. The RDS values of typical fibers are listed in Table 1 for wavelengths of interest in the S, C, and L bands. Note that many of the combinations of fiber type and desired wavelength band cannot be accommodated by currently-available DCMs, which have RDS values less than $0.009\ nm^{-1}$, as shown in Fig. 6. Also, although standard singlemode fiber (SSMF or type G.652) has the lowest RDS in all the bands, its absolute level of dispersion, too high to be displayed on the scale in Fig. 5 and Fig. 6, adds significant cost to the compensation of transmission systems.

An alternative to using a wideband DCM is to split the DWDM bands and compensate them individually.[18] This approach, however, requires an additional filter to separate the bands on entering and exiting the amplifier sites of Fig. 2, adding up to 3 dB of loss per span. This loss limits route performance according to Eq. 1 and as shown in Fig. 3b. A 3 dB added loss is equivalent to a 3 dB per channel power penalty, which essentially cuts the system reach in half.

2.5. *OPTICAL NONLINEARITIES*

One of the principal factors limiting system reach and described in Section 2.2 is the occurrence of nonlinear distortion. Nonlinearities arise since the refractive index of glass varies slightly when exposed to light, causing optical pulses to alter themselves or each other. The details of these interactions are described more fully in Chapter 13 in Volume IVB, but it suffices for the present discussion to note that nonlinear distortions decrease the OSNR at the receiver. The power penalty associated with nonlinear crosstalk can be represented by the

increase in P_{ch} required to restore the OSNR to its value in the absence of nonlinearities. However, increasing P_{ch} has diminishing returns since higher P_{ch} generates even more crosstalk, which could further degrade OSNR at the receiver. This also increases the overall system cost by requiring higher power amplifiers, for example. The goal is to optimize P_{ch} for a required receiver OSNR. A common means of measuring signal robustness is to observe the Q-factor of the received signal. The Q-factor is directly related to the OSNR and signal bit-error rate.[19]

The design of the optical fiber plays a significant role in determining the system impact of signal distortion due to nonlinear crosstalk. Fiber characteristics such as the nonlinear refractive index n_2, chromatic dispersion D, and mode-field effective area A_{eff} determine the amount of nonlinear crosstalk generated for a given DWDM channel spacing. The magnitude of nonlinear effects is dependent upon the ratio of the nonlinear index and the intensity of the optical field, which may be represented by the inverse of the effective area. This is demonstrated in the nonlinear coefficient γ given by

$$\gamma = \frac{2\pi n_2}{\lambda A_{eff}} \left[\frac{1}{W \cdot km} \right] \qquad \text{Eq. 2}$$

where λ is the signal wavelength and n_2 is the nonlinear refractive index. The effect of dispersion depends on the particular nonlinear phenomenon. For example, stimulated Raman scattering (SRS), in which the optical field couples with the vibration of the glass matrix and is partially scattered, depends on the magnitude of $D^{1/2}$. Four-wave mixing (FWM) generated signals are inversely dependent on D.[20] Distortion due to self-phase modulation (SPM), in which the leading edge of a pulse modulates the refractive index seen by the trailing edge of the pulse and causes pulse broadening in typical fiber, is exacerbated by large amounts of dispersion. Obviously, if lowest nonlinear performance in terms of SRS and FWM is the only goal, a fiber with the largest possible A_{eff} and D is desired. In this case, SSMF with an $A_{eff} = 80\,\mu m^2$ and $D = 17\,ps/nm \cdot km$ at 1550 nm would be the best choice among commercially available terrestrial fibers. In fact, SSMF has the most desirable characteristics for CATV linear transmission of a single wavelength channel, since the dominant impairments are SRS and stimulated Brillouin scattering (SBS).

As previously stated, a large value of dispersion has drawbacks; and as will be seen, the choice of optical pulse shapes along with proper dispersion management throughout the route can be used to help control impairments due to nonlinearities such as FWM, XPM, and SPM.

2.6. *DISPERSION MANAGEMENT OF OPTICAL NONLINEARITIES*

Although the onset of nonlinear crosstalk is very much dependent on the optical intensity as indicated by Eq. 2, the phase interaction of these signals with

other channels and within a channel is equally important.[21] Worst case interaction occurs when the interacting frequencies are in phase throughout the duration of signal propagation. Dispersion disrupts this phase relationship since different frequencies propagate at different group velocities and reduce the interaction length. Hence, nonlinear crosstalk can be effectively controlled by managing dispersion throughout the route,[22] for example by periodically introducing a component with negative dispersion, such as a dispersion-compensating fiber or grating module, along the route. Optimum locations of the DCMs for dispersion management are at the amplifier point of each span in Fig. 1.[23] A comparison of dispersion maps for a 1600 km system with 100 km spans are shown in Fig. 7.[22] In this work, the case of 100% per-span compensation (Fig. 7b) resulted in the lowest Q (Fig. 8) or highest nonlinear

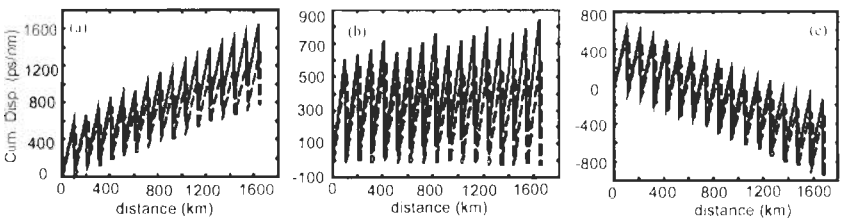


Fig. 7 Dispersion maps: (a) under (90%) per-span dispersion compensation, (b) full (100%) per-span dispersion compensation, and (c) over (110%) per-span dispersion compensation. Solid: 1559.4 nm, dashed: 1543.2 nm. Reprinted with permission from Ref. 22.

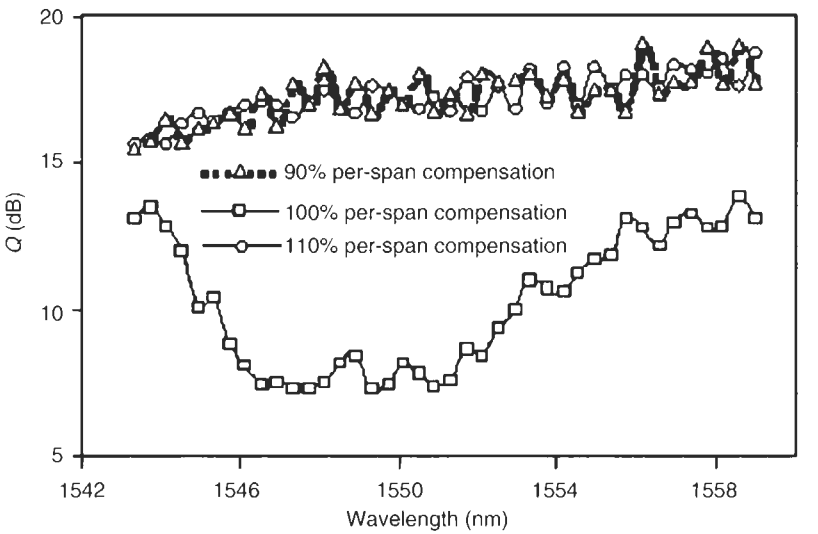


Fig. 8 Q factors for the three dispersion maps, experimental results.

crosstalk, since the signals were realigned with each other at the end of each span, resulting in an in-phase condition. The other conditions of under (90%) or over (110%) per-span compensation resulted in lowest nonlinear crosstalk and hence highest Q.

Dispersion management is a useful tool in suppressing nonlinear crosstalk in a transmission system, but too much dispersion can be costly in terms of components or other distortions. Is there an optimal amount of dispersion that balances system costs and performance? The answer to this question is greatly dependent on a wide array of system parameters and must be extensively simulated to understand the results. For example, simulation of a 5-channel 40 Gbps system at 200 GHz channel spacing over 80 km spans and 480 km total route length indicates that moderate dispersion results in optimum system performance as indicated in Fig. 9.[24] The fiber parameters used in the simulation were $A_{eff} = 55\,\mu m^2$, $n_2 = 2.4\,(mW \cdot km)^{-1}$, and 0.21 dB/km loss. These simulations looked at an array of dispersion values and line compensation ratios. The simulations also accounted for precompensation values and the effects of pulse shape, return to zero (RZ), or nonreturn-to-zero (NRZ). Optimum precompensation was found to be 10% for RZ and 20% for NRZ.

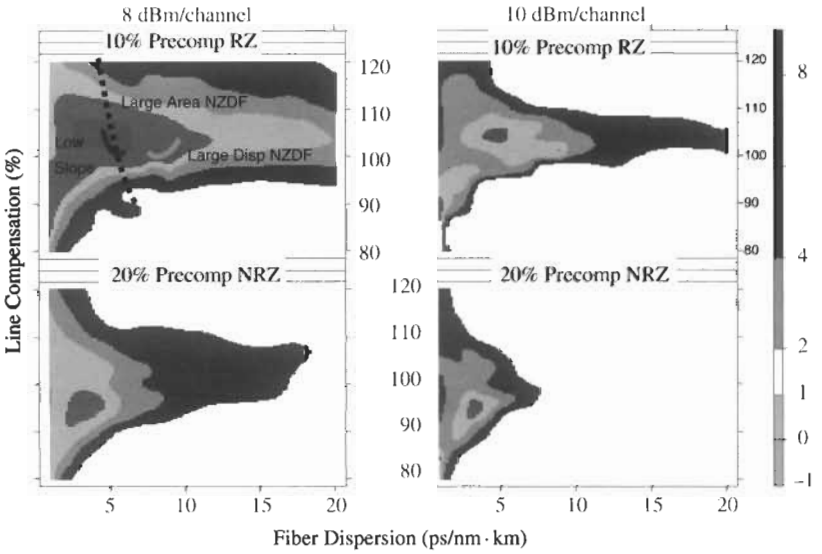


Fig. 9 Simulation results of power penalty contours (indicated in dB by bar on right) for percent per-span compensation versus fiber dispersion for a 40 Gbps DWDM system for two-channel power and precompensation levels. 5 channels, 200 GHz spaced NRZ, 6 × 80 km system. Curves indicate dispersion and compensation ratios across C-band channels (from 1530 to 1565 nm) using a high-slope DCF for large-area (dashed), low-slope (solid), and large-dispersion (solid orange) NZDF. Reprinted with permission from Ref. 24. See also Plate 1.

If the dispersion is too low, excessive XPM and FWM distort the signal. If the dispersion is too high, SPM and XPM cause unacceptable distortion. In applying the criteria that the penalty should be less than 2 dB, the optimal dispersion is in the 3–5 ps/nm · km range. This dispersion range must be sufficient to include the dispersion of the entire DWDM band of channels. Thus, a low-slope NZDF will have a clear advantage over a higher-slope NZDF as indicated by the inset curves on the 8 dBm RZ contours when using a commercially available high-slope DCF ($D = -107$ ps/nm·km and $S = -0.93$ ps/nm^2·km). Only the lower-slope NZDF matched with this DCF will fit the entire C-band of DWDM channels into the lower power penalty contours.

It is clear from Fig. 9 that reducing the signal power from 10 dBm per channel to 8 dBm per channel helped broaden the optimal dispersion range. Alternatively, increasing the A_{eff} by 1 dB (from 55 μm^2 to 70 μm^2) will increase the size of the penalty contours by 1 dB and hence broaden the optimal dispersion range. However, the resultant effect will not be enough to bring the entire C-band of the large-area NZDF into the required dispersion range, since the contour bands are plotted in 2 dB increments above the 2 dB power penalty contour. Thus, this NZDF type will require a better matched DCF to bring the system within the required region. Although the large-dispersion NZDF is within the penalty contours for the RZ case, only the low-slope NZDF could fall within acceptable penalty contours for the NRZ case. While RZ modulation may give better system margin, NRZ may result in lower system cost since it requires fewer components. In any fiber case, clearly excellent dispersion slope matched compensation is needed.

2.7. *WIDEBAND AMPLIFICATION*

The effect of stimulated Raman scattering (SRS) is to generate scattered light of slightly lower energy, or longer wavelength, than the incident light. If a probe signal is present near the scattered wavelength, the probe will be amplified. This effect can be used to advantage by propagating pump light in the transmission fiber at a wavelength suitable for creating gain in the DWDM signal band. The bandwidth of the Raman gain spectrum associated with SRS is of the order of 100 nm,[25] so distributed Raman amplification (DRA) allows for a wider transmission band than the ~40 nm bandwidth associated with EDFAs.

The efficiency of Raman gain is associated with the A_{eff} of the fiber and the GeO_2 content of the core.[26] Raising the GeO_2 content increases the Raman gain coefficient, but it would require large amounts of GeO_2 to significantly impact this value. The on-off Raman gain $G_{on\text{-}off}(f)$ at a frequency f is very sensitive to A_{eff} as given by[27]

$$G_{on\text{-}off}(f) = \exp\left\{C_R(f) \cdot P_{pump} \frac{1 - \exp\left[-\alpha_{pump} L_{span}\right]}{\alpha_{pump}}\right\} \qquad \text{Eq. 3}$$

where α_{pump} is the fiber loss coefficient at the pump and P_{pump} is the pump power. $C_R(f)$ is the Raman gain coefficient at frequency f given as

$$C_R(f) = \frac{1}{k_{pol}} \frac{g_R\left(f_{pump} - f\right)}{A_{eff}} \frac{f_{pump}}{f_{ref}} \qquad \text{Eq. 4}$$

where f_{pump} is the frequency of the pump, $k_{pol} = 2$ for depolarized pump and signal frequencies and g_R is the Raman efficiency measured at f_{ref}. Figure 10 illustrates the effect of A_{eff} on Raman gain according to Eq. 3 for equivalent α_{pump} and g_R. Here, the moderate A_{eff} of the low-slope NZDF allows for significant gain improvement compared to large-area NZDF or SSMF for the same pump power. This increased gain and subsequent OSNR improvement can be traded off for reduced pump power and hence system cost savings. Specifically, a given $G_{on\text{-}off}$ and OSNR can be achieved at a lower pump power in a transmission fiber with moderate A_{eff} compared to one with higher A_{eff}.[28]

The added gain from DRA keeps the signal power higher throughout the length of the fiber as compared to the case without DRA. This elevated signal power increases Q, but also increases the impact of Kerr-type nonlinearities (because the average signal power is increased) and noise (because noise is also amplified). The amount of Kerr nonlinearity is proportional to power density, so a larger A_{eff} will decrease the nonlinearity. The dominant noise source originates from Rayleigh backscattering of either the pump or the signal and is inversely proportional to A_{eff}. This noise is added to the ASE noise from the discrete amplifiers of Fig. 2. Further increasing pump power increases the Rayleigh backscatter noise[12] and limits the peak OSNR improvement associated with DRA. However, the net effect is to significantly improve the OSNR,

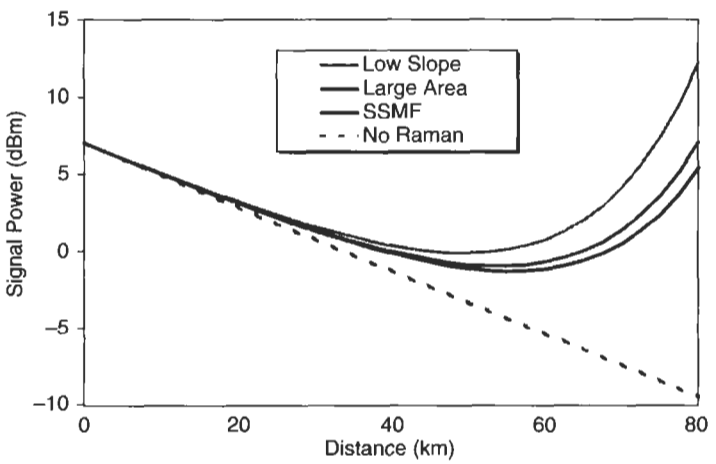


Fig. 10 Calculated signal power vs. span length of various fiber types for a 1450 nm pump of 500 mW.

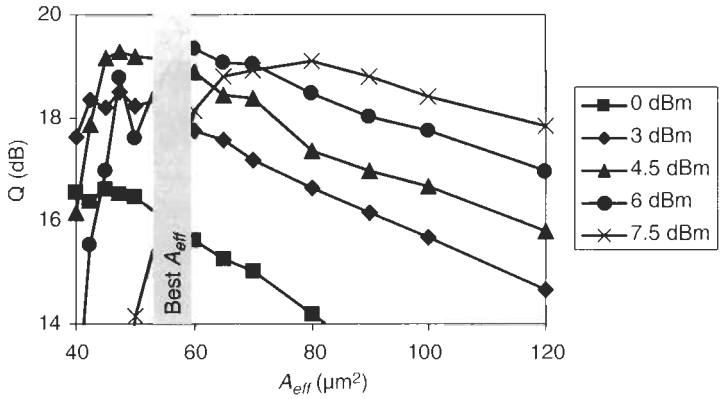


Fig. 11 Q versus A_{eff} for P_{ch} ranging from 0 to 7.5 dBm. System simulated is 7 channel × 40 Gbps, 100 GHz spacing, NRZ, 6 × 120 km, $n_{sp} = 2$ and 500 mW pump. Indicates that the best Q over the greatest range of P_{ch} is associated with moderate A_{eff}(55–60 μm²).

since the gain of the DRA reduces the need for discrete amplifier gain, which adds much more noise. Thus, the use of DRA reduces the overall effective noise figure of the system, the benefit of which is indicated by Fig. 3c.

Though reducing A_{eff} adversely affects the nonlinear crosstalk and noise performance of a DRA link, the benefits to OSNR of the efficient gain and reduction in effective noise figure positively enhance the system. To determine the tradeoff among these effects, the impact of A_{eff} on a 100 GHz channel spacing, 7 channel × 40 Gbps DRA system is illustrated in Fig. 11.[29] Here a 500 mW backward Raman pump is inserted at the amplifier at the end of a span. The simulation results indicate that moderate A_{eff} results in the best Q for a broad range of P_{ch}.

2.8. POLARIZATION MODE DISPERSION

Polarization mode dispersion (PMD) is a random effect resulting from intrinsic and extrinsic sources that causes the group velocity to vary with polarization state. Thus, over distance, a pulse will break up into multiple pulses, each with a specific polarization. This can severely constrain the bitrate-distance product of a system if left uncompensated. Intrinsic sources of PMD are minor imperfections resulting in noncircular fiber core geometry and residual stresses in the glass materials near the core region. Extrinsic sources of PMD include stress due to mechanical loading, bending, or twisting of the fiber and cable. Each of these sources induces a birefringence in the fiber, causing the otherwise degenerate polarizations of the LP01 mode to propagate at different speeds along the fiber and hence dispersing the optical pulse.

Typical optical systems allow for a 1 dB power penalty for PMD associated with ~15% of the time slot for RZ modulated pulses. As such, a 10 Gbps link can allow up to 15 ps of dispersion associated with PMD, while a 40 Gbps system can only tolerate 3.8 ps. It is important to note that not all of this PMD is allotted to the transmission fiber. Other components, such as EDFAs and DCMs, may also introduce PMD into the link. As such, there is an accumulation of PMD according to

$$PMD_{total} = \{N_{span}[(PMD_{fiber})^2 \cdot L_{span} + (PMD_{comp1})^2 + (PMD_{comp2})^2 + \mathrm{K} + (PMD_{compN})^2]\}^{1/2} \quad \text{Eq. 5}$$

where N_{span} is the number of spans in the route, PMD_{fiber} is the cabled fiber PMD in ps/(km)$^{1/2}$, L_{span} is the span distance in km, and $PMD_{comp1...N}$ are the PMD values related to the individual components located in the amplifier sites for each span. Assuming that there is only one component at each amplifier site, contributing 0.5 ps of PMD, and that the average span length is 80 km, one can find the required PMD_{fiber} needed for a given route distance and data rate as shown in Fig. 12. Here one sees that a PMD_{fiber} of 0.1 ps/(km)$^{1/2}$ is adequate for 10 Gbps RZ data rates and distances beyond 8000 km, while 40 Gbps RZ systems are constrained by PMD_{fiber} to less than 3000 km for realistic values of cabled fiber PMD unless the component PMD values are reduced significantly as well.

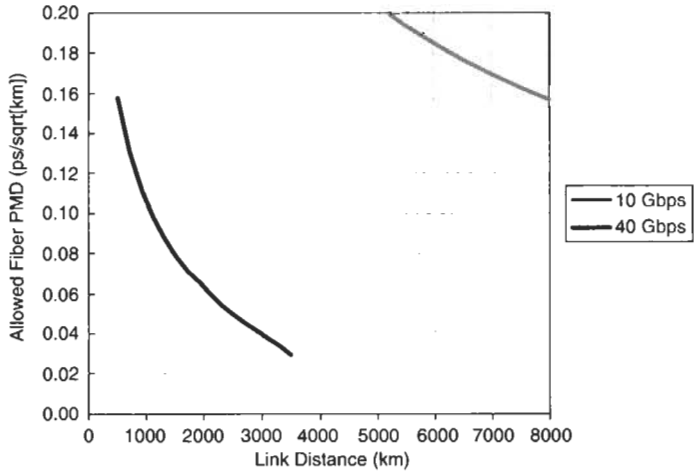


Fig. 12 Estimate of required cabled fiber PMD levels for 10 and 40 Gbps RZ transmission systems as a function of route distance, assuming 80 km spans and a discrete component with PMD of 0.5 ps located at each amplifier site.

2.9. *UNDERSEA FIBERS*

All of the effects described above become even more critical for undersea DWDM systems due to the increased route distances compared to terrestrial DWDM systems. One of the principal differences between terrestrial and undersea transmission systems is the amplifier spacing. Terrestrial systems attempt to space amplifiers as far away from each other as possible in order to reduce system costs. Undersea systems are more focused on extending to the longest optical path possible without regeneration. For a given P_{ch}, closer amplifier spacing allows for longer optical route distances as indicated in Fig. 3a. As undersea systems have adopted DWDM to increase transmission capacity, optical fiber development has focused on dispersion management.[30]

Historically, optical fiber development for WDM undersea systems has concentrated on reducing the slope of the path average dispersion in order to widen the DWDM channel bandwidth for adequate OSNR. Until recently, this was done with a pair of negative dispersion fibers compensated with a standard single-mode fiber that is also part of the transmission path. This configuration is illustrated in Fig. 13. The pair of negative dispersion fibers is used to flatten the dispersion slope of the span between the amplifiers. The first fiber is a large-mode field (LMF) fiber and allows for large P_{ch} to be launched with small nonlinear penalty; while reduced slope fiber (RSF) follows to help flatten the dispersion slope of the amplified span. Since the RSF has a smaller

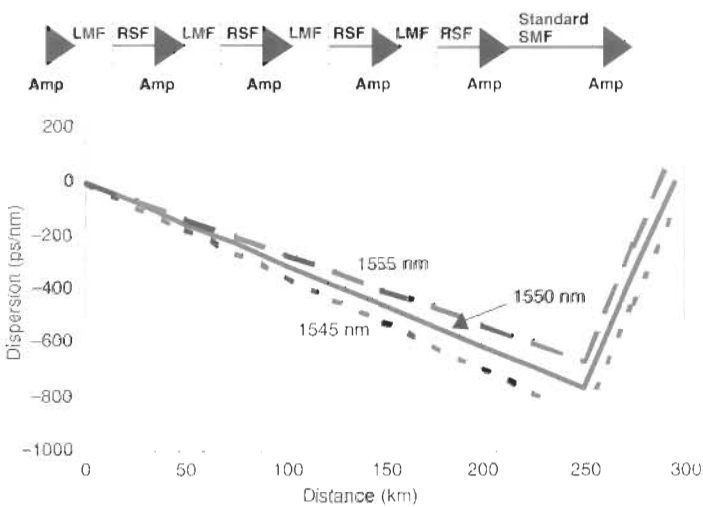


Fig. 13 Typical dispersion map of undersea fibers using LMF and RSF map with SMF for transmission and compensation.[30]

mode field than the LMF, it is used in a position where the P_{ch} has been attenuated due to loss and therefore will not generate significant nonlinear crosstalk.

After several amplifier spans, SMF is used to dispersion compensate the previous LMF + RSF spans. An advantage of using SMF as the compensating fiber is that it can also be used as part of the transmission span. In terrestrial applications, the DCF is normally lumped at the amplifier site. Such a technique would increase overall system costs by taking up valuable space in the amplifier closure as well as limiting system reach by inserting an otherwise unnecessary discrete loss element in the span. The overall slope of the spans shown in Fig. 13 can be on the order of 0.08 ps/nm^2 · km. Further reduction of the dispersion slope could be used to increase DWDM bandwidth and hence transmission capacity as indicated in Fig. 14. The limitation of this technique is that all of the fiber types in Fig. 13 have positive dispersion slope, making it difficult to further flatten the overall slope of the system.

A fiber with negative dispersion slope would solve the need for further dispersion slope reduction. Currently, an enormous amount of development is focused on mating a pair of fibers with opposite dispersion and dispersion slope, $+D$ and $-D$ fibers.[30,31] The slope-flattening capabilities of a pair of fibers for undersea application are shown in Fig. 15. Here, residual dispersion for wavelength channels at 1530 and 1565 nm are nearly indistinguishable from the channel at 1550 nm for the 300 km distance shown. The significant improvement over the LMF + RSF design is not only due to using a negative dispersion slope for the $-D$ fiber, but also to the excellent matching of the RDS of these fibers.[32]

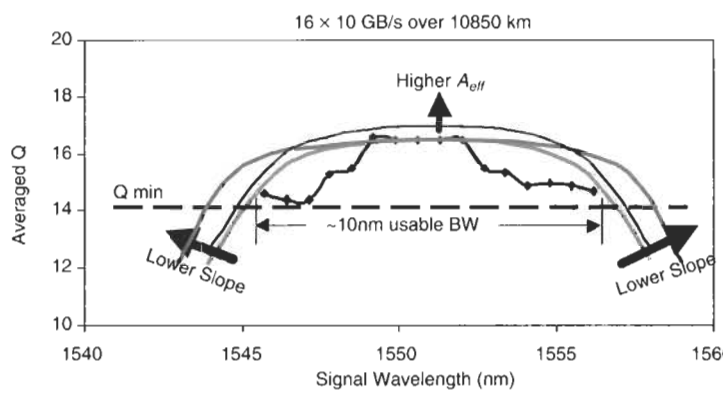


Fig. 14 Analysis indicating that reducing dispersion slope further increases usable bandwidth as opposed to increasing A_{eff}. Reprinted with permission from Ref. 30.

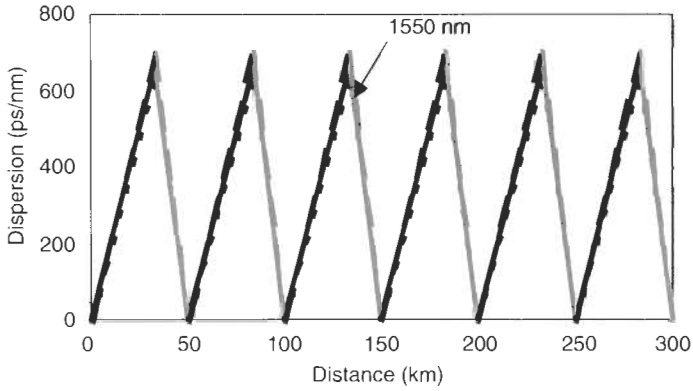


Fig. 15 Dispersion map of a $+D$ and $-D$ pair of fibers for undersea applications. The solid line indicates the dispersion of a 1550 nm channel. The dashed lines indicate the dispersion of 1530 and 1565 nm channel wavelengths.

The $+D/-D$ fiber pairs fulfill one of the enabling technologies for high-capacity ultralong-haul terrestrial and undersea systems which include[33]

- advances in modulation format to improve spectral efficiency;
- wideband amplification to provide gain beyond the conventional 30 nm band of EDFAs;
- FEC to reduce the required OSNR as indicated in Fig. 3d;
- dispersion flattened chain of optical fibers.

The same technique of dispersion management can be used to extend the reach of terrestrial systems. However, the challenge to implementing a dispersion-flattened chain of optical fibers in terrestrial applications lies in overcoming field splicing issues and administration of the field installation of the fibers. Unlike undersea systems, terrestrial systems do not deploy the transmission equipment with the fiber. Thus, it is not known a priori what systems will be used with the fiber that is installed. The challenge then becomes knowing what fiber type, lengths, and length intervals need to be installed well in advance of system development.

At present, fiber design for long-reach, high-capacity systems required for backbone applications, whether terrestrial or undersea, is a careful balancing of fiber properties such as dispersion, dispersion slope, and effective area. Although many design parameters are entailed in defining a system architecture, such as distance between inline amplifiers, use of Raman amplification, number of spans, and modulation format, the overriding goal is to reduce system cost. To that end, the fiber design choices amount to reducing the cost of dispersion compensation, increasing the available bandwidth of the system, and decreasing the cost of Raman amplification.

3. Fibers for Metro and Access Systems

3.1. *INTRODUCTION*

Access to the long-haul network typically occurs in metropolitan areas where the information is aggregated for long-haul transmission or demultiplexed for distribution. The requirements in this space are significantly different from those of long-haul transmission, since the population density and geography necessitates somewhat lower capacity than the fat long-haul pipe but greater flexibility in distribution. These so-called Metro networks take traffic off long-haul networks for distribution to enterprise, residential, and business access. This environment is much more complex operationally than long-haul due to the interplay of demographics, regulatory issues, and cost sensitivity.

Metro systems are broadly categorized as "metro access or edge" (20 to 50 km) and "backbone or express" (50 to 200 km). A third category called "regional" is sometimes used for coverage in a greater metro area (200 to 300 km) or parts of a twin city "metroplex."[34] This third category helps classify a small number of large tier-1 cities where the individual ring perimeters may be small, but where the acyclic lightpath lengths may exceed 200 km during a failure with bidirectional line switched ring (BLSR) protection.[35]

Access systems include local exchange carrier (LEC), specific business and residential fiber-to-the-curb (FTTC), and fiber-to-the-home (FTTH) systems in the 0–20 km range; they may be point-to-point (PTP) or point-to-multipoint as in passive optical networks (PON). Additionally, CATV broadband access systems that deliver video and data by placing hybrid fiber coax (HFC) nodes within a few km of the subscriber are often engineered to backhaul traffic to head-end sites that may be as far as 75 km from the fiber nodes. System aspects of metro and access applications are described elsewhere in this book. This section will focus on how the fiber requirements are impacted by network architecture.

3.2. *FIBER REQUIREMENTS*

Metro fiber requirements differ from long-haul in significant ways:

- Unlike long-haul, metro fiber must support multiservice, multiwavelength, multiprotocol traffic at multiple bitrates. These include OC-3 to OC-192, Ethernet, Fast Ethernet, Gigabit Ethernet (GbE), 10 Gigabit Ethernet (10 GbE), ESCON, FICON, Fiber Channel and subcarrier multiplexing, all transparently carried on different wavelengths.
- Metro fiber need to be forward-focused to emerging dense WDM (DWDM) and coarse WDM (CWDM) standards,[36] but also must be compatible with existing systems.

- Although 90% metro rings are less than 100 km in perimeter, maximum lightpath distance between nodes can be as high as 200 km—depending on the protection mechanism and interconnecting mesh. Higher degree of optical transparency is desirable but needs to be balanced with ease and versatility of network expansion and service upgrade.[37]
- The 1310 nm band will continue to remain critical, due to the preponderance of low-cost legacy optics as well as the emerging 10 GbE standards.
- Since the outside plant fixed costs (trenching, ducts, huts) in metro are almost twice as high as those for long-haul,[38] metro fiber is not expected to be replaced frequently. Hence, technological adaptability is the key to forestall obsolescence.

Business and residential access (including broadband HFC access) fiber requirements differ from metro in more subtle ways:

- Low cost is the prime and fundamental driver. So CWDM technologies and fibers that operate with low-cost coupling, sources, and passives are important. Along these lines, any multiplexing scheme that yields fiber-pair-gain advantage, such as a PON system, is preferred.
- (For broadband HFC access) multiple signal formats, e.g., QAM and VSB-AM, may be transmitted over the same fiber, so fiber nonlinearity must be low.

By carefully analyzing how the above requirements translate to different fiber designs, one can identify optimum fiber choices. Sometimes conflicting requirements may not be handled adequately by any one fiber type. In such case, a widely practiced hybrid cabling solution which incorporates two fiber types within a single cable can be quite cost effective.

3.3. FIBER TYPES

The fiber types being considered in metro/access applications are

- SSMF (standard single mode fiber, G652), e.g., SMF-28™, Depressed Cladding SMF, Matched Cladding SMF;
- LWPF (low water peak fiber, G652.5), e.g., AllWave™;
- NZDF (nonzero dispersion fiber, NZDSF G655), e.g., TrueWave® RS, LEAF®;
- NDF (negative dispersion fiber, NZDSF G655), e.g., MetroCor™.

All fibers may be broadly categorized on the basis of their dispersion and loss, shown respectively in Fig. 16 and Fig. 17. Both SSMF and LWPF have

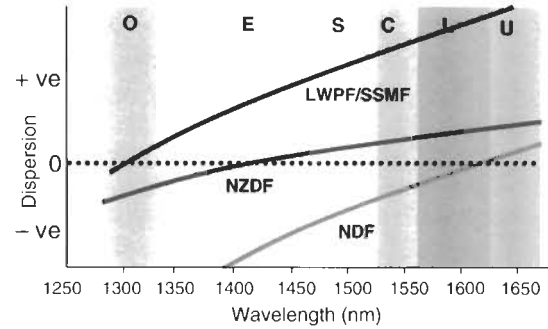


Fig. 16 Dispersion characteristics of fibers, not drawn to scale.

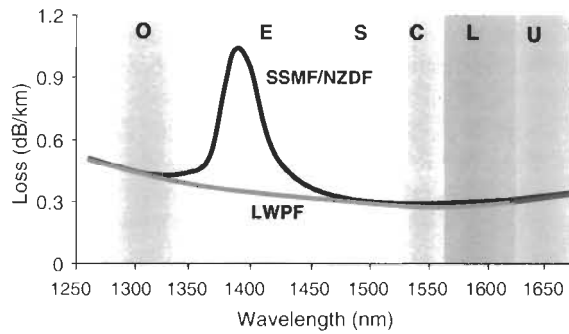


Fig. 17 Loss characteristics of fibers.

the same dispersion profile across the entire transmission band, the magnitude of which is about double the NDF dispersion at 1550 nm and about four times that of NZDF. The NDF has negative dispersion in the entire 1300–1600 nm region, whereas the NZDF crosses zero dispersion around 1450 nm.

The loss characteristics of these fibers is worth discussing due to the revolutionary advancement in the permanent elimination of the 1385 nm water peak in LWPF. Removal of this peak allows use of the low-dispersion 1400 nm region (average dispersion $D \approx 8\,\text{ps/nm}\cdot\text{km}$). For the first time, transmission in the entire 1260 to 1620 nm region is possible, increasing the usable WDM spectral capacity by >50% compared to SSMF. For LWPF, 1385 nm loss is less than 1310 nm loss (0.35 dB/km), and works within the 1310-nm span engineering practices. NDF[39] on the other hand exhibits a higher loss of 0.4 dB/km and 0.5 dB/km at 1385 nm and 1310 nm, respectively, and may need span reengineering.

We note another important fiber characteristic, effective area A_{eff}, which determines the nonlinearity threshold of a fiber.[40] As discussed above in Section 2.7, a small A_{eff} may be beneficial for distributed Raman amplified systems but detrimental as well due to higher four-wave mixing (FWM) and

Raman crosstalk penalties. Hence the nature of the applications, whether regional or broadband access, can guide us to choose a small or large A_{eff} since the impairments due to nonlinearities grow with distance. SSMF and LWPF have a larger A_{eff}, permitting reduced FWM and higher power input to the fiber which helps extend amplifier spacing and system reach for shorter metro and access applications.

3.4. DISPERSION COMPENSATION AND REACH

Like long-haul, metro systems are limited by either loss or dispersion. However, unlike long-haul, metro distances need fewer amplifier cascades and are therefore not limited by optical signal-to-noise ratio (OSNR) due to accumulated spontaneous noise. Dispersion is often the limiting factor for long distances (>200 km) or at high bitrates (~10 Gb/s) in systems that use directly modulated laser (DML) transmitters. Note that electronic regeneration is not required to compensate for loss or dispersion and is only used to demultiplex traffic at a node. To combat dispersion impairment, conventional dispersion-compensating fiber (DCF) may be used to postcompensate spans of SSMF, LWPF, or NZDF, as is done in long-haul systems. NDF, on the other hand, has $D < 0$, and relies on the precompensation of dispersion due to the transient chirp of the transmitter. DMLs typically have transient chirp which blueshifts the leading edge of the pulse and thus induces pulse narrowing in $D < 0$ fiber.[41] Note that this is a different form of chirp from self-phase modulation which redshifts the leading edge of a pulse and increases the power penalty in NDF.[42]

Figure 18 shows the uncompensated reach (maximum transmission distance possible without extrinsic dispersion compensation) as a function of bitrate for externally modulated lasers (EML). Such transmitters, either the electroabsorption type or the Mach-Zehnder lithium niobate variety, have very narrow line width and are essentially chirp-free, and so yield a consistent transform-limit value for the uncompensated reach. Near 40 Gbps and beyond, where DMLs are not suitable, the uncompensated reach falls off dramatically (less than 15 km for nonreturn-to-zero signaling for NZDF).[43] This

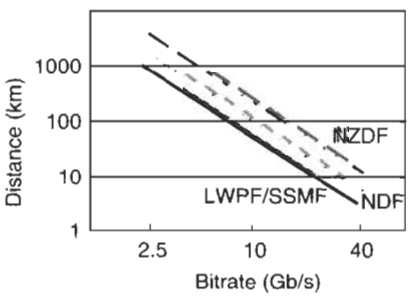


Fig. 18 Uncompensated reach with EMLs (externally modulated laser).

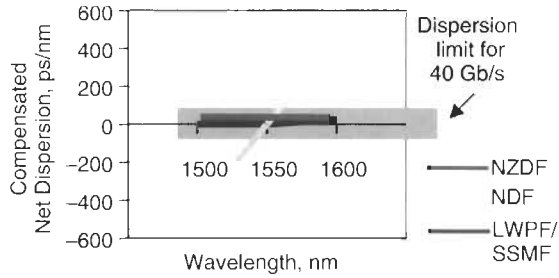


Fig. 19 40 Gbps dispersion compensation (after 100 km propagation and compensation for 1550 nm).

means that dispersion compensation is required as systems upgrade to higher bitrates, no matter what metro distances are considered. Hence, DCFs will be absolutely crucial to precisely compensate the whole EDFA band. DCF is currently available for SSMF, LWPF, and NZDF, but at this time, NDF can only be compensated with SSMF.[44] Since the ratio of dispersion between NDF and SSMF is only about two, long lengths of SSMF fiber are required. Shown in Fig. 19, the compensated net dispersion remains within the narrow ±50 ps/nm tolerance window over the entire 1500 to 1600 nm region for SSMF, LWPF, and NZDF. In contrast, the residual net dispersion for NDF shows unacceptable variation due to poor matching of the dispersion curves with SSMF and presents a significant cost penalty for 40 Gbps DWDM.

Before closing this subsection, we reexamine the use of transient chirp in DML sources to extend reach over NDF. Generally, DML chirp, whether adiabatic, thermal, or transient, differs from laser to laser.[45] To satisfy a given reach, the overall chirp must be carefully selected, requiring extensive screening, binning, and inventory management for each of the many DWDM wavelengths. This process reduces yield and therefore increases DML cost. In addition to cost, the dependence of transient chirp on laser bias, extinction ratio, and the modulating waveform rise/fall times must be addressed.[46] Pending complete characterization of these effects over laser aging and temperature variation (for systems using uncooled lasers), DCF will continue to be the more reliable technique for dispersion compensation in commercial systems.

3.5. FIBER CAPACITY LIMITS DUE TO NONLINEARITIES

Even when chromatic dispersion is adequately compensated, further system impairments are often present. In particular, two issues are important from a metro/access perspective: raw capacity (total bitrate over the system bandwidth) and isolation between different wavelength bands/channels. The former is typically constrained by four-wave mixing (FWM), which limits the number of channels placed within the system bandwidth. The latter includes effects such as Raman scattering which causes crosstalk among signals of dissimilar

modulation format and necessitates high signal-to-noise ratio. Both issues arise from the nonlinear properties of the fiber and are reduced with increasing A_{eff}. Hence, higher A_{eff}, as in SSMF and LWPF, benefits metro and broadband access applications.

FWM is a deleterious nonlinearity that arises from the optical Kerr effect in fiber (the refractive index of the glass varies slightly as the pulse passes) and places an upper limit on the closeness of channels and/or the maximum input power per channel. With the input power limited, FWM thereby limits the distance between amplifiers. Apart from a higher A_{eff}, FWM is also mitigated by higher dispersion values since crosstalk components can walk off the main signal before reception.

Table 2 shows the maximum power input (on per-channel basis) to the fiber as a function of channel spacing in the upper L-band. Compared to NZDF and more importantly NDF, the higher capacity of LWPF and SSMF in the upper L-band (~1620 nm) and the U-band, is clearly evident.

Reduced capacity can also be used to allow higher input power into the fiber, which means longer amplifier spacing and reduced outside plant cost. We remark that tightly spaced DWDM operation is far more desirable and economical in the L- and C-bands (due to the availability of EDFAs) than the 1300 nm or the 1400 nm bands. Hence, ensuring superior L-/U-band capacity/power performance for future growth should be an important criterion for judging metro fibers.

The second type of impairment, crosstalk caused by Raman scattering, occurs when a signal is Raman scattered to generate light shifted upward in wavelength. This results in random crosstalk from the lower signal wavelengths to the upper signals separated by approximately 100 nm. A lower Raman gain coefficient helps reduce crosstalk and is much desired in metro applications, especially when OSNR requirements are high. Since Raman gain coefficient is inversely proportional to A_{eff}, among other factors, LWPF/SSMF are expected to provide ~3 dB improved performance relative to NDF.[40]

Table 2 **Maximum Power Input Thresholds to Keep FWM Impairment Less Than 1 dB. 4 Spans Considered.[43] Dispersion Values at 1605 nm**

	Dispersion (*ps/nm · km*)	A_{eff} (μm^2)	*P(100 GHz)* *dBm*	*P(50 GHz)* *dBm*
LWPF/SSMF	20.5	86.6	16.9	10.9
NZDF	7.4	55.5	10.5	4.5
NDF	−1	51.5	1.5	−4.5

3.6. *THE 1300 nm MARKET CONTINUES TO GROW*

Despite the surge of DWDM in C- and L-bands in the metro backbone, the outlook for the 1300 nm band remains strong in the metro edge and access, driven by:

- the large edge/access market, estimated to be ten times that of the metro backbone, that is almost exclusively dominated by 1300 nm transmission due to availability of low-cost sources, such as uncooled DFBs, Fabry-Perot lasers, LEDs, and the upcoming VCSELs.
- recent standardization in IEEE 802.3 Gigabit Ethernet forum for two 1300 nm solutions for 10 GBE, and the huge market for very short-range (VSR) SONET.[47,48]
- the progress being made toward standardization of a full-spectrum CWDM wavelength grid to address CATV digital reverse, metro access ring and point-to-point applications (enterprise and ISP traffic aggregation) using uncooled lasers and lower-cost mux/demux with relaxed specifications as well as relaxed filter packaging requirements.[49]

It is clear that metro fiber must be compatible with 1300 nm transmission. Otherwise, a huge revenue base will remain untapped, or expensive optical transponder units (OTU) will be required to convert the 1300 nm band to that at 1550 nm. LWPF and SSMF, by virtue of low dispersion in the 1300 nm region, are ideally suited to this wavelength band, and can especially benefit from the negligible chirp afforded by VCSELs and quantum dot lasers.[50] Although insufficient data exists at the present time, NDF performance at 1300 nm may be impaired at high bitrates (~OC-192) owing to the very high dispersion coefficient. NZDF (such as TrueWave® RS with $D = -9$ ps/nm·km @ 1310 nm) can post good performance with 10 Gb/s DML over 60 kms.

3.7. *GETTING READY FOR 1400 nm BAND UTILIZATION*

The inability to utilize a wide portion of the spectrum (1360 to 1460 nm) in SSMF due to the water peak is a constraint for future bandwidth/wavelength expansion. LWPF breaks this bottleneck with full-spectrum CWDM (1260 to 1630 nm in a 20 nm grid spacing,[51] and DWDM possibilities to truly enable "managed wavelength services" as an alternative to, or in extension of, the dark fiber paradigm. By using a wavelength band separator, such as the multi-band module,[52] a carrier's carrier stands to reap revenue from additional λ, with room for 150 100-GHz channels. Although slow to market introduction, several proof-of-concept demonstrations and field trials[53] in the 1400 nm band have captured the essence of why one should exploit the 1400 nm band. These include extra wavelengths in a low-dispersion region enabling high bitrate and reach, reduced cost from low figure-of-merit filters and uncooled sources used

in CWDM, distributed Raman amplification for S-band, and better service and modulation format mix as discussed in Section 3.4. Some 1400 nm components are already being produced and used. With ITU standardization of CWDM,[51] a key enabler of volume—a wavelength grid for lasers—will be in place. Standardization of λs, rather than technology breakthroughs, is necessary for commercialization. With full spectrum flexibility, carriers who build their networks with the 1400 nm capability stand to provide the lowest cost per bit and full spectrum connectivity.[54]

3.8. METRO AND ACCESS FIBER ECONOMICS

The fibers described in Section 3.2 are roughly priced within a factor of two of each other, with SSMF being the cheaper alternative to dispersion-shifted designs, which command a premium. However, it is more important to judge the total system cost, where fiber, DCF, amplifiers, OADM, electronics, and the construction costs are all included. Since each fiber has different dispersion, A_{eff}, and operating bandwidth, there are different cost break points depending on the network topology and assumptions.

Certain broad observations can be made from extensive network cost modeling.[55] For a ring circumference of less than 200 km with 2.5 Gb/s per wavelength traffic, LWPF shows the most favorable economics (factoring in the lack of 1400 nm operation on SSMF), even when including the costs of DCF. Furthermore, the more wavelengths, the better the cost differential compared to NDF; hence growth scenarios to 5 years look economical and promising. Beyond 200 km, where dispersion compensation plays a key role, NZDF fiber is the more economic choice by virtue of its optimal dispersion profile in the EDFA band. If one were to consider 10 Gb/s per wavelength traffic, instead of 2.5 Gb/s per wavelength, 100 km is the distance beyond which NZDF is more economical than either NDF or LWPF.

In the shorter residential and broadband access areas, similar cost advantages are forecast for the LWPF[55] due to the larger number of CWDM channels it can support relative to SSMF. As discussed in Chapter 10, Section 3.1 of Volume IVB, many of the current residential access systems employing PTP and PON architectures may find SSMF adequate, naturally making it the lowest cost option. But as bandwidth upgrades are needed in the future to accommodate Gigabit and 10 Gigabit Ethernet, several interesting CWDM architecture possibilities emerge:

- Upgrade of the ITU standardized (G.983.1 and G.983.2) APON using CWDM in the 1360–1460 nm band.
- Composite PON using broadcast Ethernet in the downlink and switched-Ethernet over CWDM in the uplink. (This architecture permits native Ethernet transport, without having to build additional MAC layer to facilitate upstream TDMA ranging.)

- CWDM in both downlink and uplink. (This architecture is most versatile in that full benefits of a logical PTP such as security and privacy are realized, including cost savings from fewer feeder fibers.)

The architectures described above need the foresight of LWPF deployment today to harvest the system cost reduction later.

3.9. FUTURE TRENDS FOR METRO/ACCESS FIBER AND SYSTEMS

Fiber design advancement for metro and access has not kept pace with those in long-haul, primarily because the long-haul market has been characterized by a drive to lowest-cost-per-bit and increasing capacity with longer reach. This has pushed the limits of available technology, and continued advancement has required newer fiber designs. Although the needs of the metro and access markets are more modest, this does not preclude a metro/access fiber wish list. Among the desired features are:

- a small but finite dispersion value (~2 to 6 ps/nm · km between 1300 nm to 1620 nm), preferably negative but not essential,
- loss spectral characteristics of LWPF,
- A_{eff} characteristic of LWPF/SSMF, and
- all of the above at the lowest cost, in addition to enabling reduced interconnection cost (easy coupling to sources/detectors/passives and relaxed splice/connector tolerances).

Such a solution would entail various tradeoffs, most notably in dispersion and A_{eff}, which is why (as suggested in Section 3.1) a practical approach may be to place two fiber types inside the same cable. For example, a combination of NZDF and LWPF fibers in a hybrid cable can combine the performance of metro regional (>200 km) with that of metro backbone and access (≤200 km).

Many of today's research concepts are likely to see the light of commercial metro/access systems tomorrow. Some of the more well-known advances are

- VCSELs at 1300 nm,[56]
- closer channel spacing with ≤25 GHz in the C- and L-bands,[43]
- chirped-pulse WDM enabling >15,000 channels over the 1280–1630 nm spectrum,[57]
- distributed Raman pumping at the low-water peak to enable better use of S-band.

Such developments require that the fiber plant be compatible with any shift in technology choices, or face a huge cost penalty associated with plant upgrade.

3.10. METRO APPLICATION SPACE

Summarizing the findings of previous sections, one can observe the following about the suitability of various fiber types currently available for the metro and access market:

- NZDF is optimum for DWDM in the C- and L-bands, as traditional long-haul fiber transmission performance has remarkably and consistently demonstrated. It is thus ideal for metro regional market.
- NDF appears to be optimized for C-band for DML operation and is thus suitable for metro regional applications. However, performance in other bands is suboptimal.
- SSMF's performance is a subset of LWPF, hence greenfield deployment of SSMF provides no differentiating technology advantage. A slight first cost advantage is quickly obviated as bandwidth demands grow.
- LWPF is optimized for 1300 and 1400 bands and shares the proven performance of SSMF in C- and L-bands. It is thus ideal for the metro backbone and access markets.

A practical and highly strategic approach is to place two fiber types, such as NZDF and LWPF, inside the same cable to take advantage of optimal performance over a cross-section of distances, bitrates and wavelength windows. Since a metro network is a major investment that must provide returns to service providers for many years, carriers needing fiber for combined metro access and regional rings benefit from including both LWPF and NZDF in their cables.

4. Multimode Fiber Applications

4.1. INTRODUCTION

As transmission moves closer to the user and distances become shorter, the impairments discussed above for long-haul and metro become less significant. A different set of solutions is necessary to address the specific needs of the local area network (LAN) market and those of other short-reach applications, such as storage area networks and equipment room interconnections. As demand for data rates in excess of a Gbps grows for these applications, multimode fiber (MMF) and short-wave VCSELs are emerging as the dominant technologies.

The tremendous growth in demand for high-data rate LANs started with the success of Fast Ethernet in the mid 1990s. The two major developments spurring interest in MMF have been the adoption of the Gigabit Ethernet standard[58,59] by the IEEE in 1998, along with related standards supporting it, and the development of the 10 Gigabit Ethernet standard, which is in draft

as of this writing. Although other technologies have played a role, we will concentrate on Ethernet because it accounts for more than 95% of the LAN market.

The need for ever higher speeds in LANs is driven by the aggregation of traffic from each layer of a switched network to the one above. As data rates at the network stations multiply, so must those of the backbones at the building, campus, and metropolitan levels. There will be four physical medium-dependent (PMD) sublayers in 10 Gigabit Ethernet: 40 km serial transmission over single-mode fiber (SMF) at 1550 nm; 10 km serial over SMF at 1310 nm; wide wavelength-division multiplexing (WWDM) at 1310 nm over either SMF and MMF for 10 km or 300 m, respectively; and serial at 850 nm over MMF of various lengths, depending on fiber type. The SMF PMDs address metropolitan and campus backbones as well as high-speed links to internet service providers. The MMF solutions address building backbones in LANs, storage networks, and equipment interconnections. The reason for the inclusion of multimode PMDs is their lower cost relative to single-mode solutions. The WWDM multimode PMD is aimed at the installed base of MMF, whereas the serial multimode PMD will require a new generation of fiber to reach 300 m. The combination of the availability of inexpensive VCSELs at 850 nm, the loose tolerance packaging allowed by the ease of launching and receiving light to and from MMF, and the large expected volumes of short-reach 10 Gigabit Ethernet ports are believed to make 850 nm serial over MMF the lowest-cost PMD of the four.

There are many effects that can influence the performance of the various components of a MMF LAN link, including the laser, the optical subassembly, the fiber, the various connections in the link, and the detectors. The most important technical development in relation to the fiber in the last few years has been the elucidation of the performance of MMF when excited by laser diode transmitters instead of the previously used LEDs. This change in launch condition forced a reexamination of much of the knowledge base inherited from the LED era. We will concentrate on the new understanding of these issues gained during the development of the Gigabit and 10 Gigabit Ethernet standards.

4.2. MMF BACKGROUND

The vast majority of the installed base of MMF in the United States adheres to a standard design with core diameter of 62.5 μm and maximum fractional index differences of 2%. The performance is specified in terms of overfilled bandwidth, discussed below, with the most common requirements being 160 or 200 MHz-km at 850 nm and 500 MHz-km at 1300 nm. These grades of fiber satisfy a variety of low-speed (10–100 Mb/s) LAN application and structured cabling standards. A small fraction of installed MMF adheres to an earlier standard of 50 μm core diameter and 1% index difference, with bandwidth requirements of 400 or 500 MHz-km at both wavelengths. However,

the basic 50 μm design, with new, very stringent, performance requirements, has been revived as part of the 10 Gigabit Ethernet standard to support serial transmission at 850 nm.

The refractive index profiles $n(r)$ of multimode fibers are chosen to maximize the modal bandwidth, which is accomplished by minimizing the spread in modal group velocities. The optimal profile shape depends on the exact material dispersion characteristics of the doped silica comprising the fiber but is very close to a parabola. Given an index profile $n(r)$, the Maxwell equations may be solved at free-space wavenumber $k = 2\pi/\lambda$ to yield the propagating modes of the fiber. For LAN applications the vector corrections to the scalar wave equation are masked by the much larger differences between the multiple propagating modes, so we may restrict attention to the scalar equation[60]

$$(\nabla^2 + k^2 n^2(r))\psi = \beta^2 \psi \qquad \text{Eq. 6}$$

where ψ is either component of the transverse electric field $\mathbf{e}$ in Cartesian coordinates. The guided modes take the form

$$\mathbf{e}_{l,m,v,\mathbf{p}}(r, \psi) = f_{l,m}(r) \begin{Bmatrix} \cos l\psi \\ \sin l\psi \end{Bmatrix} \begin{Bmatrix} \mathbf{x} \\ \mathbf{y} \end{Bmatrix} \qquad \text{Eq. 7}$$

where $f_{l,m}$ solves the radial part of Eq. 6, $m = 1, 2, \ldots$ indexes the radial solutions at fixed azimuthal index l, v indexes the angular dependence, and $\mathbf{p}$ is the polarization. Each pair l, m corresponds to two degenerate modes (differing in polarization) when $l = 0$, and four degenerate modes (differing in polarization and angular dependence) when $l > 0$, as indicated in Eq. 7.

The degenerate sets of propagating modes indexed by l, m can be further grouped into degenerate mode groups (DMGs), which share a common phase velocity. This degeneracy is only exact in the case of the infinite parabolic profile, but still holds approximately for manufactured profiles. These degenerate mode groups are indexed by the principal mode number μ, defined as

$$\mu = 2m + l - 1 \qquad \text{Eq. 8}$$

The degeneracy of a mode group is twice the group index μ, except for the very highest order mode groups where the cladding plays a role. 50 μm fiber at 850 nm supports roughly 400 modes, including both polarizations and angular dependencies, which can be divided into about 20 such mode groups.

When the endface of a fiber is illuminated by a source of radiation, a field $\mathbf{E}^{(\text{source})}(r, \psi)$ is induced just inside the end face of the fiber. Working with a single Fourier component, this field can, in turn, be expanded in the bound and radiation modes of the fiber:

$$\mathbf{E}^{(\text{source})}(r, \psi) = \sum_{l,m,v,\mathbf{p}} a^{(\text{source})}_{l,m,v,\mathbf{p}} \mathbf{e}_{l,m,v,\mathbf{p}}(r, \psi) + (\text{radiation}) \qquad \text{Eq. 9}$$

The power launched into radiation escapes from the fiber and is lost, while the amplitudes of the guided modes can be computed by overlap integrals. The modal power distribution (MPD) of the source is defined to be the distribution of power launched into the principal mode groups of the fiber

$$P_{\mu}^{(\text{launch})} = \sum_{\substack{l,m,\nu,\mathbf{p}: \\ 2m+l-1=\mu}} \left| a_{l,m,\nu,\mathbf{p}}^{(\text{source})}(r,\psi) \right|^2 \qquad \text{Eq. 10}$$

In communications applications we are interested in the propagation of signals down the fiber. A pulse launched into a guided mode centered at a given frequency will travel at the group velocity of the mode and experience spreading due to chromatic dispersion. Random imperfections in the fiber will couple power between different modes propagating with the same frequency.[61] In modern MMF this power coupling is negligible between modes in different mode groups over lengths of up to many kilometers. The modes within a group, by contrast, typically completely couple within hundreds of meters.[62] Therefore, the pulses in the various modes of a mode group continually share power and ultimately merge into a single composite pulse. Because the coupling is relatively weak at LAN length scales, this composite pulse is broadened by modal, as well as chromatic dispersion.

When a spectrally narrow pulse of light, such as that from a single longitudinal and transverse mode laser, is launched into a 50 μm fiber at 850 nm, we may expect the impulse response of the fiber to be the superposition of about 20 distinct pulses, one per mode group. Each of these pulses travels with its own characteristic group velocity and spreads according to modal and chromatic dispersion. By contrast, when a spectrally broad pulse is launched into a MMF, such as one from a highly multimoded laser or LED, the pulses associated with the various laser modes overlap, and the modal nature of the propagation is blurred. The limiting case of this situation is known as the mode-continuum approximation,[63] which holds for LEDs and highly multimoded lasers. However, for serial 10 Gb transmission at 850 nm the spectral widths of transmitters are limited by chromatic dispersion to the point where the opposite limit holds. When such highly coherent sources are employed with multimode fiber, modal noise becomes a concern. This effect is present when a time-dependent speckle pattern is present in the fiber at points of mode-selective loss, such as imperfect connectors or optical couplers. A coherent source launching into multiple fiber modes ensures the existence of a speckle pattern in the fiber; it can vary with time due to changes in the laser output or mechanical changes in the fiber itself. When present, modal noise must be compensated by increasing the transmitter power.

The salient point in the preceding discussion is that the effective bandwidth of a MMF link is determined jointly by the properties of the transmitter and fiber. The two interact primarily in two ways: chromatic and modal dispersion. In the former, the spectral width of the source and the chromatic dispersion of

the fiber combine to determine the amount by which a pulse spreads during propagation. In the latter, the transverse fields of the transmitter determine the modal power distribution of light launched into guided modes of the fiber. Each guided fiber mode group then propagates with its own group velocity, giving rise to modal dispersion. The effective modal bandwidth of a fiber may thus depend very strongly on its excitation condition.

4.3. FIBER AND SOURCE CHARACTERIZATION

A variety of measurements have been developed over the years to characterize multimode fiber and the sources used to excite them.[61] For our purposes, the most important are various measures of bandwidth, differential modal delays (DMD), and quantities derivable from the near-field intensity (NFI).

The 3 dB bandwidth of a fiber is defined to be the frequency, measured in MHz, at which the modulus of the transfer function of the fiber drops to 1/2 of its peak value.[64] From Section 1.1 it is clear that this value depends strongly on the excitation conditions of the fiber, which therefore must be carefully specified. Historically, the most common launch condition has been that of *overfilled*, in which all but the highest fiber modes are equally excited. More recently, various measures of restricted launch bandwidth have been defined that specify excitation of smaller sets of fiber modes.[65] These measurements are directed at the modal bandwidth, so the spectral widths of the sources used are required to be sufficiently small to avoid substantially biasing the measurement.

A bandwidth measurement attempts to summarize the modal dispersion of a fiber with a single number. Such measurements are useful only if the fiber is used in a manner consistent with the launch used for the bandwidth measurement. A differential modal delay (DMD) measurement,[61,65,67] by contrast, attempts to completely characterize the modal delays of a fiber in order to predict fiber performance under arbitrary launches. Ideally, pulses would be launched into individual mode groups and their delays measured, but such launches are not practical. Instead, a DMD measurement consists of scanning a single-mode fiber across the end face of the multimode fiber under test and recording the temporal response to a short pulse at each of a set of positions. Examples are presented in Fig. 20 and Fig. 21. Under the assumption of complete intragroup mode coupling, such a measurement completely characterizes the modal delays of the fiber under test. The individual group delays are not immediately available from the data, because the launching SMF excites more than one mode group at a time. However, they may be estimated from the raw data.[68] DMD may also be performed in the frequency domain.[69]

Due to the sensitivity of multimode fiber dispersion on launch condition, measurements of the modal power distribution of light launched by transmitters into fiber are required. Analogous to the case of DMD, the ideal measurement would involve the selective detection of individual mode groups.

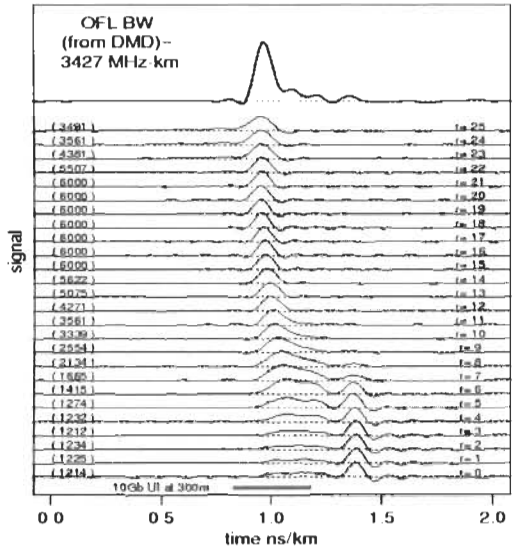


Fig. 20 DMD of a fiber with estimated overfilled bandwidth of 3427 MHz · km, which is well in excess of that required to support 10 Gb operation over 300 m. However, the modal delay structure in the low-order modes is much larger than the bit period, indicated by the solid line at the bottom.

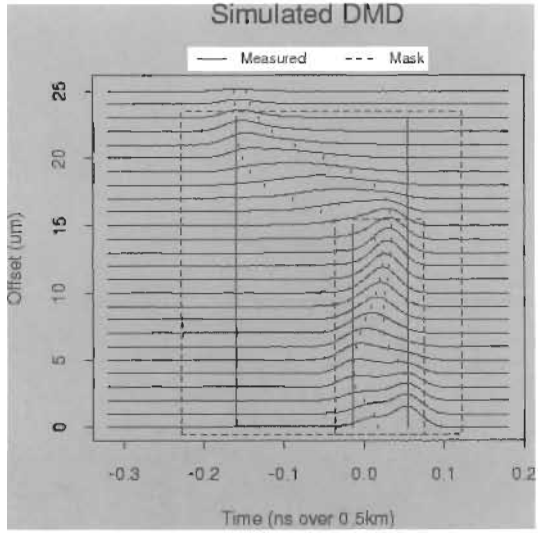


Fig. 21 Results of a simulated differential modal delay measurement, along with measured pulse widths within two specification masks. The measured width of each trace is indicated by the vertical dashes. The reference pulse was 40 ps (FWHM) in this example.

which is not practical. The standard method is instead to obtain the near-field intensity (NFI) of a length of fiber under CW excitation by the source under test, and then process it to estimate the modal power distribution.[70–76] An alternative approach, better suited to high-speed, narrow-linewidth sources, involves temporally separating the modes.[77]

For some purposes it is more convenient to work with the encircled flux (EF) defined as

$$EF(r) = \frac{1}{P_{total}} \int_0^r 2\pi r' dr' I(r') \qquad \text{Eq. 11}$$

rather than the NFI. Here, $I(r)$ is the NFI induced in a probe fiber by the source under test and P_{total} is the total power guided by the fiber.

4.4. THE GIGABIT ETHERNET STORY

Work on the Gigabit Ethernet standard began shortly after the completion of Fast Ethernet in 1995.[78] Much of the early effort of the committee went into ensuring that the known issue of modal noise would not be a problem. However, relatively late in the process an unforeseen problem was uncovered, namely unpredictable bandwidth performance with the use of laser transmitters. This resulted in the formation of the Effective Modal Bandwidth Investigation ad hoc committee. Their results raised many issues about the use of MMF with laser transmitters that influenced the development of the 10 GbE standard, as described below. The bandwidth problem discovered during the GbE standard development was that certain combinations of transmitters and fibers resulted in lower link bandwidth than was expected from the measured overfilled bandwidth of the fiber. Furthermore, the impulse response shapes produced by such combinations were such as to produce unacceptably large jitter. These problems were observed in both laboratory experiments and installed links in the field. This situation was completely unexpected to the committee; indeed, it originally surfaced during a set of round-robin measurements designed to probe how much bandwidth *improves* under laser launch relative to overfilled.

The root of the problem was a mismatch between the way fiber was specified versus how it was used. Historically, standard MMF was specified in terms of overfilled bandwidth. The overfilled launch condition matches the way in which LEDs launch into a fiber, so overfilled bandwidth is a good predictor of fiber performance under LED launch. However, the modal power distributions of the lasers (both edge emitters and VCSELs) used as Gigabit Ethernet sources can be very different from that of LEDs, and in particular can vary dramatically from laser to laser.

The overfilled launch condition puts much more power into high-order modes than low-order modes because of the larger degeneracy of the high-order mode groups. Therefore, a fiber measured to have high overfilled bandwidth is guaranteed to have well behaved high-order mode groups, but

may have very poorly behaved low-order modes. In other words, the overfilled bandwidth measurement is insensitive to lower-order mode behavior. Lasers, as opposed to LEDs, can put most of their power into these low-order fiber modes. Therefore, one should not expect a fiber with known overfilled bandwidth to behave predictably under laser launch. This was the basic fact that was discovered during the development of Gigabit Ethernet. Figure 20 displays the measured DMD of a fiber that has an overfilled bandwidth more than adequate to support 10 Gb operation over 300 m. However, the DMD contains structure that is wide in comparison to the bit period, demonstrating that this fiber would cause a link failure under some restricted launches.

Gigabit Ethernet was designed to support the installed base of fiber. Therefore, the solution to the unpredictable bandwidth problem was to study the distribution of fiber defects that caused this behavior in the installed fiber base, and ensure that laser launches were conditioned to avoid the problem regions in the fiber. The primary culprits were perturbations near the fiber axis that primarily affected the delays of the low-order modes. Such axial perturbations are common in all major MMF manufacturing methods. They are a factor only for launches that concentrate most of their power near the fiber axis, which for Gb sources translates into 1300 nm edge emitters. The solution to the bandwidth collapse problem in this case was to employ a patchcord that offset the launch into the MMF so as to avoid the center. This conditioning was found through simulation and experiment to be sufficient to rescue the standard.[79,80,81] VCSEL sources at 850 nm, by contrast, tend to be highly multimoded, and such conditioning was not found to be necessary.[82,83] Instead, the sources must meet a requirement on the coupled power ratio,[84] defined to be the ratio of power coupled into a single-mode fiber as compared to a MMF.

4.5. TOWARD A 10 Gb ETHERNET STANDARD

The July 2001 draft of the 10 Gigabit Ethernet standard[85] contained four physical medium dependent (PMD) sublayers, of which two employ single-mode fiber for distances of up to 40 km, and the other two employ MMF for LAN applications. One of these latter PMDs consists of serial transmission at 850 nm over a new generation of MMF. The other supports the installed base of MMF via wide WDM near 1300 nm. The former solution is expected to lead to the lowest total costs for new installations as 10 Gb optoelectronics technology matures. Because the WWDM PMD supports the installed fiber base rather than a new generation, we will concentrate on the serial solution.

In 1999 the TIA FO2.2.1 working group began designing specifications for fibers and transmitters that would satisfy the requirements of serial transmission at 10 Gbps. The resulting specifications, largely completed by the summer of 2001, strike a delicate balance between the requirements of the fiber and transmitter manufacturers. In this section we will describe the reasoning that led to the current draft.

In designing a set of specifications, one may trade off fiber and transmitter performance. For instance, if all group delays of the fiber are required to be essentially equal, then there can be no modal dispersion under any launch condition, so any transmitter with sufficiently small linewidth would suffice. Conversely, if the transmitter were required to have a single transverse mode perfectly matched to and focused on the fundamental mode of the fiber, and all connectors were perfect, then the system would be essentially single-mode and link performance would not depend on the behavior of the remaining modes of the fiber.

Neither of the above two extremes are practical in the sense that either would render part of the system economically inviable. The approach taken by FO2.2.1 has been to strike a balance acceptable to both fiber and transmitter manufacturers. In general terms, transmitters are required to avoid launching too much power into those fiber modes that are difficult to control during manufacture, while the remaining fiber modes, which carry most of the power, must be very tightly controlled.

The type of transmitters considered and their spectral characteristics are dictated by economics. The lowest-cost transceiver technology compatible with 10 Gb serial operation is believed to be that of VCSELs operating near 850 nm. The spectral widths must be narrow enough to accommodate the fiber chromatic dispersion at these wavelengths of roughly 100 ps/nm/km over link lengths of up to 300 m. In order to reserve most of the 100 ps bit period for effects other than chromatic dispersion the current draft calls for linewidths less than 0.35 nm (RMS). Because current VCSEL designs result in a spectrum of transverse modes that are separated by at least 0.5 nm, for all practical purposes the allowed devices will be at most few-moded. Left open so far is the question of which modes will be excited and the nature of the optics coupling the light into the fiber. These must be chosen with a view toward both the economic constraints imposed by the fiber and the relatively loose connection tolerances allowed by inexpensive connectors.

The picture of the 10 Gb transmitters that emerged in the previous paragraph has implications for the joint fiber and transmitter specifications. In particular, simulation studies[86,87] have shown that low-order mode VCSEL launches with loose connection tolerances can excite a very broad range of modal power distributions in the fiber. The implication is that very few a priori assumptions can be made about these modal power distributions in choosing specifications on the fiber modal delays. This situation is very different from that of LEDs, which can be assumed to launch the very specific modal power distribution of equal power in every fiber mode (overfilled). Without introducing some specification restricting the transmitters, the only fiber specification possible is one with extremely flat response across all mode groups. Therefore, we are led to designing a pair of transmitter and fiber specifications that restricts both without rendering either economically infeasible.

The first decision in developing a joint fiber and transmitter specification is what measurements to use. Ideally, one would use the modal power distribution of the transmitter and the mode group delays of the fiber, but measuring these quantities directly is impractical. Estimating them from practical measurements, as discussed in Section 4.3, was deemed too involved and prone to error to include in the standard. Instead, it was decided to write the source and fiber specifications directly in terms of DMD and encircled flux data. The benefits in simplicity of this approach were thought to outweigh the loss in discriminatory power of the measurements. Simulations and measurements support the view that the proposed specifications are sufficient to guarantee performance.

The transmitter specification requires that sources avoid the center and the edge of the fiber core, both of which are hard to control in existing multimode fiber manufacturing processes. It accomplishes this by requiring that

$$EF(r_{in}) \leq x_{in} \quad \text{and} \quad EF_{radius}(x_{out}) \geq r_{out} \qquad \text{Eq. 12}$$

where r_{in}, x_{in}, r_{out} and x_{out} are parameters provisionally set at 4.5 μm, 30%, 19 μm, and 86%, respectively. The definition of $EF(r)$ is the encircled flux at radius r, while by EF radius (x) we mean the radius at which the encircled flux first exceeds x. These values are intended to prevent efficient excitation both of the lowest and highest order fiber modes.

The fiber specification takes the form of a constraint on DMD measurements. In its current form, several rectangular "masks" are defined, each consisting of a radial range $[r_{in}, r_{out}]$ and a temporal width Δt_{mask}. The DMD is said to satisfy the mask if the difference Δt_{DMD} in arrival times between the earliest leading pulse and latest trailing pulse represented in the DMD traces with offsets in $[r_{in}, r_{out}]$ is less than Δt_{mask}. The difference Δt_{DMD} is defined as follows. The edges of each DMD trace are defined to be the earliest and latest times at which it attains 1/4 of its maximum power. Then, the measured DMD width Δt_{DMD} within the mask is defined to be the difference between the earliest leading edge and latest trailing edge of DMD traces within $[r_{in}, r_{out}]$. Finally, the width at the same 1/4 threshold of the pulse used to make the DMD measurement (measured by connecting the transmitter and receiver with a strap) is subtracted to yield the estimated maximal delay difference within the mask. Figure 21 shows an example of a simulated DMD measurement to which the following pair of masks has been applied:

$$\max \text{ DMD in } [0, 15] \leq 0.22 \text{ ns/km} \qquad \text{Eq. 13}$$

$$\max \text{ DMD in } [0, 23] \leq 0.70 \text{ ns/km} \qquad \text{Eq. 14}$$

This set of masks, along with the source encircled flux requirement that $EF(16\,\mu\text{m}) > 86\%$, was an early proposal for the standard.[88] This encircled flux requirement is intended to concentrate most of the power within the

smaller of the DMD masks, which has temporal width adequate for 10 Gb/s operation.

The draft fiber specification[88a] that the TIA FO2.2.1 completed in July, 2001 is quite involved and will not be reproduced here. The fiber DMD is required to pass at least one of six sets of five DMD masks. The reason for such a complicated specification is that the various fiber and transmitter manufacturers have all made different process and design choices that favor different regions of the specification space. Therefore, achieving a balance that equally spreads the pain was very difficult.

The proposed fiber specifications are not tight in the sense that there exist physically realizable fibers and transmitters that will pass the specifications but result in bandwidth below the acceptable minimum. There are two reasons for this. The first is a design decision by the committee that a specification resulting in no bandwidth failures would be unnecessarily restrictive, as failure rates of up to 1% are acceptable in the field. This is why the narrowest part of the proposed DMD mask is larger than 0.17 ps/m, the value that would guarantee a passing modal bandwidth of even in the worst case of two equally weighted maximally split pulses. This worst case scenario is very unlikely, and it is argued on the basis of simulation studies that a wider DMD mask will result in link failure rates of less than a percent.

A second reason for link failures when the specifications are met is the decision to write the specifications in terms of the encircled flux and Δt_{mask} rather than the transmitter modal power distribution and group delays of the fiber. These specification measurements do not offer as much resolution as is possible with more sophisticated tools. Furthermore, exactly which set of fiber mode groups are included in the measurement of Δt_{mask} can vary from fiber to fiber, because DMD pulse widths are defined as the times at which a pulse response attains some fraction of its maximum power. This maximum power depends on the relative values of the group delays excited by a DMD launch.

In order to avoid the type of surprise encountered during the Gigabit Ethernet standardization process, and because the proposed transmitter and fiber specifications are not rigorously tight, they were subjected to extensive simulation and experimental testing to ensure they resulted in acceptably low failure rates. The simulation study[77] involved modeling a multimode fiber link including effects of transmitters, fibers, and connections, and computing the effective modal bandwidth and ISI over a large distribution of link components. Figure 22 shows the distribution of calculated link effective modal bandwidths as a function of transmitter encircled flux for fibers that do and do not satisfy a proposed DMD mask. Clearly, the mask performs well in this case, and predicted bandwidth failures for source/fiber pairs that pass the screen are well less than a percent.

The TIA FO2.2.1 organized a series of validation measurements during the fall of 2000 to experimentally verify the performance of the specifications. A sample of 12 fibers and 21 transmitters were prepared by a range of fiber

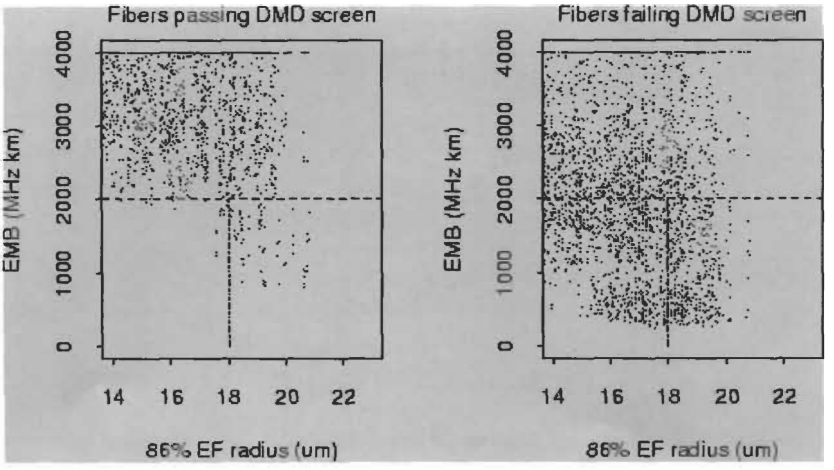


Fig. 22 Results of a simulation study demonstrating the efficacy of a proposed set of transmitter and fiber specifications.

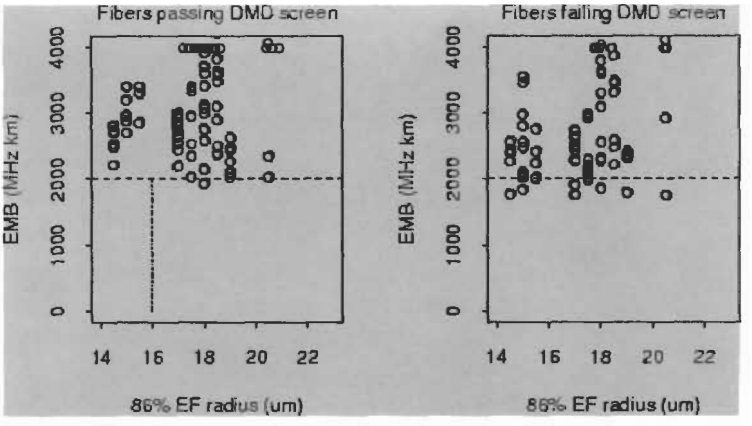


Fig. 23 Data taken by the FO2.2.1 to test the proposed framework for specifying fiber and transmitters.

and transceiver manufacturers in an attempt to capture as wide a range of behaviors as possible within the constraints of available time and resources. Each fiber was characterized by a DMD measurement, each source with encircled flux, and bandwidth and ISI data were taken on a set of transmitter/fiber pairs. In Fig. 23 we show the distribution of measured bandwidths as a function of source encircled flux 86% radius for fibers that do and do not pass the proposed specification (see Eqs. 13, 14) that was current at the time the experiment was planned. In every case, the specification worked as planned: the fiber/source pairs that met the specification exhibited bandwidths greater

than the minimum, while there are instances of low bandwidth among those pairs that failed the specification.

5. Plastic Optical Fiber

5.1. INTRODUCTION

For many years single-mode and graded-index multimode silica optical fibers have been the only practical media for optical networking. These fibers dominate by providing higher bandwidth, extremely low attenuation, and long-term reliability. The high cost and skilled labor required to install silica fiber has not impeded its implementation and optical links are now used extensively in telecommunications and enterprise data networking. However, the demand for bandwidth is now growing in the home, small office, and mobile environments. In these applications it becomes increasingly important to develop very low-cost optical links that can be quickly installed by untrained personnel. Furthermore, these data network links are often integrated in building wiring infrastructures, which makes upgradability of considerable importance. Consequently, an optical fiber medium that offers high bandwidth and ease of installation could begin to supplant copper cabling long before existing bandwidth demands necessitate such a migration.

Plastic optical fibers (POF) have offered the potential for simple, very low-cost optical links since their introduction in the 1960s. Because the elastic moduli of polymers are more than an order of magnitude smaller than that of silica, polymer fibers can have very large optical cores, yet still remain flexible. Furthermore, polymer fibers may be terminated simply by cutting, without the cleaving and polishing required by silica fibers. Consequently, the installed cost of POF systems should be quite low, due both to ease of installation and the relaxed tolerances of mechanical and optical couplings at the fiber endpoints. Because of these advantages, step-index (SI) poly(methyl methacrylate) (PMMA) plastic fibers are now used with some frequency in short distance, low data rate applications where resistance to electromagnetic interference is important. However, there are two problems with SI PMMA fibers for data networking applications. The carbon–hydrogen bonds in PMMA produce strong optical absorption at visible and near-infrared wavelengths.[89] This absorption is caused by overtones of the carbon–hydrogen stretching vibration at 3.2 μm and limits PMMA fiber to a single transmission window near 650 nm. Even within this window, attenuation is intrinsically large (~130 dB/km),[90] limiting maximum link distances to 50 m in most cases. Also, because the fiber is step-index, large intermodal dispersion severely reduces the bandwidth.

The bandwidth limitations of conventional SI PMMA fibers have been overcome by grading the refractive index profile of the core, as is done with silica

fibers. A particularly elegant technique[91] for producing such fibers is known as the interfacial gel polymerization method. In this technique a mixture of methyl methacrylate (MMA) monomer and a nonreactive, index-raising dopant are placed in a PMMA tube and heated. The MMA penetrates the inner tube wall, producing a swollen "gel" phase in which polymerization begins. The larger dopant molecules is partially excluded from the gel phase, and as polymerization progresses inward, the dopant becomes increasingly concentrated in the central portion of the preform. When the preform becomes fully polymerized, the dopant forms a graded-index core, bounded by the original PMMA tube that serves as the cladding. The dopant is effectively immobilized in the PMMA glass at temperatures sufficiently below the glass transition temperature T_g of the dopant-PMMA mixture.

Although graded-index (GI) PMMA fibers fabricated by this method offer reasonable bandwidth, 1 GHz-km or greater,[88] they have only limited utility because they are restricted to operation in the visible portion of the spectrum where few high-speed sources are available.

In 1996, Y. Koike and coworkers[92,93] overcame the wavelength limitations of POF by demonstrating graded-index plastic fibers based on a perfluorinated polymer, poly(perfluoro-butenylvinylether) (PFBVE), commercially known as CYTOP™ (Asahi Glass Co.). This material shows excellent near-infrared transparency because it contains no carbon–hydrogen bonds. While the early PFBVE fibers had negligible absorption losses at most wavelengths, they still exhibited significant attenuation due to scattering, typically around 50 dB/km at 1300 nm. Despite the high attenuation of these fibers relative to silica, PFBVE fibers have revived interest in POF for optical networking, with the expectation that extrinsic sources of scattering in the fibers can be substantially reduced.

The index gradient in PFBVE fibers is formed by partially diffusing an index-raising dopant into the polymer prior to drawing.[94] The resulting index profiles show significant diffusive tails,[95] and are quite different from the profiles required to achieve optimal bandwidth. Despite this shortcoming, surprisingly good overfilled bandwidths, ~300 MHz-km, are usually observed in such fibers.[94] We shall now describe the properties and performance of this new class of plastic optical fiber which has the capability of supporting high-speed data networking applications heretofore reserved for copper and glass fiber media.

5.2. *FIBER GEOMETRY*

A useful geometry for a PFBVE is shown in Fig. 24. The core diameter is ~120 μm, while the cladding diameter is in the range of 160–200 μm. A reinforcing polymer surrounds the cladding that extends the fiber outer diameter to 500 μm. This choice of geometry is dictated by several factors. First, the

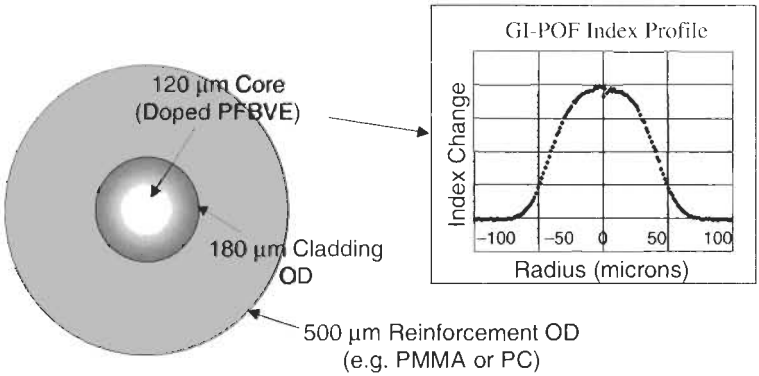


Fig. 24 Geometry and refractive index profile of a perfluorinated GI-POF.

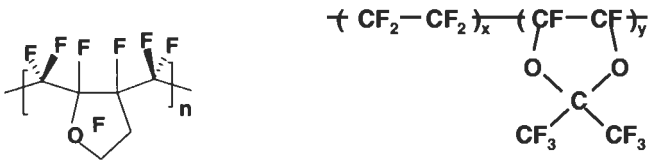


Fig. 25 Commercially available perfluorinated glassy polymers.

core should be large enough to allow interconnection with significant relaxation of the tolerance requirements placed on molded connectors. On the other hand, it should couple to existing high-speed transceivers without degradation of bandwidth performance. Finally, the outer diameter is chosen to provide acceptable load-bearing properties for handling operations such as cabling and duct installation. Inexpensive polymer glasses such as PMMA or polycarbonate comprise the reinforcement polymer. Such materials allow a fiber with a 500 μm outer diameter to withstand loads of up to about 1 kg without undergoing permanent deformation. Furthermore, these materials are inexpensive and allow the very expensive perfluorinated polymers to be restricted to the central waveguiding portion of the fiber.

5.3. *ATTENUATION*

The current revolution in plastic optical fiber is founded on the availability of glassy amorphous perfluorinated polymers, such as those shown in Fig. 25.

Since these materials contain no hydrogen, they show negligible absorption loss in the range of wavelengths preferred for short-distance optical communication, 850–1300 nm. Moreover, these materials contain bulky ring units that serve to frustrate the crystallinity typically observed in fluoropolymers. As a result, these glassy polymers also have low intrinsic scattering losses.

Although both of the polymers depicted in Fig. 25 are commercially available, only the PFBVE polymer has yet been extensively investigated for POF applications.

Although the PFBVE polymer shows no crystallinity, it exhibits measurable intrinsic scattering due to thermodynamic fluctuations of density and of orientational order. When a small-molecule dopant is added to the polymer matrix to produce a graded-index profile, light scattering from thermodynamic fluctuations of dopant concentration is also observed. The light scattering that results from these fluctuations is responsible for the fundamental limits on attenuation in a plastic fiber.

Light scattering experiments[93,96] at 632 nm have been carried out on samples of bulk doped and undoped PFBVE polymer to determine the fundamental limits of attenuation in fibers made with this system. The dopant used[96] was an oligomer of chlorotrifluoroethylene (CTFE). By studying the polarized and depolarized components of scattering in clean bulk samples of undoped and doped (10% by weight) PFBVE, the contributions of density, orientation, and concentration fluctuations to the Rayleigh scattering in these materials were resolved. Assuming the λ^{-4} wavelength dependence associated with Rayleigh scattering, the authors estimated the intrinsic losses of fibers made from these materials to be 9.9 dB/km at 850 nm, and 1.80 dB/km at 1300 nm.

The intrinsic spectral loss curve estimated for a PFBVE POF is displayed in Fig. 26.[95,97]

Spectral loss curves for OH-free single mode silica fiber, PMMA POF, and an early experimental PFBVE GI-POF are also shown for comparison. The losses of the experimental fiber are dominated by extrinsic scattering induced by processing defects, such as geometric perturbations, and/or impurities. These losses have been reduced by recent material and processing improvements so that manufacture of PFBVE GI-POF fiber with losses significantly below 50 dB/km[98] is now commercially feasible.

While absorption bands associated with the carbon-fluorine bonds of the polymer are not generally significant below 1300 nm (in Fig. 26 the small peak at 1280 nm represents the 7th overtone of the fundamental CF stretching vibration of the polymer beyond 8000 nm), other peaks are prominent at

Table 3 **Isotropic and Anisotropic Scattering Losses in PFBVE Polymer, Undoped and Doped with 10% by Weight CTFE.[96]**

Material	*Undoped PFBVE*	*Undoped PFBVE*	*Doped PFBVE*	*Doped PFBVE*
Wavelength	850 nm	1300 nm	850 nm	1300 nm
$\alpha^{(iso)}$	2.4 dB/km	0.44 dB/km	7.1 dB/km	1.29 dB/km
$\alpha^{(aniso)}$	2.8 dB/km	0.51 dB/km	2.8 dB/km	0.51 dB/km
Total Loss	5.2 dB/km	0.95 dB/km	9.9 dB/km	1.80 dB/km

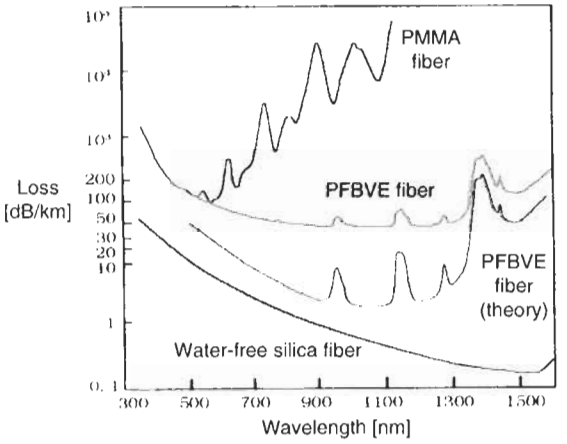


Fig. 26 Spectral loss curves for plastic optical fibers. Adapted from ref. (95).

approximately 945, 1130, and 1385 nm. These absorption bands are associated with molecular water that diffuses freely in and out of the fiber from the atmosphere. The peak at 1385 nm is quite large and serves to limit the range of applicability of PFBVE GI-POF to wavelengths below about 1320 nm.

5.4. BANDWIDTH

The bandwidth properties of GI-POFs are governed by the intermodal and material dispersion of the fiber. Intermodal dispersion is related not only to the shape of the core refractive index profile, but also to the degree of mode coupling and differential mode attenuation that may be present. While the material dispersion of PMMA is higher than that of silica, the PFBVE polymer has a substantially lower material dispersion,[93] leading to predictions of very high bandwidths (~10 Gb/s-km) for perfluorinated GI-POFs having ideal profiles.

A typical index profile for a PFBVE GI-POF is shown in Fig. 24. It is apparent that the profile is far from the ideal parabolic shape. The diffusive tails in the profile at the cladding boundary, which result from the diffusion process used to distribute the dopant to form the fiber core, are very striking. The central portion of the profile also deviates from the near parabolic shape required for high bandwidth silica fibers. Nevertheless, these fibers have quite high overfilled bandwidths of up to 500 MHz-km[98] and support very high data transmission over moderate distances.[99] Differential mode delay (DMD) experiments[100] have been carried out on these fibers to determine the source of the higher-than-expected bandwidths. Very small delay variations are observed among pulses injected into the central portion of the fiber core (roughly half the core diameter), as shown in Fig. 27.

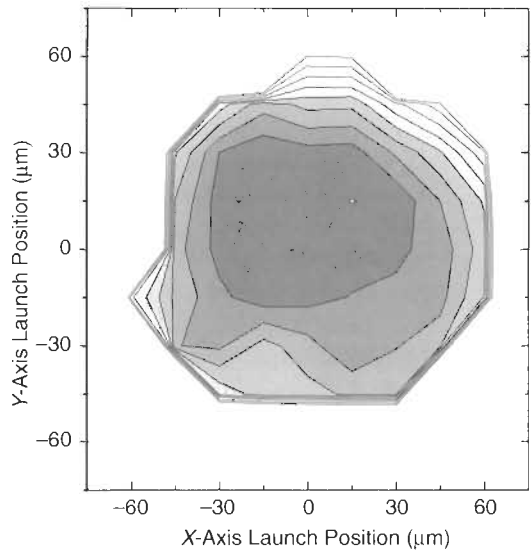


Fig. 27 Variation of pulse delay with launch position for a 118 m long PFBVE GI-POF; the darker contours represent longer delay, and the spacing between contours is 40 ps. From ref. (100).

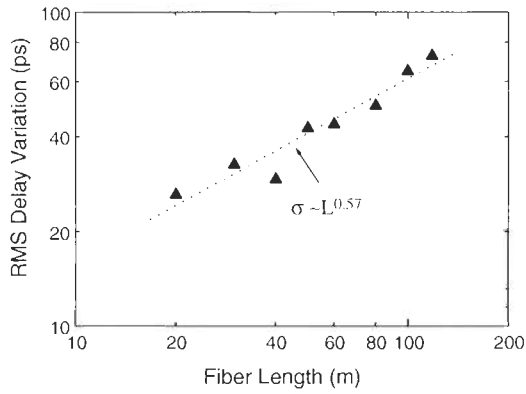


Fig. 28 Dependence of delay variation on fiber length. From ref. (100).

Furthermore, the RMS delay variation σ depends on fiber length in a way that is reasonably well described by a power law, $\sigma \sim L^{0.57}$, as shown in Fig. 28.

From this result alone, one may infer clear evidence of coupling between modes, since uncoupled modes would show delay variations proportional to the first power of length. Indeed, the observed length dependence is quite close to the $\sigma \propto L^{0.5}$ expected from a diffusive theory of mode coupling.[101]

In addition to mode coupling, differential mode attenuation (DMA) has been suggested as an important factor that determines the bandwidth of PFBVE GI-POF.[102] Indeed, DMA has been shown to be a dominant contribution to the bandwidth of GI-POF based on PMMA.[103] While the role of DMA has not yet been adequately quantified for PFBVE GI-POF, it is reasonable to assume that the highest-order modes supported by the fiber exhibit significantly higher attenuation than lower-order modes, owing to the diffusive tails that characterize the core profile. Such tails make the highest-order modes susceptible to attenuation by both macrobending and microbending deformations.

The strong mode coupling observed in PFBVE GI-POF has significant practical consequences. In particular, the very large area of low dispersion observed in the center of the fiber core (see Fig. 27) means that one might obtain a large increase in effective fiber bandwidth simply by restricting the input optical power to this area. While a similar "restricted launch" technique is commonly used to overcome intermodal dispersion in silica multimode fibers, perfluorinated GI-POF allows a qualitatively larger offset tolerance due to the larger core and strong mode coupling. With this restricted launch technique, experiments have been carried out to demonstrate transmission rates in the neighborhood of 10 Gb/s at wavelengths ranging from 850 to 1300 nm.[104,105] Since the material dispersion of the PFBVE polymer is relatively low, the fiber bandwidth is only weakly dependent on wavelength. This situation is a striking contrast to silica multimode fiber, which must be "tuned" to a narrow range of intended operational wavelength by appropriate choice of index profile. However, since the graded index profile is created by a diffused dopant, the effects of long-term aging on the bandwidth of PFBVE GI-POF remains a cause for concern.

5.5. *RELIABILITY*

A significant long-term reliability issue is the stability of the index profile at elevated service temperatures. Because the dopant material is not bound to the polymer, the stability of the index profile depends upon the effective immobilization of the dopant in the glassy polymer matrix. Suppression of dopant diffusion is largely a matter of the size and shape of the dopant molecule and the proximity of the operating temperature to the glass transition of the polymer/dopant mixture. Typical dopants used in GI-POF technology have molecular weights in the range of a few hundred to a few thousand daltons.

Dopant diffusion is essentially Fickian at temperatures well above the glass transition temperature of the polymer/dopant mixture.[106] At lower temperatures, the diffusion of the dopant more nearly follows the dynamics of the molecular motion of the polymer chains. Dopant diffusivities decrease by several orders of magnitude over a small temperature interval encompassing the T_g of the system, e.g., $-10°C$ to $+20°C$,[107] and become strongly dependent

on dopant concentration. At temperatures just below T_g, diffusion becomes greatly hindered and is coupled with molecular relaxation processes of the polymer driving toward equilibrium. As a consequence, the dopant becomes effectively immobilized with dopant diffusivities $< 10^{-15}$ cm^2/sec in the polymer glass at temperatures 20°C or more below the T_g of the dopant/polymer mixture. Thus the index profile of PFBVE GI-POF is typically found to be stable during aging for many months at temperatures up to 20°C below the T_g of the material comprising the highest index region of the fiber core.

Methods have been developed[96,106] to measure the diffusivities of dopants in polymers in the vicinity of the T_g of the dopant/polymer blends. The diffusivities in this region are highly concentration-dependent (non-Fickian). A numerical model which allows the prediction of index profile changes in GI-POFs under accelerated aging conditions has been developed.[106]

5.6. *CONNECTORIZATION*

Historically, the primary technical advantage of plastic optical fiber has been ease and low cost of connectorization. Silica optical fibers typically require relatively expensive connectors as well as trained personnel to perform the polishing and assembly required for field termination of a fiber cable. While very simple tools exist for terminating the conventional step-index POF now available, these termination methods result in relatively high connection loss, typically around 1 to 2 dB per connection. From a geometrical standpoint, the high connection losses of conventional POF are surprising, since the very large core sizes (typically 0.5–1.0 mm) of such fibers should allow excellent coupling.

White et al.[108] have investigated the origin of the high connection losses in SI-POF by examining the end faces of fibers prepared with standard blade cutting tools using confocal scanning fluorescence microscopy. By immersing the fiber end face in a liquid containing a fluorescent dye that can penetrate fine cracks in the surface, a three-dimensional image of the crack structure can be produced. In these experiments it was found that numerous branching cracks extended deeply into the end face, beyond the depth that could be imaged (roughly 60 μm). The cracks form as a result of the tensile stress concentration that occurs at the tip of the blade that is being pushed through the fiber, as shown in Fig. 29.

The tensile stress causes a crack to propagate at near sonic velocities ahead of the blade. This unstable propagation produces branching cracks that penetrate deeply into the fiber surface and cause significant scattering of rays propagating from the fiber end face with consequent high connection losses.

Unstable crack propagation can be suppressed by applying a slight longitudinal compressive load to the fiber during cutting.[109] Figure 30 shows the distribution of connection losses measured in PFBVE fiber for which the ends were prepared using the compression cutting method without any additional

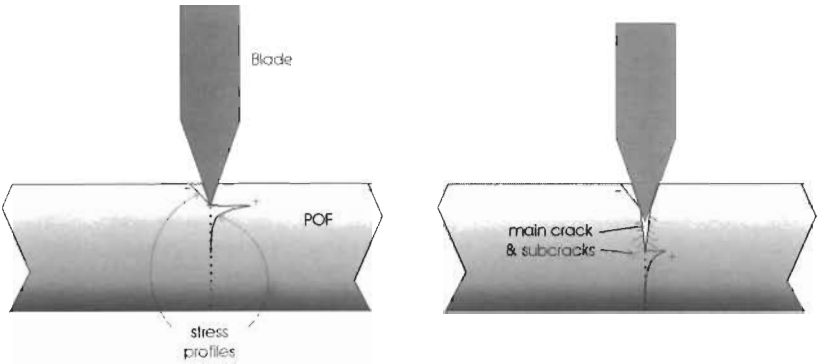


Fig. 29 Stress concentration at the blade tip during cutting of a plastic optical fiber. From ref. (109).

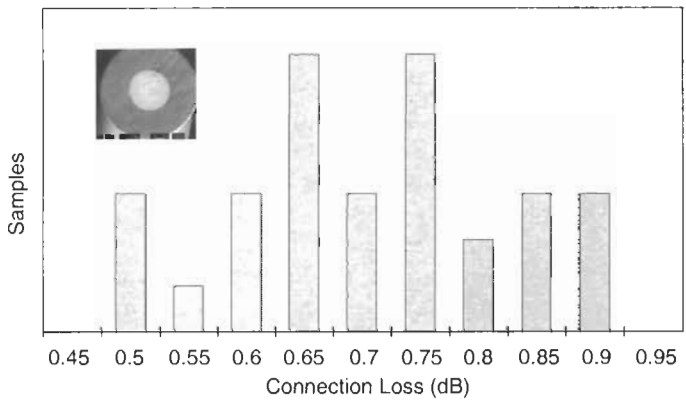


Fig. 30 Histogram of connection losses observed for fibers terminated by cutting under 1% compressive strain. In these experiments 30 compression-cut ends were butt-coupled to a polished fiber end. A scanning electron photomicrograph of a compression cut fiber end is shown in the inset. From ref. (109).

finishing. The losses generally lie in the 0.5–1.0 dB range. The compression cutting method offers the promise that personnel having no special training could connectorize POF in the field quickly with low connection losses, using a simple hand tool.

5.7. APPLICATIONS OF PERFLUORINATED PLASTIC FIBER

The raison d'être for PFBVE GI-POF is to qualitatively simplify field installation of high-speed data links, especially in the gigabit-per-second regime. In comparison to silica-based optical fibers, GI-POF eliminates the need for trained installers and specialized cleaving and polishing tools. In comparison

to high-speed unshielded twisted-pair (UTP) copper links, GI-POF simplifies the cable pulling and termination aspects of the installation process. As with all optical media, GI-POF also eliminates the potential for various forms of electromagnetic interference, which would complicate installation and use of Gb/s UTP copper links. As mentioned earlier, the ease of use of GI-POF will become increasingly attractive as high speed data links migrate outward from controlled, centralized, high-cost environments (telephony and large enterprise networks) into less controlled, fragmented consumer environments such as small office, home, vehicular data networks, and inter/intraequipment interconnection.

In vehicular applications, the older technology of step-index PMMA-based POF has gained significant acceptance beginning in the late 1990s. During these years, PMMA fibers were investigated in a number of aircraft and automotive passenger compartment data systems.[110,111] By the year 2000, every Mercedes-Benz automobile contained a step-index POF data network in the passenger compartment. Since automotive applications involve distances uniformly less than 10 meters, and since link speeds should not exceed 100 Mb/s in the foreseeable future, this older POF technology should continue to be the best fit for this niche. Although passenger aircraft demand much larger bandwidth-distance products than automobiles, aircraft applications also typically have very demanding temperature requirements. Since data links are often routed in the same trays with power cables, continuous operating temperatures of 125°C are typical. As a result, PFBVE-based GI-POF appears not to be suitable in such applications. Although other perfluorinated polymers should permit much higher operating temperatures, those materials have not been adequately developed for POF use at this writing.

In the immediate future, the most promising area for perfluorinated GI-POFs appears to be in equipment interconnection. In these applications, bandwidth demands often exceed a few hundred Mb/s and in many cases are expected to increase rapidly with time. Also, for connections between equipment cabinets, many applications must support remote deployment of different units, so that intercabinet links must be capable of supporting 100–300 meter distances. In many cases however, relatively few links are required, so that bringing in trained personnel to install silica fiber or (where applicable) copper cables and connectors may be prohibitively expensive on a per-link basis. Moreover, trained personnel will be required again whenever the link must be repaired or relocated. In such cases, GI-POF will offer an attractive and affordable way to maintain the flexibility now familiar in such applications, while also permitting much higher data rates.

In the longer term, GI-POF appears to be technically well suited for data networks in homes and small offices. GI-POF can easily support the bandwidths (potentially up to 10 Gb/s) and distances that such applications demand. Also, since these applications typically involve installation behind walls, future-proofing is paramount. Since perfluorinated GI-POF offers

high bandwidth with minimal wavelength dependence, in addition to simplicity and low installed cost, it appears to be extremely well suited for such applications. However, since these types of networks almost always require adherence to LAN applications standards, considerable work in standardization of GI-POF will be required to exploit its technical possibilities in this area.

6. Microstructured Optical Fibers

This chapter, and indeed almost all of this volume, has focused on optical fibers in which light is guided by the slight increase in refractive index afforded by doping silica with small amounts of germanium. However, a new class of fibers has generated enormous interest: microstructured optical fibers that contain air holes which run longitudinally down the fiber. Such fibers can be fabricated by drawing a stack of glass tubes, for example, rather than a solid rod. The presence of holes adds another dimension to fiber design and may lead to radical new developments for controlling and guiding light, resulting in novel fibers for transmission or photonic components.

The design and application of microstructured fibers can be broken into two categories: fibers in which light is guided by a region of higher refractive index than the surrounding material due to the well-known principle of total internal reflection, and fibers in which light is actually guided in a photonic bandgap in which the structure of the holes are such that light cannot propagate in particular directions. A simple example of a bandgap structure is a Bragg grating, which is a one-dimensional photonic bandgap where forward propagation of light in the correct wavelength range is not allowed, causing this light to be reflected. Photonic bandgap fibers are 2-dimensional bandgaps in which light cannot propagate perpendicular to the fiber axis and can thus be confirmed to regions in which the refractive index is lower than the surrounding material. Both sets of microstructured fibers offer unique features and a wide range of potential applications.

6.1. INDEX-GUIDED STRUCTURES

The principles of operation and the characteristics of index-guided microstructure fibers are quite similar to those of conventional fiber, but since the cladding contains holes with a refractive index of 1 (compared to the index of the silica cladding of 1.457 and the germanium-doped core of 1.462 in conventional single-mode fiber) the much greater index contrast and the ability to fill the holes with other materials, such as polymers or liquid crystals, allows exploration of new design space. For example, the high index contrast allows the core to be reduced from about eight microns in conventional fiber to less than one micron. This increases the intensity of the light in the core and enhances

nonlinear effects. A wide spectrum of such effects has been demonstrated,[112] including the generation of a supercontinuum which spans 550 THz, from 400 nm to 1600 nm. This result has important application in optical coherence tomography for imaging in which wide spectral range translates directly into increased spatial resolution. Another potentially important device uses Raman scattering, discussed in Section 2.7, to change the wavelength of a signal. Here, a pulse launched into a high-index contrast microstructure fiber experiences Raman gain, with the shorted wavelength of the pulse amplifying the longer wavelength, resulting in a wavelength shift of the pulse. For input pulses of 100 fs duration with average power of tens of milliwatts, a wavelength shift of more than 150 nm has been achieved in about 20 cm of fiber, with the amount of shift proportional to the input power, as shown in Fig. 31.[113]

Additional effects have been demonstrated which exploit the overlap of the optical field with the holes. For example, microstructured fibers have exhibited endlessly single-mode behavior in which only the fundamental mode is allowed to propagate for all wavelengths of light.[114] If the holey region encompasses more than about 20% of the fiber cross-section, index-guided microstructures exhibit interesting dispersive properties which may find application as either dispersion-compensating or dispersion-managed fibers.[115] In addition, by filling the holes with polymers or liquid crystals, external fields can be used to dynamically alter the properties of the fiber. For example, the temperature sensitivity of a polymer filling the holes surrounding a core can be used to tune a grating written into the core.[116] This effect has been used to increase the tuning range of a long-period grating by more than an order of magnitude.

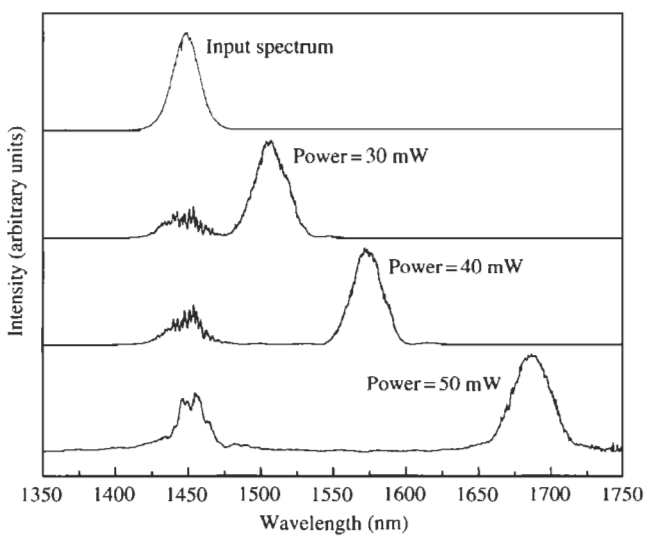


Fig. 31 Intrapulse Raman scattering is used to shift the wavelength of light.

6.2. *PHOTONIC BANDGAP FIBERS*

If the features in the cladding meet certain criteria for refractive index, size, and placement, light can be guided by a photonic bandgap effect in which propagation of certain wavelengths along certain directions is forbidden. This forces light to propagate along the fiber with properties different than would be obtained without the bandgap. For example, concentric rings of alternating high and low index materials surrounding a hollow core have been predicted to confine light to propagate through the hollow region,[117] though such a structure would very quickly radiate light in the absence of the resonance established in the cladding. Photonic bandgap fibers can guide light in regions with higher refractive index (as do index-guided fibers) or lower-index regions,

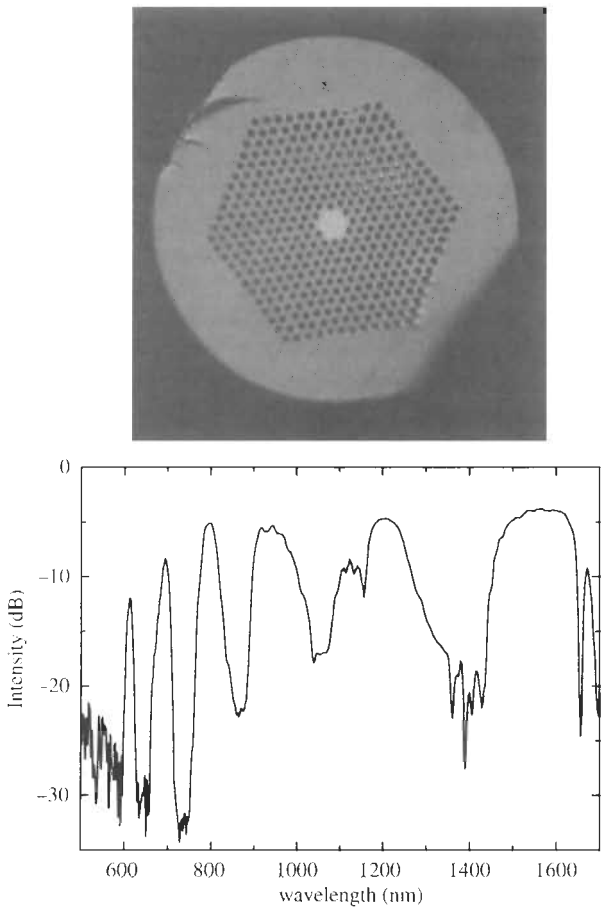


Fig. 32 Photonic bandgap structure shown on top, with resulting transmission spectrum on bottom.

with the latter currently of more interest, for reasons discussed below. A typical structure consists of a periodic array of holes in a silica matrix, with a defect present in the array to strongly confine the light, as shown in Fig. 32.

Understanding of photonic bandgap structures has evolved since they were first proposed[118] but application to optical fibers was inhibited due to lack of appropriate theoretical tools for describing their behavior and properties. Conventional means of modeling waveguides as low-index contrast dielectric structures using scalar equations is not sufficient. Instead, a range of techniques have been developed to handle the high index contrast and the defect present in a periodic array of holes.[115,119,120] Much current work has focused on predicting the geometrical requirements, which typically are expressed in terms of the ratio of the size and spacing of the holes to the wavelength of light. The primary goal has been to find designs that use the simplest material, silica, and allow propagation in air. Much of this book is concerned with understanding and indeed overcoming the properties of index-guided fibers. Effects such as attenuation, dispersion, and nonlinearity dominate the architecture of telecommunication systems and their components. A fiber in which light is guided predominantly in air would have radically different properties, and one could imagine significantly lower attenuation and nonlinearity. Such a fiber would dramatically change the nature of long-distance networks, possibly altering the basic tenets of much of this volume.

Acknowledgments

The author benefited greatly from knowledgeable input from Paul F. Kolesar and A. John Ritger in preparing this chapter.

Acronyms

ASE	Amplified spontaneous emission
CWDM	Coarse WDM
DCM	Dispersion compensation module
DGEF	Dynamic gain equalizing filter
DMA	Differential mode attenuation
DMD	Differential modal delays
DML	Directly modulated laser
DRA	Distributed Raman amplification
DWDM	Dense wavelength-division multiplex
EDFA	Erbium-doped fiber amplifier
EF	Encircled flux

EML	Externally modulated laser
FTTC	Fiber-to-the-curb
FTTH	Fiber-to-the-home
FWM	Four-wave mixing
GbE	Gigabit ethernet
GI	Graded index
HFC	Hybrid fiber coax
ISI	Intersymbol interference
LAN	Local area network
LEC	Local exchange carrier
LMF	Large mode field
LWPF	Low water peak fiber
MMF	Multimode fiber
MPD	Modal power distribution
NDF	Negative dispersion fiber
NFI	Near-field intensity
NRZ	Non-return to zero
NZDF	Non-zero dispersion shifted fiber
OADM	Optical add/drop multiplexer
OSNR	Optical signal-to-noise ratio
PFBVE	Perfluorinated polymer, poly perfluoro-butenylvinylether
PMD	Polarization mode dispersion
PMMA	Polymethyl methacrylate
POF	Plastic optical fiber
PON	Passive optical network
RDS	Relative dispersion slope, RDS = D/S
RSF	Reduced slope fiber
RZ	Return to zero
SI	Step index
SPM	Self-phase modulation
SRS	Stimulated Raman scattering
SSMF	Standard single-mode fiber
TDM	Time-division multiplexing
WWDM	Wide wavelength-division multiplexing
XPM	Cross-phase modulation

References

[1] A. Dwivedi, et al., "Value of Reach Extension in Long-Distance Networks," National Fiber Optics Engineers Conference, p. 351, Denver, CO (2000).

[2] Z. Gills and A. Judy, "New Requirements for Long Haul Optical Transmission Fiber," National Fiber Optics Engineers Conference, p. 20, Denver, CO (2000).

[3] B. Mikkelsen, "160 Gbit/s single-channel transmission over 300 km nonzero-dispersion fiber with semiconductor based transmitter and demultiplexer," 25th European Conference on Optical Communication, PD2.3, Nice, France (1999).
[4] J. L. Zyskind, et al., "High Capacity, Ultra-Long Haul Transmission," National Fiber Optics Engineers Conference, p. 342, Denver, CO (2000).
[5] B. Zhu, et al., "800 Gb/s NRZ Transmission over 3200 km of Truewave® Fiber with 100-km Amplified Spans Employing Distributed Raman Amplification," 26th European Conference on Optical Communication, 2.2.3, Munich, Germany (2000).
[6] R. Ohhira, et al., "Novel RZ signal format with alternate-chirp for suppression of nonlinear degradation in 40 Gb/s based WDM," OFC 2001, Anaheim, CA, WM2 (2001).
[7] W. Idler, et al., "Vestigial Side Band Demultiplexing for Ultra High Capacity (0.64 bit/s/Hz) Transmission of 128 × 40 Gb/s Channels," OFC 2001, Anaheim, CA, MM3 (2001).
[8] Y. Emori and S. Namiki, "100 nm bandwidth flat gain Raman amplifiers pumped and gain equalized by 12-wavelength-channel WDM high power laser diodes," Elec. Lett., **35** (16), 1355 (1999).
[9] L. Gruner-Nielsen, et al., "Dispersion Compensating Fibres and Perspectives for Future Developments," 26th European Conference on Optical Communication, 2.4.1, Munich, Germany (2000).
[10] M. Hirano, et al., "Novel Dispersion Flattened Link Consisting of New NZ-DSF and Dispersion Compensating Fiber Module," 26th European Conference on Optical Communication, 2.4.2, Munich, Germany (2000).
[11] S. Ramachandran, et al., "All-fiber, grating-based, higher-order-mode dispersion compensator for broadband compensation and 1000-km transmission at 40-Gb/s," 26th European Conference on Optical Communication, PD2.6, Munich, Germany (2000).
[12] P. B. Hansen, et al., "Rayleigh Scattering Limitations in Distributed Raman Pre-Amplifiers," IEEE Photon. Technol. Lett., **10** (1), 159 (1998).
[13] H. Kidorf, et al., "Performance improvement in high capacity, ultra-long distance, WDM systems using forward error correction codes," Optical Fiber Communication Conference 2000, Baltimore, MD, ThS3 (2000).
[14] O. Ait Sab and V. Lemaire, "Block Turbo Code Performances for long-haul DWDM optical transmission systems," Optical Fiber Communication Conference 2000, Baltimore, MD, ThS5 (2000).
[15] S. Mettler, et al., "Splicing considerations for non-zero dispersion fiber," National Fiber Optics Engineers Conference, A1-3, Denver, CO (1999).
[16] J. Trewhella, et al., "Performance Comparison of Small Form Factor Fiber Optic Connectors," IEEE Transactions on Advanced Packaging, **23** (2), 188 (2000).
[17] A. H. Gnauck and R. M. Jopson, "Dispersion Compensation for Optical Fiber Systems," *Optical Fiber Telecommunications*, **IIIA**, edited by I. P. Kaminow and T. L. Koch, 167 (Academic Press, San Diego, 1997).

[18] J. Kani, et al., "Triple-Wavelength-Band WDM Transmission Over Cascaded Dispersion-Shifted Fibers," IEEE Photon. Technol. Lett., **11** (11), 1506 (1999).

[19] N. S. Bergano, et al., "Margin Measurements in Optical Amplifier Systems," IEEE Photon. Technol. Lett., **5** (3), 304 (1993).

[20] M. Eiselt, "The Impact of Nonlinear Fiber Effects on Fiber Choice for Ultimate Transmission Capacity," OFC '99, San Jose, CA, TuE1 (1999).

[21] P. V. Mamyshev and N.A. Mamysheva, "Pulse-overlapped dispersion-managed data transmission and intrachannel four-wave mixing," Opt. Lett., **24** (21), 1454 (1999).

[22] B. Zhu, et al., "Experimental Investigation of Dispersion Maps for 40 × 10 Gb/s Transmission over 1600 km of Fiber with 100-km Spans Employing Distributed Raman Amplification," Optical Fiber Communication Conference 2001, Anaheim, CA, TuN3 (2001).

[23] Z. Gills, et al., "Impact of cross phase modulation and other nonlinearities on a high speed digital transmission system incorporating various dispersion management schemes," National Fiber Optics Engineers Conference, **2**, 607, Chicago, IL (1999).

[24] A. Judy, "Optimum fiber dispersion for multi-wavelength 40 Gbit/s NRZ and RZ transmission," 25th European Conference on Optical Communication, II-280, Nice, France (1999).

[25] H. Kidorf, et al., "Pump Interactions in a 100-nm Bandwidth Raman Amplifier," IEEE Photon. Technol. Lett., **11** (5), 530 (1999).

[26] S. T. Davey, et al., "Optical gain spectrum of GeO_2-SiO_2 Raman fibre amplifiers," IEE Poc. J., **136** (6), 301 (1989).

[27] V. Curri, "System advantages of Raman Amplifiers," National Fiber Optics Engineers Conference, B1-1, Denver, CO (1999).

[28] A. Mathur, et al., "Record 1 Watt fiber-coupled-power 1480 nm diode laser pump for Raman and erbium doped fiber amplification," OFC '00, Baltimore, MD, PD15 (2000).

[29] A. F. Judy and C. Kan, "Fiber Effective Area and Raman-Pumped 40 Gb/s Transmission Systems," 26th European Conference on Optical Communication, P1.8, Munich, Germany (2000).

[30] D. Peckham, "Fiber Designs for Undersea Systems: Dispersion Management and Impact on System Capacity," Proceedings of the 6th Fiberoptic Submarine Systems Symposium, June 15–16, 2000, Vancouver, British Columbia, Canada, published by KMI Corp, Newport, RI.

[31] B. Bakhshi, "1 Tbit/s (101 × 10 Gbit/s) transmission over transpacific distance using 28 nm C-band EDFAs," OFC 2001, Anaheim, CA, PD21 (2001).

[32] S. Knudsen, et al., "New Dispersion-Slope Managed Fiber Pairs for Undersea Fiber Optic Transmission Systems," Proceedings of the Suboptic 2001 Conference, May 20–25, 2001, paper T4.2.2, Kyoto, Japan.

[33] C. R. Davidson, "High Capacity WDM Submarine Systems," OFC 2001, Anaheim, CA, WF1 (2001).

[34] T. S. El-Bawab, et al., "Design considerations for transmission systems in optical metropolitan networks," Optical Fiber Technology, **6**, 213–229 (2000).

[35] B. Doshi, et al., "Ultra-long reach systems, optical transparency, and networks," OFC 2001, Tech. Digest (2001).

[36] An amendment to ITU G.671 distinguishes coarse WDM (CWDM) from wide WDM (WWDM). Wavelength spacings in excess of 100 nm are classified WWDM, whereas all intermediate values in between the DWDM grid (200 GHz maximum) and the WWDM grid (100 nm) are collectively specified CWDM. A 20 nm CWDM wavelength grid is the optimal when the total number of CWDM channels over the operating fiber bandwidth is traded for the combined costs of uncooled laser and wavelength mux/demux. Eichenbaum, B., and Das S. K., "Economics for Specifying a Full-Spectrum Coarse WDM Wavelength Grid," to be published, NFOEC 2001.

[37] T. E. Stern and K. Bala, *Multiwavelength Optical Networks: A Layered Approach*, Addison-Wesley, **1**, 221 (1999).

[38] P. Wong, "Network architectures in the Metro DWDM environment," NFOEC 2000, Tech. Proc. (2000).

[39] PI1302, "Corning® MetroCor™ Fiber Product Information," 10/00.

[40] C. Fludger, et al., "An analysis of the improvements in OSNR from distributed Raman amplifiers using modern transmission fibres," OFC 2000, Tech. Digest (2000).

[41] M. Sharma, et al., "Enhancing the performance of directly-modulated laser systems using negative dispersion fiber for metro applications," NFOEC 2000, Tech. Proc. (2000).

[42] F. Forghieri, et al., "Fiber nonlinearities and their impact on transmission systems," *Optical Fiber Telecommunications*, **IIIA**, ed. by I. P. Kaminow and T. L. Koch, 237–239 (Academic Press, 1997).

[43] M. H. Eiselt, "The impact of nonlinear fiber effects on fiber choice for ultimate transmission capacity," OFC 2000, Tech. Digest, 58–60 (2000).

[44] The technology of a negative slope positive dispersion fiber, the kind that will be required to compensate NDF, will remain a significant challenge. If and when such a DCF is realized, the associated cost is likely to be very high as well.

[45] L. Eng, "Time resolved spectroscopy measurement of D2570 chirp," unpublished (2000).

[46] R. A. Linke, "Modulation Induced Transient Chirping in Single Frequency Lasers," IEEE J. Quantum Electron., **QE-21** (6), (June 1985).

[47] Dataquest, "IEEE 802.3ae 10 Gigabit Ethernet worldwide revenue" (December 2000).

[48] Pioneer Consulting, "Optical edge networks: market opportunities for integrated optical network solutions in metro networks" (August 2000).

[49] J. Petiote, "Low-cost components give coarse WDM an edge," Laser Focus (January 2001). See http://www.optoelectronics-world.com.

[50] H. Jones-Bey, "Quantum-dot lasers enter telecom fray for DWDM," Laser Focus, November, 21–22 (2000).

[51] B. Eichenbaum, "Proposal for a frequency plan for optical communications networks optimized for sources where the emission wavelength may drift by several nanometers," ITU submission B01-01-17, SG 15 (2001).

[52] B. Eichenbaum, et al., "Opportunities for revenue enhancement and cost reduction in metro rings and access networks enabled by the emerging standard low water peak fiber and new architectures," NFOEC 2000, Tech. Proc. (2000).

[53] S. Das, et al., "Coarse WDM throughput of up to 20 Gb/s in the 1300–1440 nm region over 63 km of Low Water Peak Fiber," NFOEC 2001, Baltimore (2001).

[54] G. Gilder and Burger, C., "The tunable telecosm," Gilder Technology Report (December 2000).

[55] S. Das and B. Eichenbaum, "Building Metro Fiber Networks to Last: Requirements and Solutions," Lightwave, **18**(10) September (2001).

[56] P. Rigby, "Laser Blazers," see http://www.lightreading.com/document.asp?doc_id = 2580 (2000).

[57] B. C. Collings, et al., "A 1021 channel WDM system," IEEE Photonics Tech. Letters, **12**, 906–908 (2000).

[58] D. Hanson, "Gigabit ethernet optical fiber standard and its impact on the LAN industry," Proceedings of LEOS, 172–173 (December 1998).

[59] IEEE Std 802.3-1998.

[60] A. W. Snyder and John D. Love, *Optical Waveguide Theory*, Chapman and Hall, London (1983).

[61] D. Marcuse, *Principles of Optical Fiber Measurements*, Academic Press, New York (1981).

[62] S. Kawakami and H. Tanji, "Evolution of power distribution in graded-index fibres," Elect. Lett., **19**(3), 100–102 (1983).

[63] Alan R. Mickelson and Morton Eriksrud, "Mode continuum approximation in optical fibers," Opt. Lett., **7**, 572–574 (1982).

[64] Measurement of bandwidth on multimode fiber. TIA/EIA-455-204.

[65] J. S. Abbott, "Characterization of multimode fiber for 10 Gb/s operation," NIST in Special Publication 953 (2000).

[66] M. J. Buckler, "Differential mode delay measurements using single mode fiber selective excitation," Proc. Conf. Precision Electromagnetic Measurements Braunschweig, Germany, June (1980).

[67] TIA/EIA-U55-220, "Differential mode delay measurement of multimode fiber in the time domain" (2001).

[68] Steven E. Golowich, "Analysis of differential modal delay measurements for optical fiber lans," to appear, 2000.

[69] J. B. Schlager and D.L. Franzen, "Differential mode delay measurements on multimode fibers in the time and frequency domains," Technical Digest-Symposium on Optical Fiber Measurements, no. 930 in Special Publications, 127–130, NIST (1998).

[70] Y. Daido, E. Miyanchi, T. Iwama, and T. Otsuka, "Determination of modal power distribution in graded-index optical waveguides from near-field patterns

and its application to differential mode attenuation measurement," Appl. Opt., **18**, 2207–2213 (1979).

[71] S. Piazzola and G. de Marchis, "Analytic relations between modal power distribution and near-field intensity in graded-index fibres," Electron. Lett., **15**, 721–722 (1979).

[72] O. G. Leminger and G. K. Grau, "Nearfield-intensity and modal power distribution in multimode graded-index fibres," Electron. Lett., **16**, 678–679 (1980).

[73] G. K. Grau and O. G. Leminger, "Relations between near-field and far-field intensities, radiance and modal power distribution of multi-mode graded-index fibres," Appl. Opt., **20**, 457–459 (1981).

[74] D. Rittich, "Practicability of determining the modal power distribution by measured near and far fields," IEEE J. Lightwave Tech., **LT-3**, 652–659 (1985).

[75] Mode power distribution and mode transfer function measurement.

[76] Launched power distribution measurement procedure for graded-index multimode fiber transmitters. TIA/EIA-455-203.

[77] Steven E. Golowich, A. John Ritger, and William A. Reed, "A new modal power distribution measurement for high-speed laser transmitters," Proceedings of OFC, Anaheim, CA, March (2001).

[78] David G. Cunningham and William G. Lane, *Gigabit Ethernet Networking*, Macmillan Technical Publishing (1999).

[79] L. Raddatz, I. H. White, D. G. Cunningham, and M. C. Nowell, "Influence of restricted mode excitation on bandwidth of multimode fiber links," IEEE Photonics Technology Letters, **10** (4), 534–536 (1998).

[80] L. Raddatz, I. H. White, D. G. Cunningham, and M. C. Nowell, "An experimental and theoretical study of the offset launch technique for the enhancement of the bandwidth of multimode fiber links," IEEE J. Lightwave Tech., **16** (3), 324–331 (1998).

[81] M. Webster, L. Raddatz, I. H. White, and D. G. Cunningham, "A statistical analysis of conditioned launch for gigabit ethernet links using multimode fiber," IEEE J. Lightwave Tech., **17** (9), 1532–1541 (1999).

[82] Michael J. Hackert, "Characterizing multimode fiber bandwidth for gigabit ethernet applications," Technical Digest-Symposium on Optical Fiber Measurements, no. 930 in Special Publications, 113–118, NIST (1998).

[83] Petar Pepeljugoski, John Abbott, and Jim Tatum, "Effect of launch conditions on power penalties in gigabit links using 62.5 μm core fibers operating at short wavelength," Technical Digest-Symposium on Optical Fiber Measurements, no. 930 in Special Publications, 119–122, NIST (1998).

[84] Optical power loss measurements of installed multimode fiber cable plant. TIA/EIA-526-14A.

[85] IEEE 80Z.3ae Draft 3.2, 2001.

[86] Steven E. Golowich, Paul F. Kolesar, A. John Ritger, and Petar Pepeljugoski, "Modeling and simulations for 10 Gb multimode optical fiber link component specifications," Proceedings of OFC, Anaheim, CA, March (2001).

[87] Petar Pepeljugoski and Steven E. Golowich, "Measurements and simulations of intersymbol interference penalty in new high speed 50 μm multimode fiber links operating at 10 Gb/s," Proceedings of OFC, Anaheim, CA, March (2001).

[88] Steven E. Golowich, Paul F. Kolesar, A. John Ritger, and Giorgio Giaretta, "50 um multimode serial 10 Gb simulations," TIA FO2.2.1 presentation, Quebec City, Canada, June (2000).

[88a] TIA/EIA-492AAAC, "Detail specification for 850-nm laser-optimized, 50-μm core diameter/125-μm cladding diameter class Ia graded-index multimode optical fibers" (2001).

[89] W. Groh, "Overtone Absorption in Macromolecules for Polymer Optical Fibers," Makromol. Chem., **189**, 2861 (1988).

[90] T. Kaino, M. Fujiki, and K. Jinguji, "Preparation of Plastic Optical Fibers," Rev. Electrical Communications Lab, **32** (3), 478 (1984).

[91] Y. Koike, T. Ishigure, and E. Nihei, "High Bandwidth Graded-Index Polymer Optical Fiber," J. Lightwave Technology, **13** (7), 1475 (1995).

[92] Y. Koike, "Progress of Plastic Optical Fiber Technology," Proc. ECOC '96, **MoB.3.1**, 41 (1996).

[93] E. Nihei, T. Ishigure, N. Tanio, and Y. Koike, "Present Prospect of Graded-Index Plastic Optical Fiber in Telecommunication," IEICE Trans. Electron., **E80-C** (1), 117 (1997).

[94] N. Yoshihara, "Performance of Perfluorinated POF," Proc. POF Conference '97, 27 (1997).

[95] N. Sugiyama, "Perfluorinated Polymer Base GI POF for Near-IR Transmission," Polymer Preprints, **39** (2), 1028 (1998).

[96] W. R. White, M. Dueser, W. Reed, and P. Wiltzius, "New Measurement Techniques for POF Material and Performance," Proc. 7th Int. POF Conference '98, Unpublished and Postdeadline Papers (1998).

[97] X. Quan, L. Blyler, and W. R. White, "Plastic Optical Fibers—Pipe-Dream or Reality?," *Optical Polymers: Advances in Optical Polymers and Waveguides*, American Chemical Society Symposium Series (2001).

[98] K. Koganezawa and T. Onishi, "Progress in Perfluorinated GI-POF," Conf. Proc. POF 2000, 19 (2000).

[99] G. D. Khoe, L. Wei, G. Yabre, H. P. A. v. d. Boom, P. K. v. Bennekom, I. Tafur Monroy, H. J. S. Dorren, Y. Watanabe, and T. Ishigure, "High Capacity Transmission Using POF," Conf. Proc. POF 2000, 38 (2000).

[100] W. R. White, M. Dueser, W. A. Reed, and T. Onishi, "Intermodal Dispersion and Mode Coupling in Perfluorinated Graded-Index Plastic Optical Fiber," IEEE Photonics Tech. Lett., **11** (8), 997 (1999).

[101] D. Gloge, "Optical Power Flow in Multimode Fibers," Bell Syst. Tech. J., **51**, 1767 (1972).

[102] T. Ishigure, E. Nihei, and Y. Koike, "High Bandwidth GI-POF and Mode Analysis," Proceedings 7th Int. Plastic Optical Fibres Conference '98, 33 (1998).

[103] T. Ishigure, M. Kano, and Y. Koike, "Propagating Mode Attenuation and Coupling Characteristics of Graded-Index POF," Proc. Int. POF Conference '99, 106 (1999).

[104] G. Giaretta, F. Mederer, R. Michalzik, W. White, R. Jaeger, G. Shevchuk, T. Onishi, M. Naritomi, R. Yoshida, P. Schnitzer, H. Unold, M. Kicherer, K. Al-Hemyari, J. A. Valdmanis, M. Nuss, X. Quan, and K. J. Ebeling, "Demonstration of 500 nm Wide Transmission Window at Multi-Gb/s Data Rates in Low Loss Plastic Optical Fiber," Proc. ECOC '99 (1999).

[105] G. Giaretta, W. White, M. Wegmueller, and T. Onishi, "High speed (11 Gb/sec) data transmission using perfluorinated graded-index polymer optical fibers for short interconnects (<100 m)," IEEE Photonics Tech. Lett., **12** (3), 347 (2000).

[106] L. L. Blyler, Jr., T. Salamon, C. Ronaghan, and C. S. Koeppen, Reliability of Graded-Index Polymer Optical Fibers, in Reliability of Photonics Materials and Structures, E. Suhir, M. Fukuda and C. R. Kurkjian, eds., Materials Research Society Symposium Proceedings, **531**, 107 (1998).

[107] M. T. Cicerone, F. R. Blackburn, and M. D. Ediger, "Anomalous Diffusion of Probe Molecules in Polystyrene: Evidence for Spatially Heterogeneous, Segmental Dynamics," Macromolecules, **28** (24), 8224 (1995).

[108] W. R. White, L. L. Blyler, G. Giaretta, X. Quan, W. A. Reed, M. Dueser, G. Shevchuk, and P. Wiltzius, "Routes to Practical Plastic Optical Fiber Systems," *Linear, Nonlinear, and Power-Limiting Organics: Proceedings of SPIE*, Eich, M., Kuzyk, M. G., Lawson, C. M., and Norwood, R. A., eds., **4106**, 54 (2000).

[109] G. J. Shevchuk, L. L. Blyler, R. L. Decker, and W. R. White, "New Fiber Termination Methods for Low Loss Connections," Proc. Int. POF Conf. '99, Post-deadline Papers, 52 (1999).

[110] D. Seidl, P. Merget, J. Schwartz, J. Schneider, R. Weniger, and E. Zeeb, "Application of POFs in Data Links of Mobile Systems," Proceedings 7th Int. Plastic Optical Fibres Conference '98, 205 (1998).

[111] O. Ziemann, L. Giehmann, P. E. Zamzow, H. Steinberg, and D. Tu, "Potential of PMMA Based SI-POF for Gbps Transmission in Automotive Applications," Conf. Proc. POF 2000, 44 (2000).

[112] J. K. Ranka, R. S. Windeler, and A. J. Stentz, "Visible Continuum Generation in Air-Silica Microstructure Optical Fibers with Anomalous Dispersion at 800 nm," Opt. Lett., **25**, 25–27 (2000).

[113] J. K. Chandelia, B. J. Eggleton, R. S. Windeler, S. G. Kosinski, X. Liu, and C. Xu, "Adiabatic Coupling in Tapered Air-Silica Microstructured Optical Fiber," IEEE Phot. Tech Lett., **13** (1), 52–54 (2001).

[114] T. A. Birks, J. C. Knight, and P. St. J. Russell, "Endlessly Singlemode Photonic Crystal Fibers," Optics Letters, **22**, 961–963 (1997).

[115] J. Broeng, D. Mogilevstev, S. E. Barkou, and A. Bjarklev, "Photonic crystal fibers: A new class of optical waveguides," Optical Fiber Technology, **5**, 305–330 (1999).

[116] A. A. Abramov, A. Hale, R. S. Windeler, and T. A. Strasser, "Widely Tunable Long-Period Gratings," Electron. Lett., **35** (1), 1–2 (1999).

[117] Y. Fink, D. J. Ripin, S. Fan, C. Chen, J. D. Joannopoulos, and E. L. Thomas, "Guiding optical light in air using an all-dielectric structure," J. Lightwave Technol. **17**, 2039–2041 (1999).

[118] E. Yablonovitch, "Inhibited Spontaneous Emission in Solid-State Physics and Electronics," Phys. Rev. Lett., **58**, 2059 (1987).

[119] T. M. Monro, D. J. Richardson, N. F. R. Broderick, and P. J. Bennett, "Holey Fibres: An Efficient Modal Model," IEEE J. Lightwave Tech. **17**, 1093–1102 (1999); J. Broeng, et al., "Photonic Crystal Fiber: A New Class of Optical Waveguides," Optical Fiber Technology **5**, 305–330 (1999).

[120] R. M. Meade, A. M. Rappe, K. D. Brommer, J. D. Joannopoulos, and O. L. Alerhand, "Accurate Theoretical analysis of photonic band-gap materials." Physical Review B **48**, 8434–8437 (1993).

Chapter 3 | New Materials for Optical Amplifiers

Adam Ellison and John Minelly

Corning, Inc., Corning, New York

Introduction

The practical demonstration of the erbium-doped fiber amplifier (EDFA) in the late 1980s led to a great surge in research to improve existing materials and identify new materials for fiberoptic amplifiers and lasers. The motivating factor for material scientists working in this field is the limited application space for silica-based EDFAs. History shows, however, that by the turn of the 20th century many of these limits were overcome by engineering and, in fact, bandwidth demand had not pushed silica-based EDFAs to their limits. To date there is no major commercial amplifier product based on anything other than alumina-doped silica.

Research in new materials continues unabated, however, and if anything it may be accelerating. The main drivers remains the same: the bandwidth available from silica-based EDFAs is not as wide as could potentially be obtained and will soon be exhausted; and rare earth elements besides erbium that can potentially provide additional bandwidth have low quantum efficiency in a silica host. It has long been recognized that once the bandwidth window of the alumina-doped silica (Al/Si) EDFA are used up and new gain windows are needed, then at least for rare-earth doped fiber amplifiers it would be necessary to find alternative glass solutions to extend bandwidth.

Bandwidth extension is not the only application for new materials: non-CVD (chemical vapor deposition) glasses also can enable high-power fiber lasers, particularly cladding-pumped designs. In these, the ability to efficiently capture the high power, multimode ouput of broad stripe diode lasers is paramount. The highest numerical apertures that can be achieved using germania-doped silica are on the order of 0.35–0.4, and these are extraordinarily difficult to process and manufacture. For this application, all-glass double-clad fibers with high-index inner cladding and core facilitate power scaling of otherwise difficult three-level laser transitions.

This chapter reviews work in the field of materials for optical amplification and discusses the application space and context for devices based on these materials. We begin by reviewing some basic material properties required in glasses for active devices, proceed with a review of materials and the methods by which they are drawn to fiber, and conclude with a review of device

OPTICAL FIBER TELECOMMUNICATIONS,
VOLUME IVA

ISBN: 0-12-395172-0

applications. These include extended C-band, L-band, and "super-band" EDFAs, thulium-doped fiber amplifiers (TDFAs) operating in the S-band, and high-power 980 nm ytterbium fiber lasers for pumping EDFAs. Material systems reviewed include fluorides, alumina-doped silica, antimony-silicates, and tellurite.

Considerations for Commercial Applications

For better or worse, telecommunications is an industry rather than an academic concern; thus, amplifier and laser materials brought to market will tend to be those that can be mass manufactured at lowest cost while providing acceptable—hopefully outstanding—performance for the customer. In this section we consider attributes of concern for industrial applications of amplifier and laser materials. This is not intended to be exhaustive, nor does the order represent a decision tree used to weigh the merits of one material against another. It does illustrate, however, why materials that cleave close to existing telecommunications materials (e.g., Al-doped silica) are the main directions pursued today, and also the types of barriers facing new materials.

To be successful in a commercial fiber laser or amplifier application, a new material ought to satisfy the following:

1. *The specific composition should not contain ions known to absorb at the wavelengths of interest.* This criterion serves mainly to limit the range of elements in the periodic table that can be added to a glass to achieve a particular material attribute. It tends to exclude transition metals (particularly the 3d transition metals from Ti to Cu) and materials containing high concentrations of C–H, N–H or O–H bonds. It may also exclude certain combinations of rare earth elements, e.g., in glasses containing high concentrations of rare earth elements.
2. *The material should be adjusted to minimize attributes known to negatively impact amplifier performance for the application in question.* As an example, high optical nonlinearity is a major concern in WDM (wavelength-division multiplexing) applications as it contributes to four-wave mixing and crosstalk. In applications where upconversion is expected, e.g., in upconversion-pumped Tm^{3+} fiber amplifiers and lasers, photosensitivity (either photodarkening or photorefractive effect) is generally to be avoided.
3. *Rare earth element solubility must be high enough to permit practical applications, and rare earth elements should not participate in destabilizing the glass.* This criterion has nearly excluded Ge-doped silica from fiber amplifier applications. Though it is clearly an extraordinary material for obtaining low-loss telecommunications

fiber, refractory rare earth silicate crystals form at high doping levels, and at high temperatures immiscible liquids may form that are either very rich in rare earth elements or rich in germania and silica. Many other glasses show indifferent or low rare earth element solubility or poor glass stability as well.

4. *The rare earth elements in the fiber must be at concentrations low enough to avoid excessive nonradiative losses (e.g., cooperative upconversion for Er^{3+}) but high enough to produce acceptable gain against other sources of passive loss.* Ge-doped silica shows evidence for cooperative upconversion for Er^{3+} at very low concentrations, $<10^{18}$ ions/cm^3, so hundreds of meters of fiber are required to make an amplifier. Even then, fiber attributes such as polarization-dependent loss and optical nonlinearity may overwhelm any advantages conferred by low loss.
5. *There must be a commercially viable pump scheme for the optical transition of interest.* If no diode lasers are produced at the pump wavelength of interest, or if the pump power conversion efficiency is very low, then commercial interest may be small. This is the most important factor limiting deployment of Pr^{3+}-based fiber amplifiers, but it has also impacted deploying materials that are compatible only with the more expensive of two or more pump schemes. An example is Er^{3+}-doped fluorozirconate fiber amplifiers, which must be pumped at 1480 nm because the 980 nm metastable lifetime is very long, $\sim$ 1.5 ms. Er^{3+}-doped aluminosilicate fibers can be pumped at both 1480 and 980 nm, providing greater flexibility in amplifier design and, generally, lower cost.
6. *A reproducible fabrication method must exist that yields as nearly as possible exactly the same result every time it is applied.* This statement is true of commercial chemical vapor deposition processes, but it is useful to remember that these processes were not always reproducible or reliable. Most commercial processes for melted glasses are highly reproducible, but produce thousands to hundreds of thousands of kilograms of glass per day. Producing uniform, high-purity glass reproducibly in the small quantities required for photonics applications is a grave challenge, but not insurmountable.
7. *A suitable core/clad combination exists.* It is probably not an exaggeration to say that the success of Ge-doped silica in telecommunications applications lies largely with the exceptional qualities of its clad glass, pure SiO_2. The problem of identifying an appropriate clad glass for a new, promising core is often ignored until fibers are being fabricated, at which time it may assume crisis-like proportions. As an example of the impact of this problem, commercial development of fluorozirconates was delayed for several years relative to CVD-based materials while researchers sought

suitable means to produce index contrast between core and clad, and clad glasses stable enough to permit fiber draw.

8. *A reproducible fiber fabrication method must exist that readily and reproducibly converts the core/clad combination into low-loss fiber.* "Reproducible" is the key term here, as the rigid standards applied to fiber amplifiers do not permit significant variability in the fiber. This includes the pump absorption length (e.g., dB/m), mode field diameter, core centering for pigtailing; out-of-round for polarization-dependent loss; rare earth element spectroscopy; extraneous contamination during fiber draw; and passive loss.
9. *The fiber produced must have sufficient mechanical and chemical stability to permit deployment in real-world applications.* Again, the success of CVD silica-based fibers relative to fluorozirconates attests to the importance of this criterion. Silica-based fibers have tensile strengths in the neighborhood of 2.5–5 GPa and excellent durability against most kinds of chemical attack (though not hydrogen permeability—see below). Fluorozirconate fibers have tensile strengths between 0.5 and 0.8 GPa and poor chemical durability, especially with regard to corrosion by aqueous solutions.
10. *There must exist a reproducible, low-loss and reliable means to join the fiber in question to conventional telecommunications fibers.* Silica-based fibers are typically joined using arc-fusion splice, which literally melts the ends of the fibers to bond them together. This works for silica because of its low thermal expansion, which means that stresses produced on cooling of the splice are relatively low. Non-silica fibers can be fusion spliced to silica, but the installed stresses tend to be too high for commercial deployment. Fusion splices are quick and cheap, so a compelling reason must be found to overcome this advantage.
11. *The weighted attributes of pump power conversion efficiency gain bandwidth and noise figure must be competitive with any other viable technology for achieving same.* This is a commercial judgment that may involve as much sentiment or anticipation of future trends as objective reality. As an example, most fiber amplifiers employing rare earth elements other than erbium will be compared directly with competing technologies, particularly discrete Raman and semiconductor diode amplifiers. Even the band where amplification is to be obtained may be called to question; for instance, the S-band coincides with the pump lasers for distributed Raman pump lasers for the L-band.

This said, there is still great value in pursuing new hosts for amplifiers and lasers. At the very minimum, it expands the options available to solve various telecommunications problems. More generally, research in new amplifier and laser materials teaches us about the materials already in use and ways to make

them still better. Finally, the spinoff applications for new amplifiers and lasers may, in some cases, be just as interesting as the applications for which they were originally envisioned. As but one illustration of this, research into chalcogenide fibers as host for Pr^{3+} has not produced viable amplifiers, but has produced viable fiber for long wavelength infrared sensor applications.

Common Limitations

All optical amplifiers operating in the telecommunications window share a need to combat two problems: the presence of hydroxyl ions (OH^-) and pairwise interactions between rare earth ions, generally referred to as "clustering."

MINIMIZING OH^-

Humbach et al.[1] provide a recent review of the relative positions and intensities of various OH^- vibrational modes and overtone and combination bands in fused silica. Assignment of bands in other hosts is made largely by analogy with the frequencies seen in silica. The mode of greatest importance is the first overtone of the fundamental stretching mode of the O–H bond, usually denoted $2\nu_3$. The fundamental itself is generally located between 2.7 and 3.2 μm in oxide glasses, depending on which specific cation the OH^- is bonded. Twice this energy (half the wavelength) corresponds to an absorption between 1350 and 1600 nm, right in the heart of the telecommunications band for S-, C- and L-band amplification. Humbach et al. report that the extinction coefficient of $2\nu_3$ is approximately $0.0062 \times \nu_3$ in fused silica. In antimony silicates (Sb silicates), the extinction coefficient of $2\nu_3$ is approximately $0.002 \times \nu_3$. Tellurites are likely to exhibit extinction coefficient ratios comparable to that of Sb silicates. In any case, an absorption detectable in bulk glass in the vicinity of ν_3 all but guarantees substantial absorption in the telecommunications band, with concomitant impact on laser and amplifier efficiency.

Figure 1 compares absorption spectra of fused silica, aluminosilicate (15 mol% Al_2O_3), antimony silicate, and tellurite glasses over the wavelength range 2–4 μm, as indicated on the lower abscissa. The upper abscissa indicates the approximate position expected for the first overtone (the actual position of the overtone is generally at slightly longer wavelength, 10–40 nm, due to anharmonicity in the OH vibrational potential well). The wavelength of ν_3 in fused silica occurs at approximately 2.72 μm, and $2\nu_3$ appears at 1390 nm. In glasses containing heavier cations (e.g., Sb silicates, tellurites, bismuthates, etc.), ν_3 tends to be broader than in fused silica and generally appears at considerably longer wavelengths, e.g., 2.82 μm for antimony silicates and 3.1 to 3.2 μm for tellurites. The overtones of these heavy-metal-rich glasses correspondingly shift to longer wavelengths, to approximately 1420 nm in Sb silicates and approximately 1540–1580 nm for tellurites. More importantly,

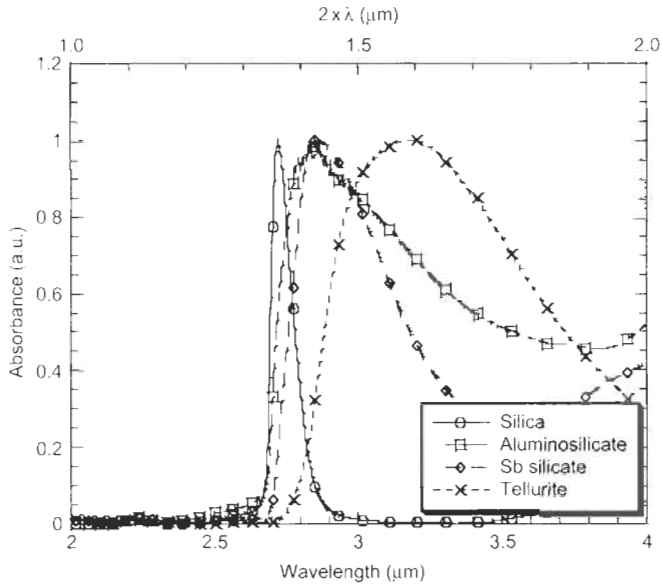


Fig. 1 Normalized absorption spectra of Al-doped silica, tellurite, and Sb-silicate bulk glasses near the OH^- fundamental stretching mode (ν_3). Upper abscissa represents the approximate wavelength of the overtone ($2\nu_3$) found in the telecommunications band.

the wavelength range spanned by the absorption increases considerably from fused silica to Sb silicates to tellurites—hence, the range of effected wavelengths for $2\nu_3$ also increases.

Figure 2 compares the absorption spectrum for the Sb silicate in Fig. 1 with equivalent glasses subjected to increasingly rigorous drying processes. The absorbance per millimeter is defined as β_{OH}, and ten times this value is the approximate absorbance in dB/mm of the band in question. For a β_{OH} of 1.45, the absorbance is 14,500 dB/m at 2.82 μm, corresponding to roughly 30 dB/m peak absorption at wavelengths corresponding to $2\nu_3$. Even at a β_{OH} of 0.014, the implied $2\nu_3$ absorption is still 0.3 dB/m, prohibitively high for applications employing pump or signal at 1.5 microns. As is apparent in Fig. 2, the ν_3 absorption is also very broad. This attribute is shared with its $2\nu_3$ counterpart, which in antimony silicates extends from approximately 1380 to 1530 nm.

Even when ν_3 is approximately 0.0, as indicated for one of the compositions in Fig. 2, it is still possible to pick up water during fiber fabrication, resulting in a persistent unbleachable loss of variable magnitude. Therefore, identifying processes that control water during glass and fiber fabrication is a major point of focus in developing new amplifier materials. Figure 3 shows an example of a passive loss (100 m to 10 m cutback) obtained from an Sb-silicate fiber. The OH^- level in the core was approximately 0.004, whence we would expect a

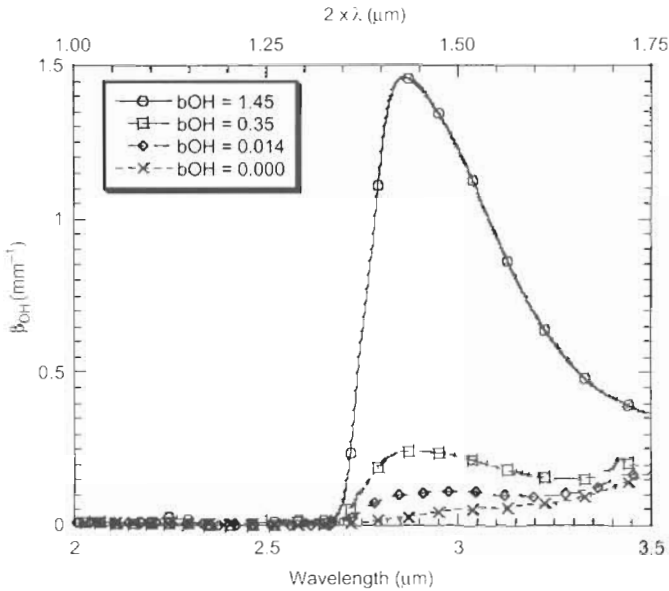


Fig. 2 Absorption spectra of Sb-silicate glasses with varying OH^- levels.

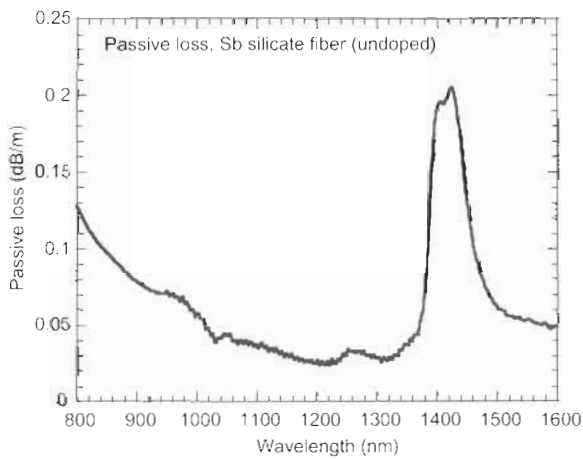


Fig. 3 Passive loss of an Sb-silicate fiber (100 m to 10 m cutback) illustrating the effect of water in the telecommunications band.

passive loss at $2\nu_3$ of approximately 0.08 dB/m above the passive loss. In fact, the passive loss at the peak of the $2\nu_3$ band is 0.21 dB/m, or approximately 0.16 dB/m above the background loss. This increased loss arises from inadequate precautions taken to ensure that the core glass did not experience an environment in which water pick-up was possible.

Water in optical fibers can have an impact exceeding that implied by passive loss, e.g., Fig. 3, because of nonradiative transfer of energy from erbium ions in their excited state to nearby hydroxyl groups. This excites high-lying vibrational modes of the hydroxyl ions, and as they relax they radiate the energy as heat rather than useful light. New experimental data and a review of several studies that have considered this effect are discussed by Houde-Walter et al.[2] Experience with antimony silicates suggests that hydroxyl-related quenching can be sensitively dependent on composition, with certain glasses showing nearly three times higher levels of hydroxyl-related quenching as others with otherwise equivalent OH^- levels in bulk glass and fiber.

Fluoride-based amplifier materials are a special case. Provided that they are melted under inert, low-water atmospheres, fluoride glasses tend to have lower OH^- contents than oxide glasses even when the starting materials have comparable water levels. Furthermore, the most important fluoride glasses for amplifiers and lasers are rich in zirconium, hafnium, indium, gallium, or aluminum, and the ν_3 stretching modes associated with OH^- bonded to these elements are found at longer wavelengths than in the oxide hosts mentioned above. Therefore, carefully prepared fluoride glasses exhibit weak $2\nu_3$ absorptions at the long wavelength end of the telecommunications band, and thus OH^- in fluoride glasses—while no less important than in oxide glasses—tends to be a less significant problem than in oxide glass synthesis.

MINIMIZING CLUSTERING

"Clustering" and "ion-pair formation" are general and not particularly accurate terms describing energy transfer between adjacent rare earth elements. At some level, energy transfer can always occur, provided that light can travel between the rare earth elements before they undergo a change in electronic state. When two rare earth element ions are close enough together, however, the interaction becomes governed by strong dipole–dipole interactions and the probability of energy transfer between the rare earth elements becomes very high. In simple terms, the energy transfer process converts pump photons into light emitted at useless wavelengths or phonon-mediated relaxation of excited states—in other words, heat.

Many studies have evaluated clustering effects on gain and power conversion efficiency in wide range of amplifier and lasers materials and for numerous rare earth elements. This is particularly true for erbium-based devices, for which there is a great deal of experimental data[3–8] and modeling work.[9–15] The proximity of erbium ions to one another is manifested in two ways—homogeneous upconversion and clustering. In the former, upconversion is a simple consequence of having many erbium atoms relatively close together under high pump power. In this case, the energy lost to upconversion processes should vary quadratically with pump power. Clustering of two or more erbium ions in close proximity will tend to produce a nonquadratic dependence of

upconversion rate on pump power, thus producing a pump penalty that may seem disproportionately large given the relatively small number of ions participating in the process.

Decreasing the concentration of the participating ions controls homogeneous upconversion. Upconversion due to clustering may arise from the way the glass or fiber is processed, or from the basic stereochemistry of rare earth element ions in the host. The former can be addressed by changes in process to produce a more uniform initial distribution of rare earth ions. The latter is an intrinsic property of the host material and places a critical limit on the amount of rare earth element ions that can be incorporated before excessive losses due to clustering are produced.

Glass Systems

We now move on to an analysis of various glass forming systems that have been proposed or demonstrated for use in fiber optical amplifiers and fiber lasers. Not all of the systems reviewed below have been demonstrated in a device, although work continues on most of the systems under consideration. Glass chemists tend to define the broadest categories of glasses in terms of the dominant negatively charged species in the glass. These anions are typically present in higher concentrations than any other element in the glass. As an example, in silica, SiO_2, oxygen is the anion, and is present at twice the concentration of silicon. Silica and other glasses in which oxygen is the dominant anion are referred to as oxide glasses. Fluoride glasses contain fluoride ions as the dominant anion, sulfide glasses contain sulfur (S^{2-}), and so on. Sulfide and other chalcogenide (Se, Te) glasses have been considered for many years as potential hosts for Pr^{3+}, but these efforts have not led as yet to fibers with good transparency at 1.3 μm. Therefore, while chalcogenide fibers may yet prove valuable for sensors or for fiber lasers at very long wavelengths, we consider them of marginal value for lasers and amplifiers operating in the telecommunications window and will not consider them further.

We begin our discussion with fluoride glasses, particularly fluorozirconates, glasses in which ZrF_4 is a major constituent. These are not necessarily the most interesting materials for optical amplifiers and lasers, but nearly every glass and fiber fabrication method in use today was either invented for use with fluorozirconates or was applied to them at one time or another.

FLUORIDE GLASSES

Fluoride glasses are promising optical materials because they often have high transparency in both the ultraviolet and infrared. The high ultraviolet transparency compared to multicomponent oxide glasses gives rise to applications for lenses and for upconversion lasers operating in the visible and near UV.

The potentially good infrared transparency results from the infrared extension of the comparatively deep UV absorption edge and the typically low phonon energy of the host glass. Projected losses for fluorozirconate fibers at their heyday were as low as 0.01 dB/km at 2.6 μm. Sadly, losses as low as this have never been obtained. Most of the problem arises not from materials constraints, but the difficulties associated with drawing defect-free fiber from these complicated materials. Nevertheless, fluoride glasses and particularly fluorozirconate glasses have been in the running for commercial deployment in fiber amplifiers and lasers for many years, and the techniques developed to fabricate them and to render them into fiber have benefited research into alternative materials, particularly tellurite glasses.

Fluorozirconates

Poulain et al.[16] reported glass formation in the system $NaF–BaF_2–ZrF_4$ near the composition $25NaF–25BaF_2–50ZrF_4$. Numerous research papers followed from workers at Universite de Rennes[17–22] and around the world (see reviews by Lucas et al.[23,24] and references therein) in which efforts were made to expand the region of glass formation and to stabilize the glasses against devitrification. Ohsawa and Shibata[25] are generally credited with the discovery of the family of very stable fluorozirconates referred to as "ZBLAN." They found that simultaneous additions of small amounts of the fluorides of aluminum and lanthanum to $NaF–BaF_2–ZrF_4$ produced a dramatic enhancement in glass stability. As a result, the typical fluorozirconate glasses used in fiber fabrication activities are often referred to by the acronym ZBLAN (Zr–Ba–La–Al–Na). Optical characterization of fluorozirconate glasses showed relatively deep UV edges and phonon edges shifted to long wavelengths relative to v-SiO_2.[26] As a result, optimum passive losses of ~0.01 dB/km were predicted for fluorozirconate fibers, and significant effort was made to render them into fiber.

There is little question that more effort has been expended to render various permutations of the basic ZBLAN composition into fiber than any material other than germania-doped silica. By 1982 fiber had been produced with losses as low as 21 dB/km,[27] too high for telecommunications applications but adequate for photonic applications (excellent reviews of this early fiber fabrication effort are presented by Tran et al.[28] and Sakaguchi and Takahashi[29]). By 1992, interest in using ZBLAN for telecommunications fiber was fading due to the complexity of drawing long lengths of low-loss fiber, but interest was mounting in using it to make fiber lasers and amplifiers. The optical properties of Pr^{3+}, Nd^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, and Tm^{3+} in ZBLAN were investigated by Wetenkamp et al.[30] They reported that multiphonon decay rates below $3200\,cm^{-1}$ were less in ZBLAN than in various oxide host materials, notably SiO_2 and YAG. This opened the possibility of using ZBLAN as a host for phonon-sensitive rare earth elements, notably thulium and praseodymium, to

make IR-active lasers and amplifiers. Raman spectra of ZBLAN and related glasses[31] show maximum vibrational frequencies of approximately 600 cm^{-1}, implying that they should improve on oxide-based materials as hosts for rare earth elements that are sensitive to high-frequency vibrational modes, notably Pr^{3+} and Tm^{3+}. This same attribute hampers the use of ZBLAN for erbium-doped fiber amplifiers because the 980 nm emission lifetime is very long, ~1.5 ms, and 980 nm pumping is extremely inefficient.

Composition

An approximate reference composition for ZBLAN is $53ZrF_4$–$20BaF_2$–$4LaF_3$–$4AlF_3$–$19NaF$, but there are numerous variations on this basic composition in the literature. In general, glasses that can be successfully drawn to fiber contain between 50 and 60 mol% ZrF_4, 20–30 mol% BaF_2, 0–20% NaF or LiF (the alkali-free glass is generally referred to as "ZBLA"), 3–5% AlF_3, and 3–5% LaF_3 or GdF_3. Glass can be fabricated in the simple systems ZrF_4–BaF_2–NaF and ZrF_4–BaF_2–$(La, Gd)F_3$, but these tend to be less stable than the more complicated compositions. Additions of LaF_3 and AlF_3 suppress the liquidus temperature (the temperature at which crystals first appear on cool-down) in the system, improving glass stability on cooling, and also frustrate crystallization of $BaZrF_6$, the first-forming devitrification phase on reheating.[32] This approach to glass stabilization has been dubbed the "confusion principle."

Changing the identities of various components in the basic ZBLAN composition has been investigated extensively. Lanthanum can be replaced by other large trivalent cations, particularly other rare earth elements. In sodium-free glasses, for example, Gur'ev et al.[33] found that the region of stable glasses shifts away from ZrF_4 and toward RF_3 (where R is the rare earth element) as the ionic radius of the trivalent rare earth element decreases. The extent to which these changes influence the usefulness of the glasses for fiber fabrication was not evaluated. Qui et al.[34] evaluated the ability of different alkaline earth cations to stabilize ZrF_4–BaF_2–ZnF_2–AlF_3–YF_3 glasses. They found that replacing barium with the smaller alkaline earths magnesium, calcium, and strontium caused a large decrease in $T_x - T_g$, the difference between the crystallization temperature (T_x) and glass transition temperature (T_g). As $T_x - T_g$ decreases, the working temperature range in which one can obtain crystal-free fiber also decreases, and hence the compositions have lower thermal stability.

Studies were also performed on relatively stable multicomponent fluorozirconates comparable to the canonical ZBLAN composition shown above. Lebullenger et al.[35] characterized the stability of sodium-bearing and sodium-free ZrF_4–BaF_2–LaF_3–$AlF_3 \pm NaF$ glasses. They found that replacing small amounts of barium with cadmium or lead slightly improved glass stability and increased the refractive index, with the index impact being much larger for lead than for cadmium. Replacing barium with potassium in sodium-free

compositions produced a modest enhancement in glass stability, but replacing sodium by potassium or rubidium in a sodium-bearing glass decreased glass stability. These trends were confirmed in the report of MacFarlane et al.,[36] who also showed that replacement of sodium by lithium produced a modest improvement in glass stability but a decrease in refractive index. Hasz et al.[37] showed that zirconium could be completely replaced by hafnium with no material impact on glass stability, and very little impact on glass properties save that it decreases refractive index of the glass.

In light of these observations, most investigators have approached the problem of creating the refractive index contrast required for core/clad structures by adding a small amount of lead fluoride (PbF_2) to the core glass and partially or completely replacing zirconium with hafnium in the clad glass. A definitive study by Poignant et al.[38] showed that substitutions of PbF_2 and HfF_4 appeared to caused tensile strength of ZBLAN-based fibers to decrease from about 600 MPa to 450 MPa. The source of the decrease was traced to a transition from compressive to tensile stresses across the radius of the fiber, and formation of PbF_2 crystals at high lead concentrations. Nevertheless, numerical apertures (N.A.s) as high as 0.45 were demonstrated, indicating that small core, high N.A. fiber was potentially achievable, an important prerequisite for amplifier and laser applications. Kogo et al.[38a] used a combination of additions of PbF_2 to the core and HF_4 to the clad to obtain a near-perfect mechanical match between core and clad glasses. These were then fabricated into a high-N.A. ($\Delta = 2.9\%$), single-mode fiber (cutoff wavelength ~1200 nm) with a modest passive loss of 0.3 dB/m at 1.3 μm. Results such as these demonstrated the potential for fluoride fibers in fiber amplifiers and lasers.

Glass Synthesis and Purification

Initially, fluorozirconate glasses were fabricated simply by melting the requisite fluoride constituents in suitably inert crucibles (gold, platinum, and carbon) at elevated temperatures, 900–1000°C. It was quickly realized, however, that the stability and purity of the glass was strongly affected by trace constituents. The processes explored or developed for improving purity of starting materials and enhancing glass quality are nearly too numerous to count. They fall into three basic categories: (1) processes used to purify the starting materials; (2) processes used to control reaction of the melted glass with the surrounding atmosphere; (3) processes used to purge the melted glass of transition metals.

Mitachi et al.[39] provided one of the first detailed studies of ways to improve the quality of glasses obtained from as-received starting materials. The methods explored included recrystallization, solvent extraction, chemical vapor deposition, distillation, and sublimation. The authors concluded that recrystallization, distillation, and sublimation were the best approaches for removing transition metal impurities. Surprisingly, despite many years of

detailed research, very little has emerged to change these conclusions, though the sophistication of the approaches has improved considerably.

Robinson[40] and Poulain et al.[41] provide detailed descriptions of processes used to obtain high purity fluoride precursors. In its most basic form a precursor material is reacted with an aggressive fluorination agent to obtain a pure fluoride. In the case of aluminum, the precursor is typically aluminum chloride or chloride hydrate. In the case of zirconium, its oxide (ZrO_2), oxychloride ($ZrOCl_2$),[42] oxynitrate ($ZrO(NO_3)_2$),[43] or fluoride hydrate ($ZrF_4 \cdot H_2O$)[44] are used. These precursors are typically taken through one or the other of two distinct processes. In the first process, the precursors are reacted at elevated temperature with flowing anhydrous hydrofluoric acid (HF), a process that takes the starting materials directly to fluorides. The second is to react the starting materials with an excess of ammonium bifluoride ($NH_4F \cdot HF$) at elevated temperature, followed by a further increase in temperature to burn off traces of ammonium.[45] Lanthanum fluoride (LaF_3) is precipitated nearly stoichiometrically from aqueous solutions by titration of a solution of a soluble salt (e.g., $LaCl_3 \cdot 6H_2O$) with hydrofluoric acid. Sodium fluoride is obtained by direct reaction of its hydroxide or carbonate with HF, followed by addition of alcohol to precipitate the salt.

Zirconium and aluminum fluorides are typically distilled or sublimed to eliminate oxygen (the refractory oxides are left behind) to obtain ammonium fluorozirconate and ammonium fluoroaluminate, which is further decomposed and sublimed. The main risk in the latter approach is retention of ammonium, which causes unwanted absorptions in the final glass. Robinson et al.[46] and Drexhage et al.[47] introduced the now widely-accepted concept of reactive atmosphere processing, wherein the glass is melted in the presence of a gas that is highly reactive with respect to water, strongly fluorinating, or both. In the particular example of the study by Drexhage et al., it was shown that an IR absorption feature near 1400 cm^{-1} attributed to oxygen dissolved in the fluorozirconate glass was substantially removed or eliminated outright by melting in the presence of CCl_4. Other gases that are employed in reactive atmosphere processing are Cl_2, CCl_4, SF_6, NF_3,[48] and CF_4. While originally described as a method to enhance glass melting, reactive atmosphere processing is now typically used on all fluoride precursors prior to melting.

There is a radically different way to obtain the batch materials used for glass melting, which is to obtain the fluorides via sol-gel processing.[49–51] In this approach, organometallic precursors (alkoxides) are mixed with alcohol and combined with an acid solution to promote hydrolysis of the alkoxides (replacement of an alcohol group by a hydroxyl ion) and condensation into a gel. The gel is dried and reacted with anhydrous HF to produce a fluoride batch. The batch is then melted using methods described below. The resulting glasses have no more water or oxygen than conventionally melted materials and in other ways appear to have properties akin to those of conventionally melted glasses. The main advantage to this approach would be to disperse rare

earth dopants at nearly a molecular level, thus reducing the risks associated with rare earth element clustering due to incomplete processing.

However obtained, the purified materials are combined in the appropriate amounts and are melted in a vitreous carbon, platinum, or gold crucible. Melting takes place in an oxygen- and water-free atmosphere, typically in a glove box. Zirconium has a tendency to form reduced species, which in turn show up as dark specks within the glass, particularly if the melting takes place in a vitreous carbon or graphite crucible. Reactive atmosphere processing serves to eliminate reduced zirconium contaminants.

A final purification step may be performed in the crucible itself: Zhiping et al.[52] demonstrated substantial reductions in transition metal impurity levels in ZBLANs by electroplating transition metals out of the melt at relatively low temperature for long periods of time. Murtagh et al.[53] showed that this approach could be extended to other fluoride glasses.

As an alternative to crucible melting, a method has been developed to perform containerless melting of fluoride glass rods, a process called gas film levitation.[54] In this approach, the fluoride melt is suspended over or within a porous membrane through which gas flows. This approach avoids unwanted reactions between the fluoride melt and the crucible material and delivers glass in a form suitable for further processing into fiber. There is some indication that this approach can be used to reduce scattering losses.

There is an entirely different means of obtaining ZBLAN from the methods described above: plasma-enhanced chemical vapor deposition. This method was demonstrated by Fujiura et al.[55] Metal organic precursors are delivered with SF_6, CF_4, or NF_3 into a plasma contained within a ZBLAN tube. The authors report improvements in glass stability relative to glasses obtained from batch melting. Fiber losses were not reported. Sakaguchi has modeled the challenges associated with collapsing the ZBLAN tube prior to fiber fabrication in detail.[56]

Fiber Fabrication

To a first approximation all fibers are produced via one of two routes: the fiber is pulled directly from melted glass, which we will refer to hereafter as the direct draw approach; or the fiber is pulled from a rod of reheated glass, which we will refer to as the preform draw approach. Fiber is the only major glass product that is produced on a batch basis—nearly all glasses produced commercially are made in a continuous process, whereby batch materials are delivered into the hot end of a melter and formed glass comes out the cold end. Direct draw methods can be likened to this sort of continuous process in that one can continue to feed glass into the forming crucible from some other source. Preform draw processes are intrinsically batch-based, and thus there are great challenges associated with guaranteeing uniformity and consistency of the product.

In either a batch or continuous process, the happiness of the glassmaker is best maintained when the glass chemist provides glasses whose viscosity (resistance to flow) varies slowly as a function of temperature, and whose thermal stability on cool-down or heat-up is sufficient to permit a stable process. The viscosity versus temperature curves for ZBLAN and related glasses have been reported by Hasz and Moynihan,[57] Braglia et al.,[58] and by Zhang et al.[59] The viscosity curves reported by Hasz and Moynihan are reproduced in Fig. 4. In most commercial glass melting processes the glass is melted at a temperature corresponding to about 20 Pa · s (200 poise), whereas the glass transition temperature is about 10^{12} Pa · s (10^{13} poise). For v-SiO_2 this viscosity range encompasses a temperature variation of more than 1300°C, and Pyrex™ experiences a similar change in absolute viscosities over a 800°C variation in temperature. As is evident in Fig. 4, fluorozirconate glasses typically pass through this range in viscosities in a span no greater than 200°C, sometimes less. The steepness of the viscosity curve of ZBLAN vs. temperature creates significant challenges for fiber fabrication, particularly as regards the reproducibility of the outcome.

The glass transition temperature for ZBLAN varies with composition, but is typically on the order of 260–300°C; and T_x, the temperature at which crystals first appear on reheating, is typically 350–380°C, depending on the exact composition under investigation. The ideal viscosity at which to draw fiber depends on many factors: the size the preform, the draw speed and tension, the size of the hot spot in the furnace, and the length of the thermal gradient

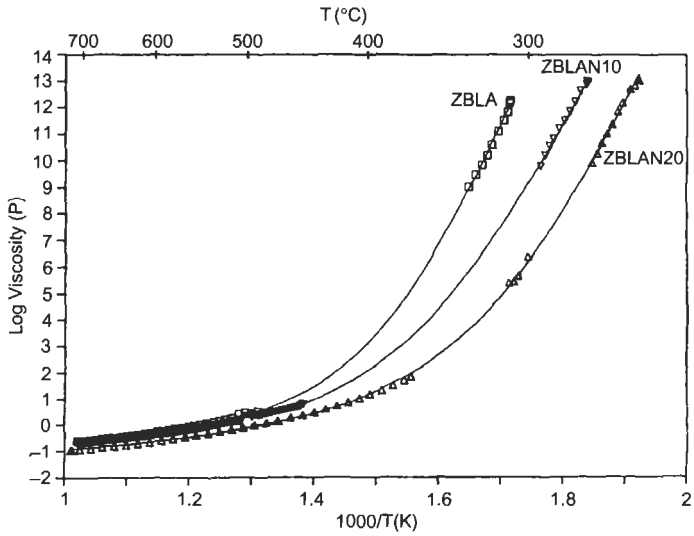


Fig. 4 Viscosity versus temperature curves for ZBLAN and related glasses. From the study of Hasz and Moynihan, reprinted with permission.[57]

away from the hot spot, to name but a few. In general, however, 10^4–10^5 Pa · s (10^5–10^6 poise) will generally work reasonably well. From Fig. 4, these viscosities correspond to temperatures that are fairly close to T_x for ZBLAN. It is not surprising, then, that devitrification during fiber draw is perhaps the single greatest limitation to the widespread use of ZBLAN fiber. On the other end of the temperature/viscosity spectrum, the liquidus temperature for ZBLAN is ~600–700°C, depending on the composition. At these temperatures, molten ZBLAN has a viscosity on the order of 0.1 Pa · s (10 poise), somewhat less than maple syrup. This is far below the viscosity at which one might hope to maintain control during a fiber draw process. It is not surprising, then, that nearly all ZBLAN fiber produced today is prepared by some variation of a preform draw process.

A number of studies have evaluated the time-temperature dependence of ZBLAN devitrification, and a summary of the main draw issues is presented by Hirai et al.[60] Their main results are shown in Fig. 5. The best fiber will have a small bending failure radius and low loss. These conditions are obtained when ZBLAN is drawn to fiber at relatively low temperature, high speed, and high tension. At the other end of the spectrum, if ZBLAN is only briefly cooled below its liquidus temperature, then crystal-free fiber is possible. However, fiber fabrication then must take place at relatively high temperature, at low speed, and at near-zero tension, which greatly complicates the task of maintaining fiber geometry.

Figure 6 is a synopsis of several methods employed in the past. Direct draw methods are considered first. The classic direct draw fiber fabrication method is double-crucible. Double-crucible fiber fabrication uses two concentric crucibles filled with molten core glass (inner crucible) and clad glass (outer crucible). The core and clad are drawn through orifices at the bottom of the crucible to form fiber with a core : clad geometry. As described, this approach is of little value for ZBLAN, because the viscosities at which stable fiber draw are typically obtained (200–800 Pa · s) correspond to temperatures well below

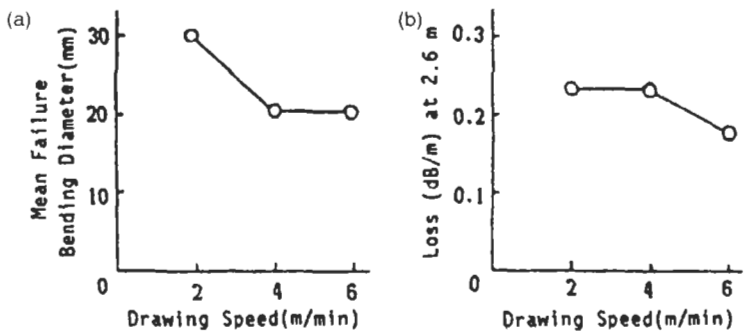


Fig. 5 Effect of draw speed of ZBLAN fiber on (a) mean bending diameter at failure, and (b) loss at 2.6 μm. From the study of Hirai et al., reprinted with permission.[60]

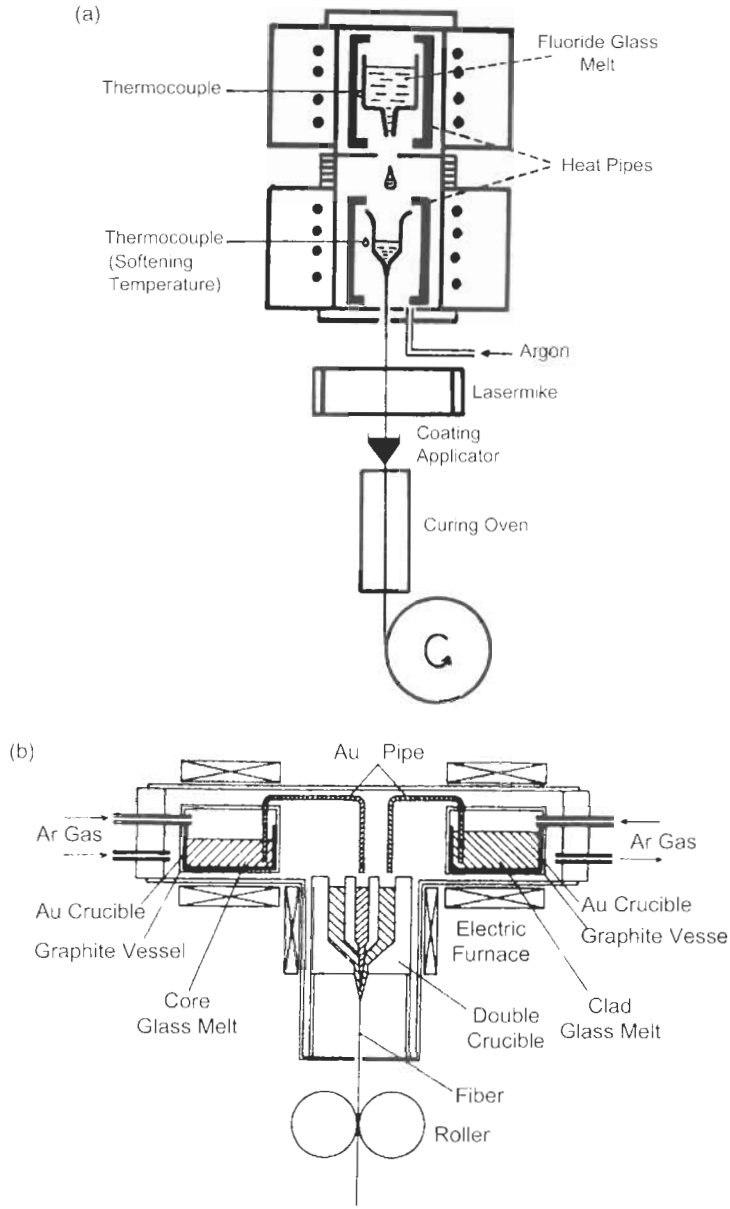


Fig. 6 Various methods used to draw ZBLAN and other soft glasses into fiber: (a) double-stage crucible fiber draw apparatus of Tran et al.;[61] reprinted with permission. (b) injection double-crucible fiberization of Tokiwa et al.;[63] reprinted with permission. (c) build-in casting of preforms from Mitachi et al.;[66] reprinted with permission. (d) rotational casting of preforms, from Tran et al.;[67] reprinted with permission. (e) combined suction casting with redraw in tapered preforms, from Kanamori et al.[70] reprinted with permission.

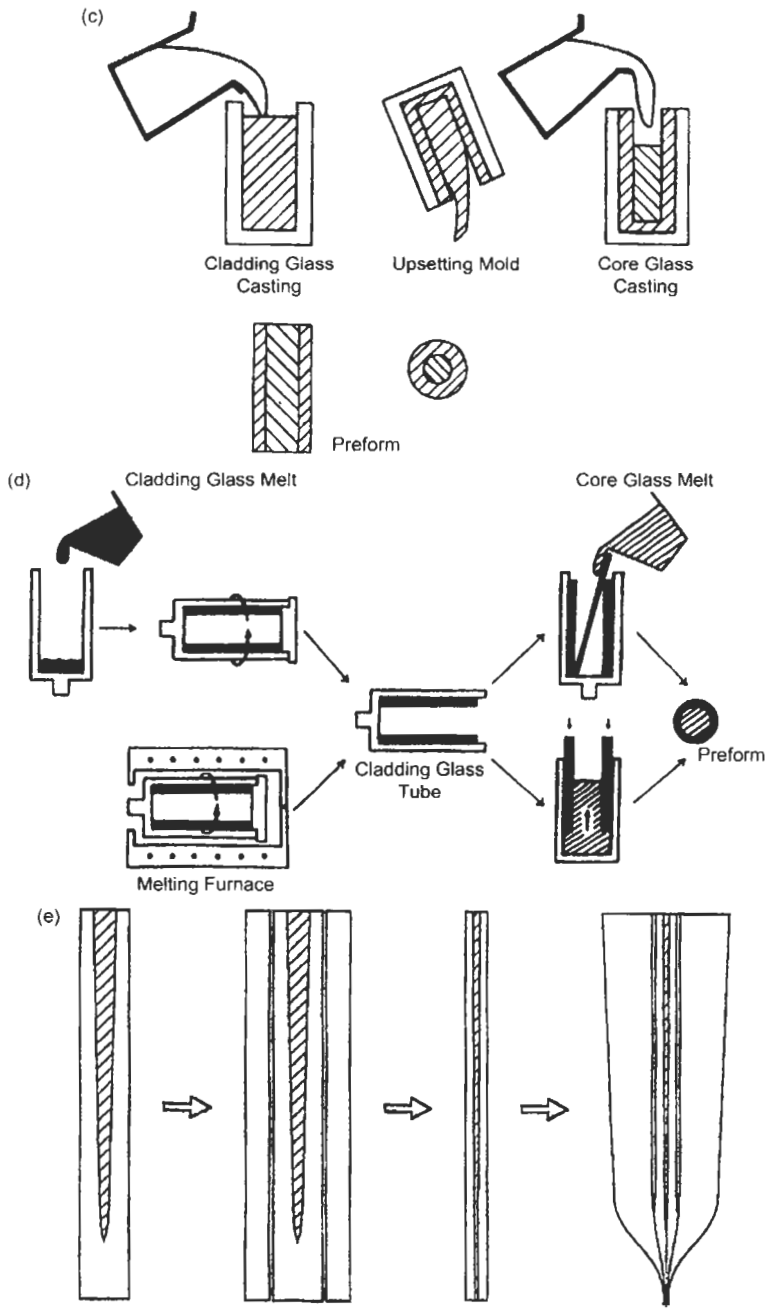


Fig. 6 (*Continued.*)

the liquidus temperature for ZBLAN. The double stage crucible approach[61] was developed to overcome this difficulty, and is illustrated in Fig. 6a. The upper crucible delivers melted glass to a forming crucible at lower temperature, and a fiber is pulled from the base of the forming crucible. Since relatively little glass is kept in the forming crucible at any time, the process can be run so as to minimize time spent below the liquidus and hence minimize devitrification during fiber draw. The thermal sensitivity of ZBLAN makes it critical to control all interfaces between the glass and crucible materials. Furthermore, air bubbles can be pulled into the fiber before having a chance to rise out of the forming crucible. These contribute to high scattering losses.

Mimura et al.[62] proposed an alternative crucible draw scheme in which the glass is injected from a crucible into a delivery tube. The exit end of the delivery tube is mounted to a cooled aluminum plate to impart some viscosity to the glass. The authors report obtaining 200 m lengths of 130 μm diameter core fiber (no core/clad structure) that was free of crystals, but no loss figures were reported. An extension of this method reported by Tokiwa et al.[63] permits simultaneous injection of both core and clad (Fig. 6b) to form core/clad structures. The authors report fabricating 600 m of crystal-free single mode fiber by this approach. Losses were not reported; however, photomicrographs of the fibers show the presence of bubbles in the clad and/or core, suggesting high scattering loss.

The first attempts to make fiber from fluorozirconates employed preform draw methods, and they are still the most widely used methods to this day. One of the first reported efforts to draw fiber from a fluorozirconate preform was performed by Mitachi and Manabe,[64] followed in short order by a preform draw fiber consisting of a fluorozirconate core and a fluorocarbon plastic clad.[65] The next problem was to form a core-clad structure in which the clad was also a fluorozirconate. Mitachi et al.[66] introduced the build-in casting method (Fig. 6c), which involves pouring a clad glass into a mold, waiting until the sides vitrify, discarding the still-molten interior glass and replacing it with core glass. This produces a core-clad geometry in an elongated form that is suitable for preform draw fiber fabrication. It is difficult, however, to obtain the large clad-to-core ratios typically needed for amplifier and laser applications, generally 25 : 1 or greater. Furthermore, it is difficult to control the geometry of the final preform. In modified build-in casting,[12] the core is poured on top of the clad and a hole is opened at the bottom of the mold. The clad at the center of the mold drains out and the core follows, producing a core/clad structure and an excellent glass/glass interface. This method has been used to produce relatively low loss fiber (~30 dB/km at 1.5 μm). Yet another variation on this approach is suction casting.[12] As in modified build-in casting, core is poured on top of the clad, but the clad contracts into a reservoir as it cools, drawing the core into the cooling well that forms. Today, this method has been used to make low-loss single-mode fiber from fluoride glasses, and is one of the most widely used methods.

The rotational casting method developed by Tran et al.[67] (Fig. 6d) has been used extensively to produce preforms with reasonable geometry and acceptable core/clad ratios for characterization work. In this method the clad is cast into a mold which is then rotated at great speed (~3000 rpm) to centrifugally cast the clad into a hollow cylinder. The core glass is then poured into the clad tube and carefully annealed into a final preform. This method produces very good geometry, but like build-in casting it lends itself to making fibers with relatively small clad : core ratios. A variation on this theme, reduced pressure casting,[68] forms the core and clad under reduced pressure, thereby reducing or eliminating air bubbles from the preform. This method was used to produce very low loss core/clad ZBLAN fiber (0.65 dB/km minimum, ~5 dB/km at 1.5 μm).

In each of the previous approaches, a clad is delivered first, followed by a core. It is also possible however, to form a rod of core glass and wrap clad around it. The most basic approach is to place a rod of core glass into a tube of clad glass, a method aptly referred to as rod-in-tube. This tends to produce many scattering centers at the core/clad interface, and so is not a preferred method for obtaining core/clad fibers. Alternatively, one can make a preform by one of the methods described above and slip it into a tube of glass to increase the clad : core ratio. A method for doing this called overclad drilling is described by Braglia et al.[69] While this preserves a good core/clad interface (assuming it was good in the original preform), the defects along the clad/overclad interface can still produce weak fiber. However, this method permits fabrication of single-mode fibers, and frees the experimenter to use any compatible glass as the overclad. An extension of this is to redraw the overclad preform into a single rod with three layers of glass. Kanamori et al.[70] combined suction casting with redraw in specially tapered preforms (Fig. 6e) to obtain 500 m of low-loss, single-mode ZBLAN fiber suitable for Er-doped fiber amplifier applications.

Finally, the clad glass can also be formed around a solid core. Mitachi et al.[71] describe a method wherein the core glass is formed into a rod, centered in a cylindrical mold, and molten clad glass is delivered around it. The large changes in volume associated with the cooling clad and the reheating core produced a poor core–clad interface, and the authors report that the fibers showed high loss. An extension of this method was developed by Kortan et al.,[72] in which a narrow rod of core glass was rapidly plunged into a cylinder of cooling clad glass. They report fiber losses of several dB per meter but attribute this largely to impurities and draw-induced defects.

Fiber Losses

It was recognized early on that formation of oxide crystals during glass synthesis or during fiber fabrication produced high scattering losses and weak fiber.[73] The second overtone of the hydroxyl (OH^-) stretching mode is located within

the telecommunications wavelength range at approximately 1450 nm and has a very large absorption cross section.[74,75] Thus, most of the early work on ZBLAN fiber resulted in methods that eliminated oxygen and water, but did little to eliminate other contaminants, notably transition metals and tramp rare earth elements. The reactive atmosphere processing methods described above have produced nearly complete elimination of water and oxygen from ZBLAN glasses, and improved handling methods have kept these levels low all the way into the fiber.

As an aside, it should be noted that complete elimination of oxygen is not always a good thing. Oxygen actually serves to stabilize many fluoride glasses, including ZBLAN in certain formulations. As an extreme example, fluoroaluminate glasses have been shown to be dramatically stabilized by the addition of small amounts of P_2O_5.[76] The presence of oxygen certainly limits the minimum loss achieved at long wavelengths (e.g., 2 μm or greater), but is of no practical importance in the telecommunications wavelength range. It is the suspicion of one of us (Ellison) that commercial manufacturers of ZBLAN for active optical devices deliberately add small amounts of oxides to improve glass stability.

As noted above, a second major source of attenuation is contamination by tramp transition metals, particularly 3d metals. Transition metal absorption cross-sections in a fluorozirconate glass were determined by Ohishi et al.,[77] and their figure illustrating absorption losses per ppm is reproduced in Fig. 7.

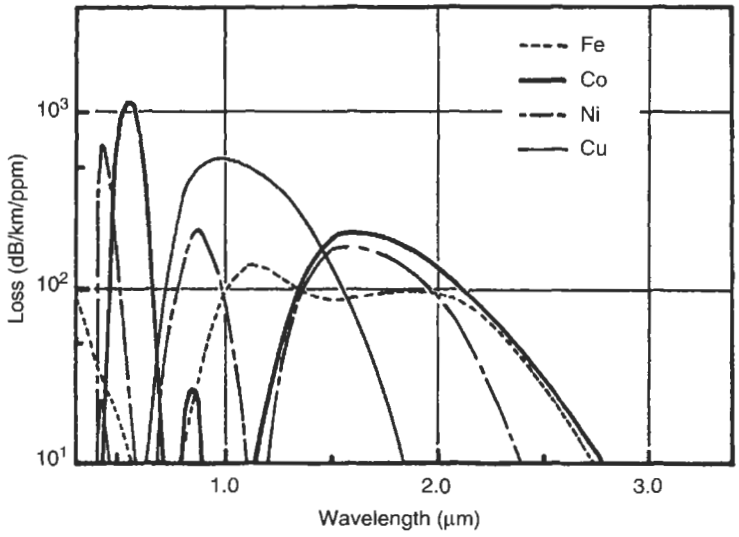


Fig. 7 Absorption losses in ZBLAN induced by transition elements, in dB/km/ppm. From the study of Ohishi et al., reprinted with permission.[77]

Iron, cobalt, nickel, and copper all have absorptions in the telecommunications wavelength range with magnitudes on the order of 100–200 dB/km/ppm. Aasland et al.[78] showed that the oxidation state of transition metals, and hence the nature and wavelength dependence of their IR absorptions, are in some cases sensitively dependent upon whether the glass was processed under nitrogen (N_2) or trifluoroammonia, NF_3—that is, whether the atmosphere was inert or oxidizing (reactive). After synthesis in an inert atmosphere, most transition metals studied were present in two valences, whereas after synthesis in a reactive atmosphere only one valence was measured for each metal studied. Jewell et al.[79] showed that fiber losses in the near IR were likely to be entirely dominated by transition metal absorptions, whereas at shorter wavelengths intrinsic scattering losses were dominant. The methods for preprocessing fluoride batch materials described above have nearly eliminated transition metal contamination in the best ZBLAN glasses and fibers.

A third source of loss in lanthanum-rich ZBLAN is absorptions from tramp transition metals. Ohishi et al.[80] measured absorption spectra for all optically active rare earth elements in a simple fluorozirconate glass. Table 1 shows the absorption in dB/km/ppm for various rare earth elements with absorptions in the telecommunications wavelength range. The natural abundance of rare earth elements drops dramatically with increasing atomic number, so heavy rare earth elements such as Er and Tm are unlikely to be present in high concentrations in lanthanum sources. On the other hand Pr is similar in size to La and can be a significant contaminant in lanthanum sources. The separation chemistry associated with the lanthanide elements has a long and illustrious history, far beyond the scope of the present review. Suffice it to say that methods and technologies exist to produce lanthanum oxide at very high purity, albeit for a price.

A fourth source of absorption loss is vibrational modes of complex anions in the glasses. Poulain and Saad[81] deliberately doped a fluorozirconate glass

Table 1 **Absorption Cross Sections for Rare Earth Elements with Optical Transitions in the Telecommunications Wavelength Range**

Ion	*Peak Wavelength (μm)*	*Absorption Cross Section (dB/km/ppm)*	*Optical Transition*
Pr^{3+}	1.56	91.3	$^3H_4 \rightarrow {}^3F_3$
	1.46	54.7	$^3H_4 \rightarrow {}^3F_4$
Sm^{3+}	1.51	28.2	$^6H_{5/2} \rightarrow {}^6H_{15/2}$
	1.39	29.9	$^6H_{5/2} \rightarrow {}^6F_{5/2}$
Er^{3+}	1.54	38.6	$^4I_{15/2} \rightarrow {}^4I_{13/2}$
Pr^{3+}	1.68	26.6	$^3H_6 \rightarrow {}^3F_4$

with small amounts of phosphate, sulfate, carbonate, nitrate, and titanate anions. Data for sulfate and phosphate anions were discussed in some detail. The authors concluded that sulfates and phosphates had a drastic impact on optical transparency in the near IR, and would have to be kept well below the 1 ppm level were transmission-quality fiber ever to be obtained. As noted above, the loss levels acceptable for active fiber devices are typically much higher than could be tolerated in telecommunications applications. Furthermore, the reactive atmosphere melting methods described above go a long way toward completely eliminating oxide contaminants in ZBLAN, whatever their form. Finally, some level of oxide may actually be useful if it promotes glass stability.

France et al.[82] evaluated the impact of residual ammonium in ZBLAN on fiber losses. Recall from above that a classic synthesis route for preparing high-purity fluorides was to react precursor materials with ammonium bifluoride, then decompose the ammonium compounds at elevated temperatures. It is very difficult to establish that all ammonium has been eliminated at a ppm level. The authors show data for a fiber in which the core glass was prepared via bifluoride consumption of ZrO_2 in which the fundamental N–H stretching mode is clearly discerned at approximately 2.9 μm. An overtone of this mode would appear near 1.45 μm, and would certainly be in a position to impact losses in the telecommunications wavelength range. Fortunately, careful handling of precursor materials generally results in nearly complete elimination of ammonium ions from ZBLAN.

Absorbers such as transition elements are referred to as intrinsic sources of loss because they are present in the glass before it is ever made into fiber. Extrinsic sources of loss are scattering centers produced in preform fabrication or fiber draw. Nearly all the research and development efforts in fluoride fibers are directed at reducing and controlling extrinsic losses, and there are many research reports on this subject. Good models of defect formation that take advantage of data concerning crystal growth kinetics have been reported by Hopgood and Rosman[83] and Sakaguchi.[84] The Sakaguchi model is particularly complete because it also includes the thermal history of the glass during formation of the preform. The variables included whether the preform was a rod of a single glass or a core/clad structure produced by rotational casting or by suction casting, the diameter of the preform, the temperature profile in the furnace, and the speed with which the fiber is drawn. The predicted losses for suction casting were far lower than those for rotational casting, because the glass spends less time at high temperature during preform fabrication. For preforms produced by equivalent methods, fibers produced from thick preforms were predicted to have much higher scattering losses than fibers produced from thin preforms due to longer exposure to high temperatures during fiber draw. For the same reason, scattering losses were predicted to be lower for fibers drawn at high speed, as a given preform spends less time in the hot zone of the draw furnace. Unfortunately, small preforms do not lend themselves to

producing large quantities of fiber; on the other hand, the demands for fiber in photonic devices tend to be small.

A final source of extrinsic loss is air bubbles in the core or clad glass or trapped between the core and clad. Reduced pressure casting tends to reduce or eliminate such defects, as do the various forms of build-in casting. A method for eliminating gas bubbles in ZBLAN-based rod-in-tube preforms proposed by Harbison and Aggarwal[85] is to employ hot isostatic pressing. The authors show photographs illustrating the substantial reduction in bubble density produced by this approach. The main benefit, aside from a reduction in scattering losses from rod-in-tube fibers, is the possibility of obtaining greater lengths of fiber without breakage.

Fiber Strength and Reliability

One of the greatest concerns raised by use of fluorozirconates in telecommunications applications is their poor durability and the poor mechanical strength of the fibers obtained from them. The strength of a fiber is an important factor in photonic applications, because fibers are typically wound around small reels contained within the device, producing bend stress, and the connections to the other fibers and/or components generally produce a small amount of installed (tensile) stress. The handling of the fiber during testing and as reels are assembled can cause crack propagation from small surface defects and cause failure in the device long after it has left the plant. Anyone with experience in working with fiber reliability experts will appreciate that there is virtually no way that fiber can be strong enough, much less too strong, for their applications.

Fiber strength can be broken into two separate parts, the intrinsic strength of the material subjected to stress, and the actual strength of a fiber as obtained from any particular process. The intrinsic strength arises at its most fundamental level from the relative strengths of bonds in the glass and the resistance of the glass to deformation under an applied load. Since the state of a glass is affected by the way in which it is processed, the intrinsic limit tends not to be attainable, and the means by which it can be quantified are not known. The actual strength of a fiber is determined largely by stress imposed during the draw—particularly if the core and clad glasses have different glass transition temperatures or different coefficients of thermal expansion—but also by defects in the fiber or on its surface. The actual strength is always less than the intrinsic strength.

It has been recognized for many years that the single most important factor limiting the strength of fluorozirconate fiber is crystalline defects formed during preform fabrication and/or during the fiber fabrication process.[86] The worst location for crystalline defects is on the surface of the fiber, as crack propagation is aided and abetted by contact with air. Numerous approaches have been developed to reduce surface crystallization of fluorozirconate fibers.

Pureza et al.[87] polished preforms to a mirror-like finish and etched the polished preform in an acidic solution of zirconium oxychloride ($ZrOCl_2$). They showed that this raised the failure strength of ZBLA fiber to approximately 570 MPa, with peak values of approximately 820 MPa. The authors also found that polishing by itself produced much weaker fiber than a combination of polishing and acid leaching. This comports with the observations of Delben et al.,[88] who showed that polished ZBLAN exhibits a surface layer with lower mechanical hardness than bulk ZBLAN. Acid leaching presumably strips this layer off, leaving clean glass. Orcel and Biswas[89] found that drawing fibers at comparatively low temperature in an atmosphere of dry helium thionyl chloride ($SOCl_2$) improved strength by suppressing surface devitrification. The proposed mechanism is suppression of water in the furnace by the thionyl chloride. The highest strengths reported in these reports are similar to the highest strengths reported today, 500–600 MPa. For reference, silica-based amplifier fibers tend have tensile strengths on the order of 7–11 GPa.

The comparatively low durability of fluorozirconate glasses raises fiber reliability issues. Most amplifier manufacturers expect their passive components to survive environmental testing laid out in the Bellcore Technical Advisory TA-NWT-001221.[90] The most demanding of these tests include monitoring losses in fibers during temperature cycling (−40°C to +75°C, 90% relative humidity, 500 cycles), damp heat exposure (75°C and 90% relative humidity for 5000 hours), and after low-temperature storage (−40°C for 5000 hours). Though pumped optical amplifier fibers are not passive components, they too are generally expected to pass the Bellcore specification. It has been known for some time that fluorozirconate fibers cannot survive high-temperature, damp heat conditions for such long periods of time. Recently, Fujiura et al.[91] and Nishida et al.[92] at NTT have described a hermetic package for fluorozirconate fiber. In substance, the fiber is epoxied to a bobbin etched with a V-groove, pigtailed to silica fiber, sealed in a package under an inert atmosphere, and packed again into an aluminum case (including the pigtails) that is back-filled with epoxy resin. The authors show that this package survives the Bellcore tests with no obvious failures in multiple tests. This provides hope that there will be a commercially viable means to deploy fluorozirconate fibers in active devices.

Applications

Fluorozirconate fibers have proved to be efficient hosts for several rare earth elements-doped fiber laser/fiber amplifier applications. Brierley and France[93] were the first to report an active device based on fluorozirconate fiber, in this case a neodymium fiber laser operating at 1050 nm. This was followed by many new developments in very rapid succession: a Nd-doped fiber laser[94] and a fiber amplifier[95] operating at 1.3 μm; a holmium-doped fiber laser operating at 1.38 and 2.08 μm;[96] erbium fiber lasers operating at 1.5 μm[97] and 2.7 μm;[98]

a ytterbium-sensitized erbium fiber operating at 1.5 μm;[99] a Yb fiber laser operating from 1.015 to 1.14 μm;[100] thulium-doped fiber lasers operating at 1.9 μm,[101] 2.3 μm,[102] and 0.8 μm;[103] and praseodymium-doped fiber lasers operating at 1.3 μm.[104,105]

Of these many applications, those involving amplification at 1.3 μm, 1.48 μm, and 1.55 μm have proved of greatest interest for telecommunications applications. Initially, neodymium-doped fiber amplifiers generated considerable interest because of the possibility of using the $^4F_{3/2} \rightarrow {}^4I_{13/2}$ transition at 1.35 μm to amplify signals near the dispersion minimum for germania-doped silica transmission fiber.[106,107] However, the $^4I_{3/2} \rightarrow {}^4G_{7/2}$ excited state absorption (ESA) from the $^4F_{3/2} \rightarrow {}^4I_{13/2}$ transition saps gain and adds to noise figure in the wavelength range closest to 1.3, particularly in preamplifiers.[108] Alternative pumping schemes[109] and compositional adjustments to the fluorozirconate host[110] have shown some promise, but interest largely faded after demonstration of Pr^{3+}-doped fiber amplifiers. By the early nineties, systems tests were being performed on Pr-doped fluorozirconate fiber amplifiers,[111] and methods to optimize fluorozirconate fibers as Pr^{3+} hosts[112] and amplifier pumping schemes[113] were underway (a summary of lab results and a field test conducted by BT Labs is presented by Whitley).[114]

Amplification at 1.48 μm is achieved using the $^3F_4 \rightarrow {}^3H_4$ transition of trivalent thulium, Tm^{3+}. In this application, fluorozirconates have substantial advantages over silica-based amplifier materials in that the highest-lying phonon modes are below 700 cm^{-1}, leading to very long lifetimes for the $^3F_4 \rightarrow {}^3H_4$ transition.[115] Komukai et al.[116] were probably the first to demonstrate upconversion pumping (1.064 μm from a Nd^{3+} : YAG laser) of thulium ions in a fluorozirconate fiber to obtain 1.48 μm output. Despite investigations of alternative pumping schemes,[117] variations on this approach have proved to be the method of choice for producing high-performance 1.48 μm amplifiers.[118] Recent developments in Tm-doped fiber amplifiers are discussed in detail below.

While amplification at wavelengths other than 1.55 μm is potentially important for future applications, most of the world has already adopted erbium-doped fiber amplifiers (EDFAs) for mainstream optical telecommunications. Ronarc'h et al.[119] reported one of the first efforts to make a practical EDFA (1480 nm pump) from a fluorozirconate fiber, and the ASE (amplified spontaneous emission) spectrum showed very little gain ripple over a broad wavelength region. Clesca et al.[120] compared gain flatness of Er-doped fluorozirconate fiber with Al-doped silica and showed that at equivalent bandwidth (~23 nm), a fluorozirconate-based fiber amplifier showed less gain ripple than a silica-based amplifier (fluorescence spectra of erbium in ZBLAN and an optimized aluminum-doped silicate are compared in Fig. 8). If nothing else, this study elevated to high importance the effort to produce new amplifier materials with lower gain ripple. The main problem with fluorozirconate

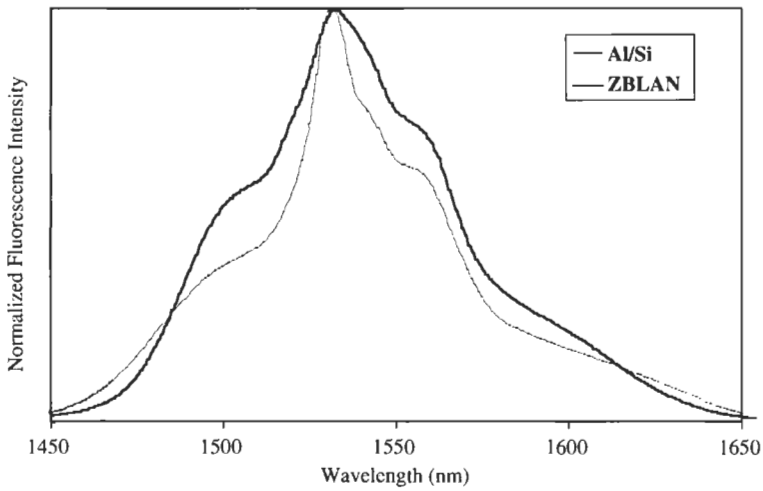


Fig. 8 Normalized 1550 nm Er^{3+} emission spectra of ZBLAN and Al-doped silica glasses.

fibers used in this application was that 980 nm pump power conversion efficiency is very low in fluorozirconates compared to Al-doped silica due to the long lifetime of the $^4I_{15/2} \rightarrow {}^4I_{11/2}$ transition, approximately 1.5 ms in ZBLAN vs. 3–5 μs in silica-based materials. This is important because it is generally desirable to run the preamplifier at high inversion and the power amplifier at lower inversion to obtain optimal noise figure. Inversions above 60% are not possible using 1480 nm pumps. Various schemes have been proposed to diminish the lifetime of the 980 level, notably introducing Ce^{3+} as a codopant.[121]

Benefits of Research in Fluorozirconates

As discussed later in this review, rare earth element-doped fluorozirconate fibers have generated considerable interest for fiber amplifier and laser applications even as hope fades for their use in telecommunications fiber. To some extent, the photonic applications grew out of the effort to deliver telecommunications-quality fiber, rather than the other way around, so new applications continue to appear. Perhaps the greatest legacy of the efforts to draw good fiber from fluorozirconates is that numerous methods and processes have been invented that can be applied to many other glasses with very poor thermal stability, poor durability, or that require complicated synthesis methods. Other fluoride glasses are the most obvious beneficiaries, but methods developed for fluorozirconates have also been applied to chalcogenides and most recently to tellurites, some of the most recent and interesting oxide hosts on the photonics scene.

Fluoroaluminates

Given the many years of research devoted to fluorozirconate glasses, researchers are often surprised to learn that the first multicomponent fluoride glasses identified were, in fact, fluoroaluminates, and that they were discovered many years ago. In 1947 Sun[122,123] reported glass formation in a beryllium-free multicomponent fluoride system. These glasses contained varying amounts of alkaline earth, lead, and/or lanthanum fluorides, but aluminum fluoride appeared to be an essential constituent. Though aluminum fluoride generally did not exceed 34 mol% in these glasses, they have been referred to as fluoroaluminates since shortly after their discovery.

For those familiar with oxide glasses, or even fluorozirconates, fluoroaluminates seem highly unstable and, hence, undesirable for commercial applications. There is more here than meets the eye, however, in this broad family of glasses. The most stable fluoroaluminate glasses are obtained at relatively high concentrations of the fluorides of alkaline earth cations, particularly calcium and magnesium fluorides. The very low water solubility of these fluorides contributes to greater durability than ZBLAN. Furthermore, the refractive indices of fluoroaluminates typically fall in the range 1.4–1.5 (sodium-D), similar to silica but lower than ZBLAN, and their Abbe numbers are generally comparable to or greater than that of silica. Therefore, fluoroaluminate glasses are attractive for applications demanding low optical dispersion, low refractive index, or combinations thereof, and remain an active area of research.

Recently, Kitamura et al.[124] have shown that the optical dispersion of fluoroaluminate glasses is approximately 60% that of v-SiO_2, and demonstrated as well the possibility of good UV transmission for these glasses. Raman spectroscopy of simple fluoroaluminates[125] shows that the IR band edge should be similar to fluorozirconates, which, combined with the good UV transmission, implies the possibility of making very low-loss fibers.

Composition

Compositionally, there are many variations on a general theme. The basic fluoroaluminate glass contains approximately 30–40 mol% AlF_3 and 55–70 mol% alkaline earth fluorides, often with 5–15% yttrium fluoride (YF_3) added to enhance glass stability. Cardoso et al.[126] identified a particularly stable composition, $37AlF_3$–$12MgF_2$–$15CaF_2$–$9SrF_2$–$12BaF_2$–$15YF_3$, and examined its devitrification behavior as a guide for fiber fabrication.

Glass Synthesis and Fiber Fabrication

Fabrication of fluoroaluminate glasses is similar to that of fluorozirconates. Like fluorozirconates, the glasses must be melted in an oxygen- and water-free atmosphere at temperatures between 750 and 1000°C. Unlike fluorozirconates,

reactive atmosphere processing to avoid reduction of glass constituents is generally not necessary, though it does remove water and oxygen and may enhance the stability of the glass. Glasses are formed by quenching the melt into a heated mold. Because the thermal stability of normal fluoroaluminates is considerably less than that of ZBLAN, care must be taken to cool the melt as quickly as possible to avoid devitrification. Once made, the glasses are relatively stable against attack by water or oxygen and can be handled, shaped, and polished in air—indeed, in contrast to ZBLAN, fluoroaluminate glasses can be polished using water as a flux.

Because fluoroaluminates have much lower thermal stability than ZBLAN, they have a much greater tendency to devitrify during fiber draw. Consequently, there has been much less effort on drawing fluoroaluminates to fiber than for more stable fluorides, notably ZBLAN. Iqbal et al.[127–130] made core/clad fibers from ZrF_4-stabilized fluoroaluminates and obtained relatively low losses (~0.2 dB/m) within the telecommunications wavelength band. There is evidence that fluoroaluminate fiber may have greater mechanical durability than fluorozirconate fiber, though this interpretation is somewhat dependent on the method used to perform the test.[131]

Applications

Naftaly et al.[132] examined Nd^{3+} spectroscopy in an alkali fluoride-stabilized fluoroaluminate glass and fibers. The core glass, $40AlF_3$–$6MgF_2$–$22CaF_2$–$6SrF_2$–$6BaF_2$–10LiF–10NaF includes alkali fluorides for improved glass stability and glass ionicity. The authors used a phosphate-stabilized fluoroaluminate clad glass and drew fiber from which gain was obtained at 1.3 microns. Losses were not stated. The authors concluded that efficient Nd^{3+} 1.3 μm amplification might be possible using fibers made from these or closely related materials.

Fluoroberyllates

Much as silica is the quintessential oxide glass, beryllium fluoride, BeF_2, is the quintessential fluoride glass. It has been recognized since before 1932 that beryllium fluoride forms a glass.[133] Zachariasen conjectured that BeF_2 would be the only single-component fluoride glass to be found, which remains the case to this day. Beryllium fluoride is readily vitrified from its melt, though handling in an inert atmosphere is generally required to avoid forming BeO. In its purest forms, the UV absorption edge of BeF_2 is below 200 nm, and its IR absorption edge lies at longer wavelengths than SiO_2. It also has the lowest optical dispersion of any known glass. Furthermore, numerous fluorides can be added to BeF_2 to improve its durability without severely compromising its desirable optical attributes (see the review by Gan.[134]) One might expect enthusiastic efforts to draw it to fiber, and indeed some effort has been made. However, beryllium compounds tend to be acutely toxic, so the argument to

develop BeF_2-based photonic fibers has seemed less than compelling. Nevertheless, interest in fluoroberyllates continues for their potential applications as deep UV transmission materials,[135] so photonic applications may follow as well.

OXIDES

The glasses of everyday experience—bottles and jars, windows, cookware—are all oxide glasses; indeed, most are a very specific variety of oxide glass called soda-lime silicate ("soda" and "lime" are common names for sources of sodium and calcium). Soda-lime silicate glasses are durable, are composed of very cheap ingredients, but have rather pedestrian attributes as far as rare earth element spectroscopy is concerned. The most important fibers for active devices in telecommunications are alumina-doped silica and silica co-doped with alumina (Al_2O_3) and germania (GeO_2), referred to hereafter as Al/Si or Al/Ge/Si, respectively. More recently, tellurite fibers have been used to build EDFAs, and antimony silicate fibers have been used in EDFAs and TDFAs. Each of these will be considered in turn below.

Al/Ge/Si and Related Materials

When we think of optical amplifiers, we generally have in mind silica doped with aluminum, germanium, and erbium. It is interesting in this regard that no other ion performs particularly well in any sort of silica: for example, the highest lying phonon modes of silica-based materials drastically impacts the efficiency of Tm-doped silica fibers operating at 1.4 μm; Nd-doped silica fibers are very efficient lasers at 1.056 μm but very inefficient at 1.33 μm; other rare earth elements that are potentially active in the infrared show no activity at all in silica-based materials because of high rates of phonon relaxation, Pr^{3+} being one of the most obvious. On the other hand, Er^{3+} in alumina-doped silica is one of the wonders of nature, an extremely low loss, highly efficient host with adequate bandwidth to handle the enormous optical network traffic of the present day. It is completely appropriate that other materials be held to measure against this exquisite material, but a grim task lies for materials scientists to improve on its properties.

Namikawa et al.[136,137] were the first to report that rare earth ions in pure SiO_2 were prone to nonradiative crossrelaxation. This process, soon thereafter referred to as "clustering" (see above) was observed for neodymium ions but was shown in later work to apply to erbium as well. In one of the great, perhaps underappreciated breakthroughs for modern telecommunications, Arai et al.[138] showed that clustering of neodymium ions in silica was drastically reduced by additions of small amounts of aluminum or phosphorus to pure silica. The amounts of aluminum or phosphorus required were on the order of 10–15 times greater than the neodymium ion concentration (e.g., mol% or ions/cc). It was later shown by Morkel et al.[139] that additions of alumina to

silica in the cores of neodymium-doped optical fibers dramatically reduced excited state absorption at the $^4F_{3/2} \rightarrow {}^4I_{13/2}$ transition near 1370 nm. Both studies concluded that aluminum and neodymium were spatially associated in the glass, and that this served to redistribute neodymium ions relative to aluminum-free silica.

More recent studies have demonstrated close association of aluminum and rare earth elements in alumina-doped silica,[140–142] including in glasses prepared by sol-gel (as opposed to chemical vapor deposition) methods. One of the first and perhaps most interesting applications of this new-found understanding of the role of aluminum in Ge-doped silica was to make "loss-less" optical fiber. In this approach, one would prepare telecommunications spans' worth of Al/Ge/Si fiber doped with very low concentrations of erbium, pump the erbium, and thereby overcome the losses inherent in the fiber. A fiber suitable for this application was first reported by Craig-Ryan et al.[143] At roughly the same time, however, researchers were attempting to build efficient erbium fiber lasers and amplifiers operating at 1.55 μm, and were soon very successful. The rest, as the saying goes, is history.

Composition and Rare Earth Element Spectroscopy

It is truly a miracle that such a small amount of a codopant should have such a dramatic impact on the optical characteristics of rare earth elements in telecommunications-type fiber. Indeed, the role of aluminum is so overwhelming that silica can be regarded as simply a structural component and germania as a means to control refractive index. Aluminum doping also dramatically impacts erbium spectroscopy, as discussed in detail by Desurvire.[144] In pure silica or Ge-doped silica, the 1.5 μm emission and absorption bands are very narrow and consist of two prominent bands, or peaks. Addition of alumina to silica or Ge-doped silica causes broadening of the emission and absorption spectra. In the emission spectrum, this produces a broad shelf from approximately 1535 to 1555 or 1560 nm with ~50% of the intensity of the primary emission band near 1530 nm. As aluminum concentration increases, the 1530 nm band also shifts to shorter wavelengths, expanding the shelf of emission intensity. Through this wavelength range, the absorption spectrum is diminishing smoothly and rapidly, and thus at approximately 65% inversion one obtains a broad, flat gain spectrum from about 1530 to 1560 nm.

There are three basic types of erbium-doped Al/Ge/Si fiber; broadly categorized as follows:

Designation	**Al_2O_3 (*wt%*)**
Type I	0.7–1.2
Type II	2.0–4.5
Type III	>4.5

Mole for mole, Al_2O_3 raises the refractive index of silica about as much as GeO_2, so increases in Al_2O_3 content are typically offset by decreases in GeO_2 content. In most cases, the concentrations of Al_2O_3 and GeO_2 are adjusted so as to obtain a numerical aperture of 0.2–0.24 relative to a silica in a single-mode, step-index fiber.

The width of the 1550 nm gain band and the magnitude of the structure within the band are strong functions of aluminum content, but only up to a point. The payback for continuing to increase impact of alumina diminishes substantially when the content exceeds approximately 5–6 wt% Al_2O_3. Ultimately, it is the shape of the gain curve that matters to the applications engineer. A series of gain curves for three silica-based fibers, pumped at 980 nm, are shown in Fig. 9. In this case, the spectra were normalized by varying the length of the fiber to yield approximately the same gain at 1528 and 1560 nm. It is clear from this figure that the gain curves differ more when alumina varies from 2.1 to 4.7 wt% than when the concentration is changed from 4.7 to 8.3 wt% Al_2O_3.

The stability of the fiber core host-glass decreases with increasing alumina content. Risbud et al.[145] and later Schumucker et al.[146] showed that at intermediate Al_2O_3 concentrations, aluminosilicate glasses undergo amorphous phase separation, even when cooled at rates comparable to those obtained under high-speed fiber draw conditions. At higher concentrations, mullite crystals (approximately $2Al_2O_3 \cdot SiO_2$) were obtained under the same rapid-quench conditions. Mullite formation during consolidation is difficult to avoid in glasses derived from chemical vapor deposition processes (see below); even if avoided then it may occur anyway during long residence times outside of

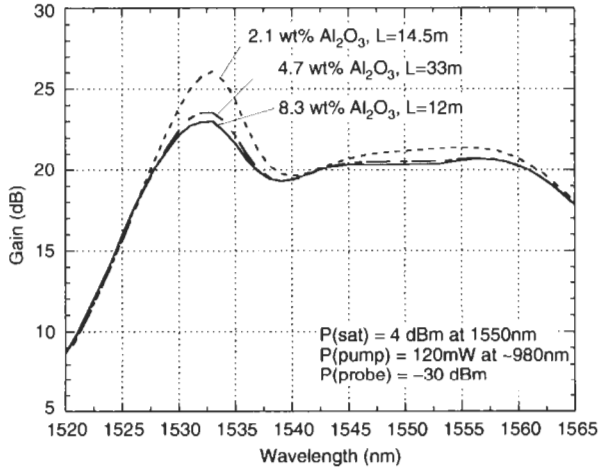


Fig. 9 1550 nm erbium gain curves for Al-doped silica fibers as a function of alumina content.

the hot zone in a fiber draw furnace. As a result, one must balance the need for broad and flat gain with the need for low passive loss, gain per unit length of the fiber, and ease of processing. Most commercial Al/Ge/Si fibers contain 3–6 wt% alumina.

As noted above, phosphorus can also "decluster" rare earth elements and in some cases may beneficially impact the breadth of gain attainable from the fiber. Betts et al.[147] evaluated Ge-doped silica-based fibers codoped with phosphorus and aluminum. The P and Al concentrations were not given, but from the refractive index data provided it appears that P_2O_5 was loaded at a level of 3–4 wt%. The authors report that less structure was seen in the 1550 nm Er^{3+} emission spectrum in a P-doped fiber than in Al-doped fibers, though gain spectra reported elsewhere in the paper show considerable structure, more than typically seen in Al-doped silica fibers with approximately the same codoping levels. The P-doped fibers also showed gain to very long wavelengths, a result confirmed in more recent work by Kakui et al.[148]

A great deal of work has been performed using fibers simultaneously codoped with aluminum and phosphorus. The compound $AlPO_4$ undergoes all the polymorphic phase transformations of SiO_2 and when added to SiO_2 forms extremely stable, high-viscosity glasses. Aluminum and phosphorus together work each to stabilize the other while maintaining the benefits of declustering rare earth elements. Unfortunately, dopant levels are seldom provided, perhaps for proprietary reasons, though as germanium is generally present levels of 2–4 wt% each are probable.

Glass Synthesis and Fiber Fabrication

Virtually all Er-doped fiber in use in telecomunications systems—indeed, virtually all optical fiber manufactured in the world today—is produced from preforms fabricated by a variant of the chemical vapor deposition process, usually referred to by the acronym CVD. CVD was first discovered by Frank Hyde[149] at Corning Inc. when he sprayed silicon tetrachloride ($SiCl_4$) through a burner flame. The silicon tetrachloride reacted with oxygen under the intense heat of the burner, producing pure silica glass particles and hydrochloric acid (HCl) in addition to other combustion products. The combination of high heat, carbon monoxide, and chlorine levels actually strips certain contaminants from the flame, particularly transition metals. When the silica particles are kept at high temperature for extended periods of time, they smoothly densify into solid, defect-free glass. Corning's High Purity Fused Silica (HPFS^{TM}) is produced via this process to this day. Unfortunately, this method traps high levels of water in the silica (800 ppm or more), and so is unsuitable for producing low-loss optical fiber in the telecommunications window.

A great deal of work was performed in the late 1960s and early 1970s to identify a means to obtain dry, contamination-free glass with a higher refractive index than silica to serve as the core in optical fiber. The breakthrough came

in 1973 with the patent by Keck, Schultz, and Zimar[150] for what is now known as the outer vapor deposition (OVD) process. In this process, the silica particles ("soot") are sprayed onto a ceramic rod (called a "bait" rod). The rod is rotated constantly, and the soot builds up in successive layers. The composition of the soot can be varied dynamically to form a core soot, then a clad soot, or still more complicated structures. The cylindrical body that results, formed of many layers of soot, has a bulk density of 0.2–0.5 g/cc, as compared to 2.2 grams for pure SiO_2—in other words, it is mostly free space.

The soot blank, as it is called, is subjected to rapidly-flowing halide-bearing (e.g., Cl_2) gas at elevated temperature to strip out transition metal contaminants and water, then consolidated into dense glass at higher temperature, typically 1250–1500°C. Since chlorine gas could access all areas of the soot blank, removal of undesirable contaminants is highly efficient. Today, minimum losses in Ge-doped silica fibers produced by the OVD process are approximately 0.17 dB/km, very close to the nominal Rayleigh limit for Ge-doped silica of 0.15 dB/km.

One potential disadvantage of the OVD approach for erbium-doped fibers is that the rates of combustion of the various precursor materials are very unlikely to be equivalent. As a result, it is unlikely that erbium ions will be perfectly uniformly distributed throughout each of the soot particles produced in the combustion process. Rather, the erbium is quite likely to form small dense nuggets that differ little from erbium oxide itself. This limitation can be overcome in part by using low concentrations of erbium in the combustion process by changing the identities of the precursor materials or the flame conditions so as to cause more uniform combustion, or by adding the erbium in a solution with one or more of the other precursor materials.

Outer vapor deposition is not the only means to produce astoundingly low-loss silica-based fiber. MacChesney and DiGiovanni[151] provide an excellent and very thorough review of two other basic approaches, the inner vapor deposition (IVD) method and the inadequately named "modified chemical vapor deposition" (MCVD) method. In the IVD method, combustion of the silica precursor compound takes place within the interior of a tube, rather than over the surface of a bait rod. A burner moves over the outer surface of the tube, providing a hot spot that facilitates collection of soot along the length of the tube. After sufficient soot is deposited, the soot inside the tube is subjected to halide drying and is consolidated into dense glass. A vacuum is applied to the tube, the tube is heated to high temperature, and the central void (the center line) is collapsed. The fact that the inner surface of the center line never contacts anything other than air may provide some advantage for access of the halide drying/decontaminating gas to access all regions of the core. It also has the advantage that core can be laid down directly onto the surface of a clad tube, thereby minimizing total thermal processing of the soot

after laydown. This is particularly important for relatively unstable materials such as Al-doped silica.

Nevertheless, IVD as described suffers from the same sorts of limitations as OVD as regards direct laydown of erbium along with the rest of the fiber components. This is overcome to some extent, however, by the fact that IVD results in a dense glass tube filled with low-density soot, making it is possible to infiltrate the soot within the tube with a solution containing the dopant ion of interest. This was, in fact, the means by which efficient Er- and Al-codoped silica fibers were first obtained[152] (see the review by Ainslie[153] for more detail, particularly as concerns erbium concentration quenching). In this approach, a solution containing erbium and aluminum salts is infiltrated into the soot constituting the core of the optical fiber. The soot is dried and then subjected to halide gas treatment and consolidation. The solution presumably coats many of the soot particles with the aluminum and erbium salts, producing a very low-concentration layer rich in Al and Er over an enormous surface area, and is mingled with the rest of the glass during consolidation. This helps minimize erbium ion clustering. However, clearing all traces of OH^- out of the soot is complicated by solution doping processes, and it requires special attention to detail to avoid making this as great a problem as erbium ion clustering.

In the MCVD method, a silica precursor reacts directly with oxygen in a plasma within a tube. The reaction product is laid down directly as dense glass. Provided that the reactants are free of contaminants, one obtains high-purity silica-based glass without first producing a soot blank. This approach has the disadvantage that organic precursors must be avoided whenever possible, as it is not possible to extract them efficiently during laydown of the glass. Furthermore, it can be difficult to control the composition of the material along the length of the tube because different parts will have different thermal histories. It is also somewhat difficult to employ MCVD for relatively unstable compositions, because the newly formed glass remains at elevated temperature for most of the laydown process. Finally, since silicon tetrachloride ($SiCl_4$) is the preferred precursor for silica, chloride contamination is a nearly inevitable feature of materials produced by MCVD processes. This can contribute to photosensitivity and attenuation due to chlorine-related optical defects in the near UV.

On the other hand, anything that can be metered into the plasma can be laid down in glass. This includes the anions reacted with the metal precursors as well as the metals themselves. MCVD is probably the only efficient means to lay down alkali cations, which are otherwise stripped via chlorine drying in OVD and IVD processes. It is also one of the most effective means of adding fluorine to silica. While neither of these components induces valuable spectroscopic changes to erbium ions, there is little doubt that such changes could be obtained via other dopants. This leads us to conclude that MCVD is underutilized as a means to produce fiber for erbium-doped fiber amplifiers—more effort should be made to explore the compositional richness available through this method.

While all of the world manufactures erbium-doped Al/Ge/Si fibers through variants of the methods described above, this has not deterred workers from other approaches. Particularly noteworthy are sol-gel approaches, such as illustrated by the studies of Matejec et al.[154,155] In this approach, liquid organometallic precursors are added to water and hydrolyzed, forming a water-rich, low-density semisolid called a gel. The prototypical precursor for SiO_2 is silicon tetraethyl orthosilicate (silicon tetraethoxysilane), generally referred to by the acronym TEOS. Simplistically, the reaction between TEOS and water is as follows:

$$Si(OCH_2CH_5)_4 + 2H_2O \rightarrow Si(OH)_4 + HOCH_2CH_3 \quad (1)$$

$Si(OCH_2CH_5)_4$ is TEOS · $Si(OH)_4$ is referred to generically as silicic acid, and the other reaction product is ethanol. The silicic acid undergoes a polycondensation reaction, which goes nearly to completion upon heating the gel to high temperature. This reaction is summarized as follows:

$$Si(OH)_4 \rightarrow SiO_2 + 2H_2O \quad (2)$$

Other organometallics undergo similar hydrolysis and condensation reactions, although as with CVD it is very difficult to find precursors or process variations that permit all to undergo all steps at equivalent rates. Nevertheless, as with MCVD, nearly anything that can be made into a water- or alcohol-soluble salt can be added to a sol-gel, so the compositional flexibility of this process is huge compared to CVD. Indeed, one might argue that we would be making optical fibers by sol-gel were it not for the early success of CVD. This approach also deserves more attention by materials scientists interested in next-generation amplifier materials.

Regardless of how it is made, the preform is drawn to fiber using a conventional draw tower. The draw temperature varies with the exact makeup of the clad glass: silica clad generally requires temperatures on the order of 2100°C, whereas phosphorus- or fluorine-doped clad glass permits lower draw temperatures. Since the preforms can attain much greater length and diameter than any other material considered in this review, the amount of fiber obtained from the preform tends to be quite large as well, on the order of tens of kilometers per preform. Typical draw speeds are 1–5 meters per minute. The fibers can be coated with UV-curable polyacrylate coatings as is used in telecommunications applications, or it can be hermetically sealed using sputtered carbon or metallic coatings.

Fiber Losses, Strength, and Reliability

As noted in the beginning of this chapter, the two general sources of fiber loss are transition metals and water. Transition metal impurities were studied in detail by Schultz,[156] and a figure from his paper illustrating the impact of

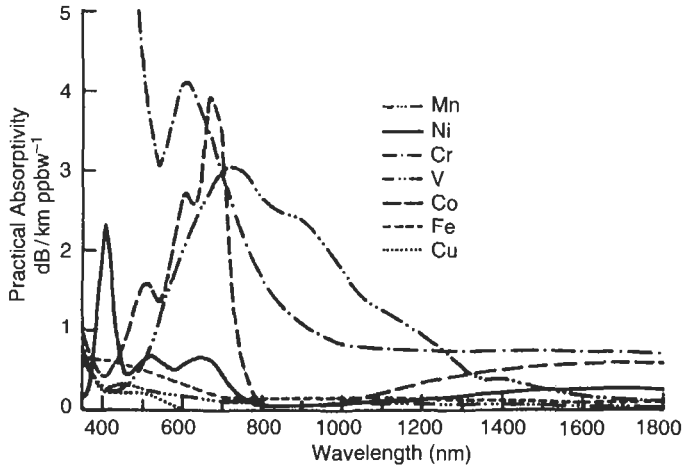


Fig. 10 Absorption losses in Al-doped silica induced by transition elements, in dB/km/ppm. From the study of Shultz, reprinted with permission.[156]

transition metal impurities throughout the visible and near IR is shown in Fig. 10. The most notable contaminants in the infrared are chromium (Cr), cobalt (Co), vanadium (V), and nickel (Ni), though iron is a ubiquitous contaminant and so may have a disproportionately high importance. In practice, these tramps are very effectively scavenged during the halide purification step in OVD and IVD processes and tend to be removed fairly efficiently during plasma CVD as well. Therefore, they tend not to be the limiting factor in fiber losses.

Hydroxyl ions may be more difficult to remove than in conventional Ge-doped silica transmission fiber. This is because the free energy of formation of aluminum hydroxide is very high compared to that of silicic or germanic acid, and thus it is more difficult to strip water away. Phosphoric acid also has a very high free energy of formation, and when heated it partially polymerizes into a material called polyphosphoric acid that is notoriously difficult to strip of water. Core glasses containing Al, P, or both also tend to be subject to rehydration after consolidation, so great care must be taken between consolidation and collapsing the center line. Again, since center line collapse occurs as a separate step in the OVD process, but is the final step in consolidation in the IVD process, IVD has a minor processing edge in this regard. As with many aspects of fiber fabrication, good hygiene and technique can overcome these obstacles for either the OVD or IVD process.

Unlike fibers made from fluoride glasses or alternative oxide glasses, devitrification during fiber fabrication is not of great importance for silica-clad fibers, save that at very high alumina contents care must be taken to avoid mullite formation. Pure silica has one some of the highest elastic constants of any glass,

and since fiber strength tends to scale with elastic constants, silica-based materials tend to make very strong fiber. Other than devitrification materials such as mullite, break sources in silica fiber tend to originate from the surface of the preform, and so with careful handling very high fiber strengths are obtained. Typical values are 3–4 Gpa—compared with 0.5–0.8 GPa reported for the best fluorozirconate fibers. High strength lends itself to ease of handling: the fiber is more forgiving when twisted or bent, when a load is applied along its length, or when it is accidentally pinched or abraded. Therefore, industries that manufacture fiber-based devices tend to put a premium on high fiber strength, whereas the end user is mainly concerned with reliability.

Reliability certainly benefits from high fiber strength, but of equal (or greater) importance is how the strength varies when a relatively low level of stress is applied for a very long time, as in the coil in an amplifier. An indication of this performance is obtained from a metric called the dynamic fatigue coefficient, or n-value, which is determined by applying stresses to fiber at a very slow rate until failure is obtained. A plot of the percentage of fiber pieces that fail vs. applied stress is called a Weibuld distribution, and from the slope of the trend one obtains the dynamic fatigue coefficient. This is generally determined after the fiber has been "aged," or after ambient or accelerated conditions for days or weeks after fiber draw. When the dynamic fatigue coefficient is high, the fiber behaves as though it stretches under applied stress, whereas when it is low it acts brittle. Brittle fiber typically results from break sources across the fiber surface, which produce minute cracks that are propagated by corrosion. Therefore, highly durable materials tend to produce high dynamic fatigue coefficients, though other factors are very important as well, such as the distribution of stress across the fiber surface.

A rough metric for comparing two fibers is the product of the simple tensile strength and dynamic fatigue coefficient. For normally coated silica-clad telecommunications fiber, the tensile strength is typically on the order of 5 GPa, and the dynamic fatigue coefficient is approximately 20, so the fiber reliability metric is 100. A hermetic coating on clean fiber eliminates the possibility of corrosion of any kind, and so the n-value becomes effectively infinite.

Applications

Although originally investigated for its use in Nd-doped fiber lasers and amplifiers, Al/Ge/Si and variants are now used almost entirely for erbium-doped fiber amplifiers and lasers. Though early investigators used 800 nm diode pumps,[157] the 980 and 1480 nm absorptions were quickly shown to produce highly inverted, high-power fiber lasers[158] and amplifiers.[159] Most modern amplifiers use a combination of 980 and 1480 nm single-mode diode pump lasers, whereas erbium fiber lasers tend to use 980 nm diodes alone. Al-doped and Al/P-codoped silica fiber amplifiers are used throughout the world in telecommunications networks. The literature is filled with applications for

Er-doped Al/Ge/Si amplifier fibers; many of these will be discussed later in this chapter. Phosphorus- or P/Yb-codoped silica fibers are used in high-power 1530 nm lasers.

Tellurites

The origins of tellurite glasses are obscure. Tellurium oxide (TeO_2) itself is a poor glass former, requiring heroic quench rates to avoid forming crystalline TeO_2. Addition of oxides of monovalent or divalent cations, particularly those of alkalis, barium, zinc, and lead, causes a dramatic stabilization of the glasses, such that glasses are readily obtained from compositions with as much as 80 mol% TeO_2 almost without regard to the exact identity of the dopant. As such, multicomponent tellurite glasses have been known for at least 30 years; however, they remained largely laboratory curiosities, interesting because of their high refractive indices, but for not much else.

All of this changed rather dramatically in 1994. If a research paper could launch a thousand ships, then the seminal study of Wang et al.[160] concerning optical applications of tellurites might be the one to do it. Prior to this study, very little work had been performed to analyze rare earth element spectroscopy in tellurite glasses. Since then, not only have many new tellurite systems been characterized both in terms of properties and rare earth element spectroscopy, but working devices and system tests have been performed in support of the most promising application for tellurite glasses, erbium-doped fiber amplifiers.

Wang et al. identified very stable compositions in the system Na_2O–ZnO–TeO_2 ($5Na_2O$–20ZnO–$75TeO_2$ was found to be particularly stable) and not only characterized nearly all properties relevant to fiber draw, but actually fabricated moderate loss (~1 dB/m) fiber from a particular core/clad combination. They also characterized the spectroscopy of neodymium, praseodymium, erbium, and thulium in the preferred core glass composition, and determined more generally the effects of various glass components upon erbium 1.55 μm emission lifetime. Finally, they characterized the effect of holmium codoping in thulium-doped tellurites to reduce the lifetime of the 1.9 μm transition in favor of the 1.48 μm transition. It is not too much of an exaggeration to say that there was little left to do in the system in question but make low-loss fiber and test devices in systems.

Composition and Rare Earth Element Spectroscopy

The study of Wang et al. catalyzed a large effort to evaluate rare earth element spectroscopy in various tellurite systems. Wang et al.[161] reported 1.3 μm emission characteristics of neodymium and praseodymium in tellurite glasses. They concluded that ytterbium codoping of praseodymium-doped glasses was effective in facilitating energy transfer from the $^2F_{3/2}$ level of Yb^{3+} to the $^1G_4$

level of Pr^{3+}, hence improving pump power absorption. Man et al.[162] performed a Judd-Ofelt analysis of Pr^{3+} in $8Na_2O$–$20ZnO$–$72TeO_2$ glass and concluded that the 24 μs lifetime of the $^1G_4 \rightarrow {}^3H_5$ (1.3 μm) optical transition and 2.6% calculated quantum efficiency made tellurite fiber competitive with fluorozirconate fiber as a host for Pr^{3+}. More recently, Tanabe et al. evaluated sensitizing 1.3 μm emission from Pr^{3+} by codoping Pr-doped tellurite glasses with Yb^{3+}. The authors report that energy transfer from Yb^{3+} to Pr^{3+} was greater than 90% even at very low Pr/Yb ratios, resulting in a substantial increase in fluorescence intensity from Pr^{3+}.

Of greater relevance to modern telecommunications systems are the spectroscopies of Er^{3+} and Tm^{3+}. Figure 11 compares the emission spectrum of a well-optimized tellurite glass developed by one of us (Ellison) at Corning Inc. with that of a well-optimized aluminum-doped silicate glass. The aluminosilicate emission spectrum basically fits inside that of the tellurite host, and the tellurite glass shows much stronger emission intensity at long wavelengths. This illustrates the promise of tellurite hosts for C- and L-band EDFA applications. The 1480 nm emission spectrum of a tellurite host is very similar to that of ZBLAN, and calculated quantum efficiencies are nearly the same. Thus, tellurites would seem more attractive than ZBLAN or Al-doped silica for L- and, perhaps, C-band EDFAs, and possibly as attractive as ZBLAN for Tm-doped fiber amplifiers (TDFAs).

McDougall et al.[163] calculated Judd-Ofelt parameters for Er^{3+} and Tm^{3+} in binary tellurite glasses as a function of rare earth element concentration and showed that radiative lifetimes for both remained high to quite high rare earth element concentrations, suggesting high solubility and the potential for

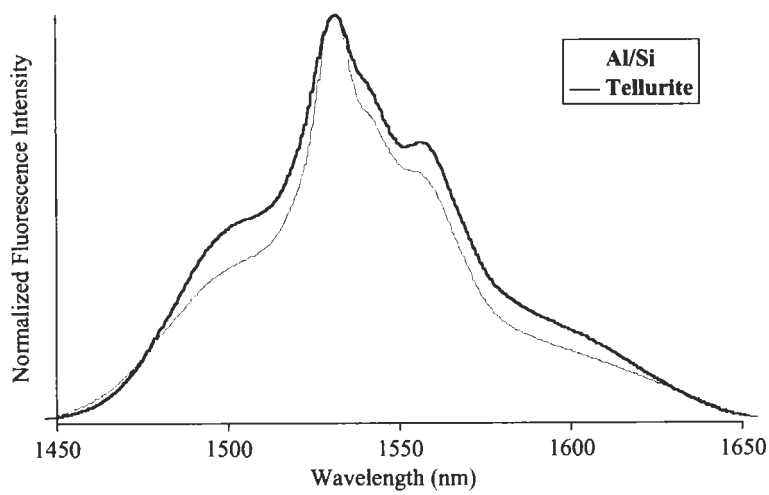


Fig. 11 Normalized 1550 nm Er^{3+} emission spectra of tellurite and Al-doped silica glasses.

low levels of rare earth element clustering. Le Neindre et al.[164] examined Er^{3+} 1550 nm emission linewidth in a base composition equivalent to the preferred composition of Wang et al. but containing two or more alkali cations. They showed that the broadest emission bands were obtained in glasses containing a mixture of alkali cations, particularly mixtures of potassium and lithium oxides. Jha et al.[165] evaluated Er^{3+} emission lifetimes as a function of OH^- and Er_2O_3 concentration and the absolute concentrations of Na_2O and ZnO. They found that OH^- had a significant impact on erbium emission lifetimes, though at concentrations so high that OH^- absorption at 1450 nm would be of comparable or greater concern. They also found a complex relationship between the full-width-at-half-maximum (FWHM) of the erbium 1550 nm emission band and bulk composition, with a mixed sodium-zinc glass providing the broadest spectrum.

The main limitation of using tellurite fibers for EDFAs is the very long 980 nm emission lifetime, typically on the order of 500 μs in glasses with optimally broad 1550 nm emission spectra. Various methods have been employed to reduce the 980 nm emission lifetime. One of the more promising involves codoping Ce^{3+} into the Er-doped tellurite host glass. In this approach, one takes advantage of a phonon-mediated nonradiative energy transfer between the Er^{3+} $^4I_{11/2}$ (980 nm) transition to the Ce^{3+} $^2F_{5/2}$ then back to the Er^{3+} $^4I_{13/2}$ (1550 nm) transition. The authors conclude that at modest cerium doping levels (~1–2 mol%), a large improvement in energy transfer is obtained without excessive impact on the 1550 nm emission lifetime.

Consideration of Tm-doped tellurite glasses for TDFAs is a more recent application. Naftaly et al.[166] and Shen et al.[167] have recently compared Tm-doped tellurite and fluorozirconate fibers and conclude that tellurites may provide certain advantages in practical applications. Choi et al.[168] investigated emission spectra of Tm-doped tellurite glasses codoped with Dy^{3+} in an effort to identify means to suppress amplified spontaneous emission (ASE) power at 810 nm. They conclude that a layer of Dy^{3+}-doped glass close to the Tm-doped core might be effective in suppressing ASE. Cho et al.[169] concluded that the same approach might work for fluorine-doped tellurite glasses. It should be noted that a similar scheme involving Tm-Ho codoping in ZBLAN was demonstrated Sakamoto et al.[170] to be effective in suppressing 800 nm ASE. Tanabe et al.[171] proposed a solution much like that of Choi et al. involving a Nd^{3+}-doped layer next to a Tm-doped core in a ZBLAN-based fiber. Both may deserve attention for tellurite-based TDFAs as well.

Finally, Jiang et al.[172] evaluated the emission cross-section of the $^2F_{5/2} \rightarrow {}^2F_{7/2}$ (980 nm) transition of Yb^{3+} ions in tellurite glasses. They identified compositions in the system $ZnO–La_2O_3–TeO_2$ with large emission cross-sections and long emission lifetimes that appear to be well suited for Yb fiber laser applications. Though this is somewhat off the beaten path for conventional photonics applications, Yb fiber lasers may have a role to play as inexpensive, high-power pump sources to replace 980 nm diodes and as pump sources for

cascaded Raman amplifiers. The 980 nm pump source option is discussed later in this chapter.

In addition to generating numerous studies concerning rare earth element spectroscopy, the ground-breaking study of Wang et al. also generated numerous efforts to improve on the original compositions. For example, Duverger et al.[173] showed that the intensity of the highest-lying Raman mode in MO-TeO_2 tellurites decrease systematically with increasing MO concentration, and likewise (though not so strongly) as the atomic mass of M increases from Mg to Zn to Pb. This suggested that glasses with lower TeO_2 contents might be more effective hosts for phonon-sensitive rare earth elements. Braglia et al.[174] determined the viscoelastic properties and thermal stabilities of xNa_2O–20ZnO–(80–x)TeO_2 glasses and melts as a function of sodium oxide concentration. They demonstrated that there is a tradeoff between the viscosity of a glass at any given temperature and the stability of the glass, with the most stable glasses (most sodium-rich) being less viscous than the least stable glasses. However, T_x–T_g values were all in excess of 130°C and reached values as high as 179°C. This is to be compared with values of 80–100 for stable fluorozirconate glasses. This demonstrated that a wide range of compositions in this system could potentially be drawn into low-loss fiber.

Tellurium oxide has also been explored as a codopant in systems consisting largely of a different glass former. Of these, one of the more interesting systems are those in which TeO_2 is added to a germanate base glass, referred to generically as tellurium germanates. Pan and Morgan characterized Raman spectra and glass stability[175] and optical transitions of Er^{3+} ions[176] in glasses from the system PbO–CaO–TeO_2–GeO_2. At TeO_2 contents of 30 mol% or greater, no crystallization temperature (T_x) was seen, indicating a high level of glass stability. Raman spectra showed a decrease in the intensity of the highest lying vibrational modes as TeO_2 content increased, an observation borne out in increased upconversion efficiency for erbium ions dissolved into the glass matrix. Feng et al.[177] examined the effect of hydroxyl ions on erbium 1550 nm fluorescence lifetimes in Na_2O–ZnO–GeO_2–TeO_2–Y_2O_3 glasses. They found that hydroxyl ions had a small effect on emission lifetimes, but that it could be removed in any case by bubbling the melts with a mixture of carbon tetrachloride (CCl_4) and oxygen.

In some tellurite base glasses, it is possible to replace a substantial fraction of the oxygen atoms with halides, generally two halide ions for one oxygen ion. Sahar and Noordin[178] examined glass formation and physical properties in the system ZnO–$ZnCl_2$–TeO_2, and identified compositions with good glass stability at high (nominal) chlorine contents. No analysis of rare earth element spectroscopy was provided. Sidebottom et al.[179] analyzed Raman spectra, Eu^{3+} phonon side-band spectroscopy and erbium and neodymium fluorescence lifetimes in ZnO–ZnF_2–TeO_2 glasses as a function of fluorine content. The side-band spectra showed close agreement with the Raman spectra (Fig. 12). They found that replacing oxygen by fluorine substantially increased

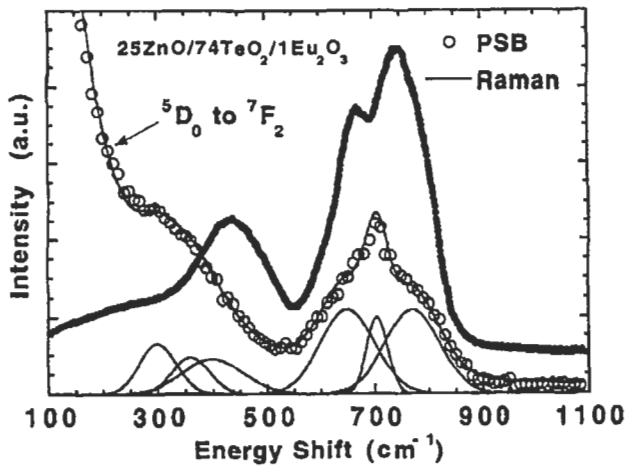


Fig. 12 Comparison of Eu^{3+} phonon sideband and Raman spectra in a tellurite host. Reprinted from Ref. 179 with permission.

phonon-sensitive lifetimes, but reduced glass stability as well. They interpreted this as a result of incorporating fluorine into the rare earth element environments; however, fluorine generally contributes to lower OH levels in oxide glasses, but OH measurements were evidently not performed. Ding et al.[180] identified a broad range of relatively stable glasses in the system PbX_2–TeO_2, where X = F, Cl or Br. The glasses containing bromine in particular exhibited long 1550 nm emission lifetimes and broad 1550 nm emission spectra, and also had the highest T_x–T_g values of any of the glasses referenced in their study.

Glass Synthesis and Fiber Fabrication

Tellurite glasses are much simpler to fabricate than fluoride glasses because they are stable when melted or reheated in air. Indeed, tellurium is prone to reduction when melted in inert atmospheres, so a relatively oxygen-rich environment is a prerequisite to obtaining clear, colorless glass. As noted above, tellurite glasses can be melted under conditions akin to reactive atmosphere processing provided that excess oxygen is also present. These methods can be used to drive down OH^- levels, to partially scrub transition metal contaminants, and when bubbling is employed, to improve glass homogeneity. Glasses are typically prepared from high-purity oxides and (in the case of alkalis and alkaline earths) carbonates, and may be subjected to heat treatments prior to melting to extract residual OH^- from the powdered batch materials. Tellurite glasses are stable in gold crucibles, though refractory or fused silica crucibles can also be used with success. Once fabricated, the glasses are very stable in air and can be handled without recourse to very dry or inert atmospheres,

though handling under such conditions may reduce contamination and hence improve fiber losses.

Given similar melting temperatures, viscosity curves, but somewhat improved glass stability, it is not surprising that those fiber fabrication methods that work best for fluorozirconate glasses also work well for tellurites. The first core/clad fiber reported by Wang et al. was produced by a rod-in-tube method, with the tube fabricated by suction casting. Where methods are described at all, the preferred method appears to be reduced pressure casting followed by redraw and overcladding with tubing produced by suction-casting or preform drilling. As noted by Wang et al., the bending strength of simple rod-in-tube tellurite fiber improves on that of ZBLAN, suggesting that reliability may be less of an issue for tellurites than for fluoride glasses. As with fluorides, the main break source in tellurites is likely to be crystals formed during fiber draw, and devitrification products are also likely to prove the limiting factor in reducing losses in tellurite fibers.

Applications

Photonic applications for tellurite fibers have focused mainly on C- and L-band fiber amplifiers. Mori et al.[181,182] reported the first complete characterization of amplifier performance of erbium-doped tellurite fibers. Small signal gain spectra from their fiber showed broad, relatively flat gain in the L-band extending from 1560 to 1620 nm with approximately a 20% maximum point-to-point variation in gain intensity (gain ripple). The small signal gain spectra of Al-doped silica and fluorozirconate fibers fit neatly inside that of the tellurite, illustrating the potential of these materials for broadband L-band amplifiers. This L-band application is most germane; the 980 nm emission lifetime of tellurite fiber is too long to permit efficient 980 nm pump for C-band applications. On the other hand, 1480 nm pump makes little contribution to noise figure in the L-band. The data of Choi et al.[183] suggest that 980 nm pump power conversion efficiency may be enhanced by codoping with Ce^{3+}, but no device data are reported. Chryssou et al.[184] provide modeling results indicating that Er-doped tellurite fiber amplifiers might hold a decisive advantage compared to Al- or Al/P-doped silica fiber amplifiers. More recent results concerning tellurite fibers in amplifier applications will be discussed later in this chapter.

The most important factor limiting use of tellurite-based fibers in real systems is the extraordinarily high refractive indices and nonlinear coefficients of tellurite glasses. A high refractive index creates substantial challenges with regard to pigtailing as it is critical to minimize reflection losses in amplifier design. A high nonlinear coefficient manifests itself in cross-phase modulation and four-wave mixing, a consequence of the high intensity of pump and signal light in fiber amplifiers. Indeed, high-nonlinear-coefficient fibers can actually be used to produce light amplification relying entirely on the physics that produces four-wave mixing, a process referred to as parametric amplification.

Four-wave mixing produces contamination of one signal by another and tends to increase in magnitude directly with fiber length. The persuasive study of Marhic et al.[185] shows that cross-phase modulation produced in a 2.4 meter tellurite EDFA is roughly equivalent to that produced in 33 meters of Al-doped silica EDFA. Sakamoto et al.[186] show that by expanding the mode field and increasing the erbium doping level in a tellurite EDFA, four-wave mixing can be reduced, but at the expense of power conversion efficiency. Whether erbium levels can be raised so high that this problem can be surpassed is not obvious as of this writing.

A second, perhaps less serious problem concerns photosensitivity of tellurite glasses under UV exposure. Prohaska et al.[187] show sodium-zinc tellurite glass (the preferred glass for fiber draw) undergoes an irreversible increase in optical absorption in the visible and UV after 2 minutes of exposure to a 248 nm KrF excimer laser operating at a fluence of 50 mJ/cm^2 at 30 Hz. These dosage levels are very high, yet power densities in high numerical aperture cores can approach this fluence level. Given the demonstrated propensity for erbium and thulium ions in tellurite glasses to upconvert to visible and near UV wavelengths, it is possible that induced photodarkening may be a long-term or systemic problem for tellurite glasses. Clearly more research is needed on this subject.

ANTIMONY SILICATES

In 1997, one of us (Ellison) found that erbium in antimony silicate glass (Sb/Si) exhibited promising 1530 nm spectroscopy, appearing to be at least comparable to well-optimized tellurite glasses. A major effort was undertaken to convert compositions in this system into low-loss optical fibers for use in amplifiers and lasers. Antimony silicates make high-quality, low-loss fiber and can accommodate high levels of rare earth elements without significant clustering, but have lower refractive than tellurite glasses and so do not suffer from equivalent optical nonlinearity.

Composition and Rare Earth Element Spectroscopy

Antimony oxide is commonly used as a flame retardant in the textile industry and as a fining agent in glasses. It is a poor glass former itself, tending to devitrify on cooling unless rapidly quenched. At roughly the same time that good rare earth element spectroscopy was found, it was also shown that glass formation takes place throughout the binary antimony trioxide-silica (Sb_2O_3–SiO_2) system. Further, a small amount of B_2O_3, SiO_2, or GeO_2 is all that is needed to produce stable Sb_2O_3-rich glasses, and at intermediate levels the glasses are quite resistant to devitrification. This permits manufacturing high-quality glass in multi-kilogram quantities.

At the other end of the binary system, liquidus temperatures plunge as Sb_2O_3 is added to SiO_2, and melts throughout the system (except nearly

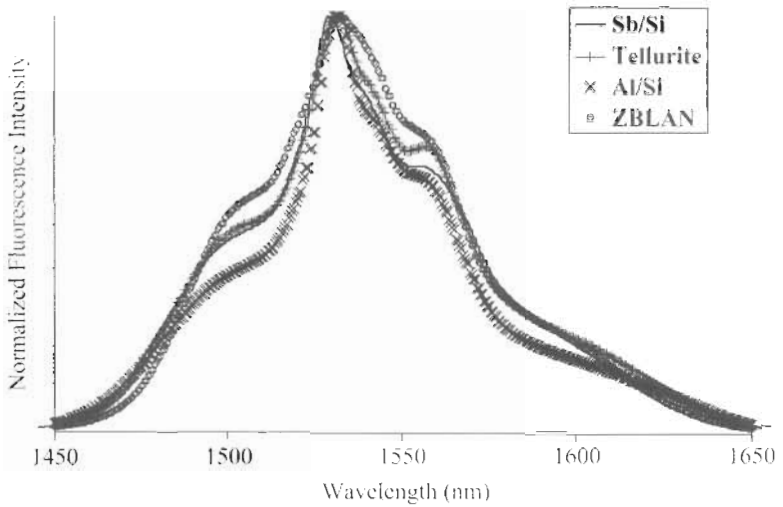


Fig. 13 Normalized 1550 nm Er^{3+} emission spectra of Al-doped silica, tellurite, ZBLAN, and Sb/Si glasses.

pure Sb_2O_3) are quite stable against devitrification when cooled below their liquidus temperature. Sb_2O_3 has Lewis acid/base characteristics that are extremely similar to silica, and so glasses throughout the system have very high durability compared to fluoride or tellurite glasses.

Beyond a certain Sb_2O_3 level, the spectroscopy of a rare earth element shows little evolution with Sb_2O_3 content, though the level varies from one rare earth element to another. At any particular Sb_2O_3 concentration, erbium 1550 nm spectroscopy in particular is quite insensitive to the presence of other components in the glass, though this is particularly true of glasses with 30 mol% or more Sb_2O_3. This lends considerable flexibility to adjusting the properties of the glass to obtain specific goals (e.g., viscosity, durability, stability) without adversely impacting the desired optical performance of the fiber amplifier.

Figure 13 is a comparison of 1550 nm emission spectra of Er^{3+} in well-optimized Al-doped silica, tellurite, ZBLAN, and Sb/Si glasses. The peak of the emission in Sb/Si is blue-shifted relative to tellurite or Al/Si hosts, nearly to the peak position seen in ZBLAN, but on the long wavelength side extends as far as seen in tellurites. The emission cross-section at long wavelengths is comparable to that of the tellurite glass, but much greater than that of Al/Si or ZBLAN. The spectroscopy of Tm^{3+} and Yb^{3+} in Sb/Si is discussed later in this chapter.

Glass Synthesis and Fiber Fabrication

Antimony was once considered a useful means to raise the refractive index of the core glass in silica-based telecommunications fiber. Shimizu et al.[188,189] first

reported making Sb-doped silica fibers by vapor axial (chemical) deposition, and achieved minimum fiber losses of 7 dB/km. More recently, Susa et al.[190] reported making Sb-doped silica fibers using sol-gel methods. Losses were much higher than in fibers produced by VAD, which the authors attribute to charge-transfer interactions between +3 and +5 antimony ions in the glasses. On the other hand, low OH contents (<3 ppm) were obtained through careful processing, illustrating the promise of this technique for producing homogeneous glasses in this system.

At Corning Inc., antimony silicate glasses are currently manufactured from high-purity oxide batch materials melted in refractory crucibles. It has been found that glasses melted in air must be held for some period at temperatures greater than 1000°C so as to minimize formation of cervantite, Sb_2O_4, a crystal containing equal concentrations of Sb^{3+} and Sb^{5+}, which is stable well above its melting point for extended periods of time. Indeed, cervantite appears to be the most stable form of antimony oxide in air between approximately 400 and 950°C. The presence of even traces of Sb^{5+} in a glass produces yellow coloration noticable to the eye and high loss in the infrared.

Antimony is considered toxic, having some chemical affinity with arsenic. It is not clear what valence state is responsible for this toxicity, but in any case most regulatory agencies in most countries have strict limits on workplace exposure to antimony. Despite its apparent chemical affinity with arsenic, Sb_2O_3 is not particularly volatile, and fuming above antimony-bearing melts is low unless the concentration of Sb^{5+} is very high. This means that while careful air-handling is required to avoid contamination, exotic melting conditions (e.g., glove box or controlled atmosphere) are not required; again, this is an asset for large-scale production.

As with charge transfer interactions between antimony ions of different valences, charge-transfer interactions between antimony of any valence and transition elements produce strong coloration and, apparently, unusually high loss in the infrared. Great care must be taken to avoid transition metal contamination even at the tens of parts per billion level. Antimony oxide is typically obtained by either wet chemical precipitation from aqueous solutions or from oxidation of its principle ore, stibnite (Sb_2S_3). The former introduces substantial amounts of water but can be used to obtain very high-purity material. The latter tends to remove most water but leaves the transition metal impurities largely intact: this is particularly true for copper, the most prevalent transition metal other than iron in antimony sources. Copper produces strong green coloration in antimony silicate glasses at a level of 1–2 ppm.

Antimony silicate glasses have been drawn to fiber by numerous means. As noted above, VAD preforms and sol-gel rod-in-tube methods have been applied. Preforms have also been prepared using variants of build-in casting, which permits use of an antimony silicate as a clad glass. Fibers made from preforms with antimony silicate clad glasses tend to have fairly poor strength, on the order of 0.8–1.2 GPa. This is due largely to surface defects

produced by oxidation of antimony and formation of cervantite (draws performed under oxidizing conditions), reduction of antimony and formation of colloidal antimony metal (draws performed under reducing conditions), or silica-polymorph devitrification (draws performed at elevated temperature under oxidizing conditions). In addition to contributing to poor tensile strength, these defects also contributed to unacceptable fiber losses, similar to those reported for ZBLAN-based fibers.

In part to address this problem of weak fiber, a triple crucible approach was developed to prepare antimony silicate fibers from precursor glasses. Three concentric crucibles are nested together. Core and clad glasses are loaded into the innermost crucibles and a glass equivalent to the clad but loaded with infrared absorbers is loaded into the outermost crucible. The crucible positions with respect to one another are adjusted on-the-fly until adequate geometry is obtained.

An optical photomicrograph of a fiber produced by triple crucible is shown in Fig. 14. The overclad glass is discolored relative to the clad and core due to the high concentration of transition and rare earth element absorbers. Typical fiber attributes are as follows:

Strength:	1.8–2.2 GPa
n-value:	>50
Minimum loss:	30–50 dB/km
Core centering:	<0.5 μm
Length:	4–10 km

The tensile strength is much higher than tellurite or fluorozirconate glasses but much lower than CVD-delivered silica fiber. This is offset in part by the very high dynamic fatigue coefficient (*n*-value), which, using the metric discussed

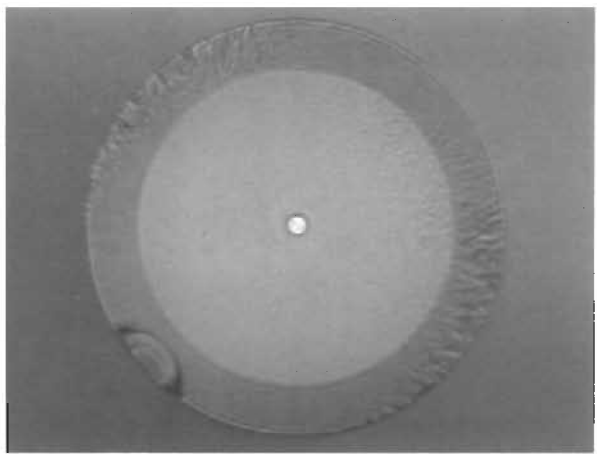


Fig. 14 Optical photomicrograph of a Sb-silicate fiber produced by direct draw triple crucible fiberization.

previously, implies mechanical performance in the field comparable to silica-based fibers. The maximum length of fiber that can be obtained from a single triple crucible fiber draw is not known, but is believed to be on the order of 80–100 km.

Applications

Though originally developed for erbium-doped fiber amplifiers, Sb/Si fibers are attractive hosts for Tm^{3+} and Yb^{3+}. Each of these applications will be discussed in detail below.

Device Applications of New Materials

Now that we have reviewed the material properties of rare earths in various host glasses, we switch to examining the application of these materials to practical devices. We begin with a review of the evolution of the EDFA, emphasizing the interplay between spectroscopy and the engineering of gain flatness (for valuable comparison of the range of erbium spectroscopies obtained in various oxide and fluoride hosts, consult the review by Miniscalco[191]).

ERBIUM DOPED FIBER AMPLIFIERS

It is a wonder of nature that the EDFA gives efficient amplification in the same wavelength region where silica fibers have their lowest loss. Over the last decade long-haul systems have evolved from single-channel low bitrates to multichannel systems at data rates as high as 40 Gbits.[192,193] This bandwidth explosion has been achieved by squeezing more and more bandwidth out of the EDFA. Furthermore, to avoid nonlinear impairments the gain over the EDFA bandwidth must be kept flat typically to <0.5 dB over the full range for a gain of 25 dB. To achieve this, materials with broad Er^{3+} gain and high efficiency are combined with accurate gain flattening techniques.[194]

Erbium Amplifier Bands

Trivalent erbium gives gain over a broad bandwidth typically extending from 1530 nm to 1605 nm. However this bandwidth is not efficiently accessible in a single amplifier. It is usual to pair a so-called C-band EDFA operating in the 1530–1560 nm region with an L-band amplifier in the 1570–1605 nm range. It will be shown that a material optimized for C-band will not necessarily serve well in the L-band. In fact for ultimate bandwidth different compositions are indicated for each amplifier.

The gain spectra of an EDFA depend on the average inversion of the gain media. The inversion can be controlled by the combination of pump power relative to signal power and the fiber length.[195] Figure 15 illustrates how the

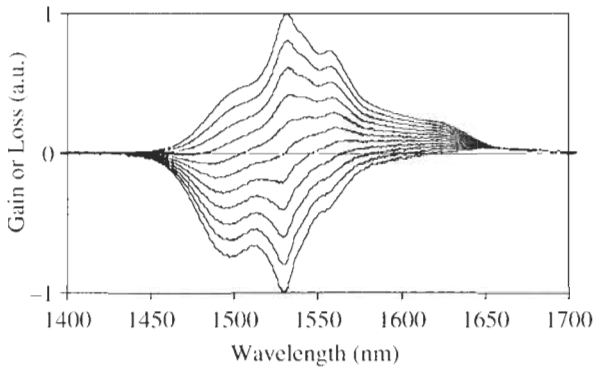


Fig. 15 Normalized gain shape for a generic erbium fiber for inversions of 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0.

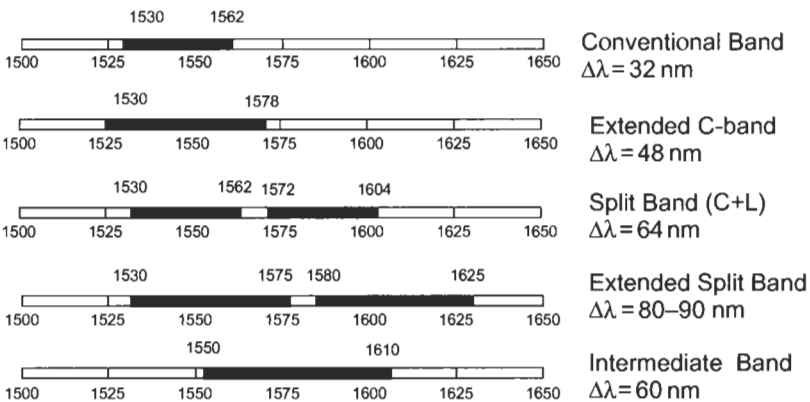


Fig. 16 Evolution of erbium amplifier bands.

spectra evolve with increasing inversion. At low inversion the erbium-doped fiber exhibits loss over most of the band and a low gain at the extreme red end of the band. As the inversion increases, gain is obtained at shorter and shorter wavelengths. Note that the rate of change of gain increases at shorter wavelengths. This is known as gain tilt, a phenomenon that degrades the performance of an amplifier operating away from its design point.

The evolution of erbium-doped amplifier bands is illustrated schematically in Fig. 16. The conventional or C-band corresponds to wavelengths close to the gain peak. This band typically extends from 1530 nm to 1562 nm but can be extended by compositional adjustment. Conventional band amplifiers operate typically at an inversion in the range 0.6–0.65. The long or L-band EDFA, on the other hand, represents gain in the long wavelength tails of the gain spectrum. Such amplifiers typically operate at an inversion level of 0.4

and have relatively flat gain. Once more compositional adjustments can extend the L-band.

There have recently been some reports of wideband EDFAs that straddle the traditional C- and L-bands. Perhaps unsurprisingly such amplifiers operate at inversion levels intermediate between C- and L-bands, i.e., around 0.5. The advantage of such amplifiers is the absence of a band splitter in the signal path. While these amplifiers tend to need deeper gain-flattening filters, they can have advantages in terms of intrinsic noise figure and ease of systems upgrade.

Conventional Band EDFAs

A conventional EDFA operates in the range 1530 to 1560 nm with the bandwidth being set by the ability of the amplifier designer to engineer a solution based on the following:

- the intrinsic spectroscopy of the erbium in the host material;
- the ability to manufacture gain-flattening filters of given depth and slope;
- the pump power at his disposal.

It is therefore desirable to have a host glass with broad and flat gain within the specified bandwidth of an amplifier, thus minimizing the need for filtering and saving pump power. The conventional band of an EDFA evolved naturally from early single wavelength amplifiers designed to operate at the 1532 nm erbium gain peak or at 1550 nm, the loss minimum of single-mode fibers.[196–198]

Al/Ge/Si EDFAs

The first erbium doped fibers were fabricated in a germanosilicate host glass very similar to that used for transmission fiber. However this host had a very narrow C-band spectrum and suffered from serious concentration quenching.[199] An alumina-doped germanosilicate host quickly superseded it as the material of choice.[200]

Aluminosilicate amplifiers can give unfiltered bandwidths of around 18 nm[201–205] and with gain-flattening filters are the material of choice for amplifiers of up to 32 nm bandwidth in the C-band.[206–208] At this bandwidth filter depths of approximately 4–5 dB are sufficient to flatten the gain of a typical 26 dB gain long-haul amplifier.[209]

Fluorozirconate EDFAs

Fluorozirconate fibers based on the ZBLAN composition were initially proposed as a means of broadening EDFAs without the need for filtering. The material does exhibit impressive intrinsic gain flatness. Bayart et al.[210] demonstrated 25 nm bandwidth with a gain ripple of less than 5%. This was a

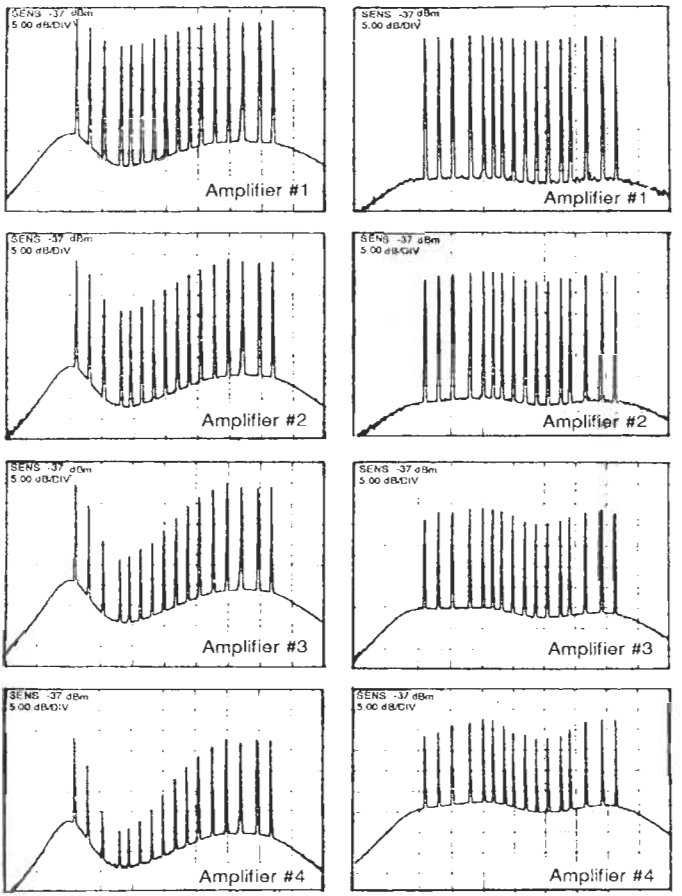


Fig. 17 Evolution of gain ripple in a cascade of silica (left) and fluoride (right) EDFAs.[211]

bandwidth improvement over aluminosilicate of approximately 50%. It was also shown that in a three-amplifier cascade over similar bandwidth a gain flatness advantage of 12 dB could be achieved relative to cascaded silica-based EDFAs.[211] This is illustrated in Fig. 17.

The fluorozirconate fibers remain the material with the flattest intrinsic ripple in the C-band. However its useful bandwidth is severely limited to the aforementioned 25 nm. Although a modest improvement in bandwidth to 28 nm was achieved by hybridizing with an Al/Si amplifier,[212] it soon became clear that gain-flattening filter technology combined with improved amplifier architectures relaxed some of the requirements on the material properties. Consequently erbium-doped fluorozirconate fibers did not succeed commercially.

Sb-silicate C-band EDFAs

The success of filtering technology does not mean there is no benefit to finding materials with intrinsic spectroscopic advantage. It is really the product of gain ripple and bandwidth that is the valuable attribute. Consequently, just as improvements in filtering techniques offset the benefits of the Er-doped fluorozirconate fiber, so a material with broader gain spectra could push the limits of a given filter technology. The new antimony-silicate material described earlier is flatter and broader than aluminosilicate and has some advantages in C-band amplification.

The first reported application of the antimony silicate material by Ellison et al.[213] was a materials-based approach for flattening the Er-fiber gain for a 32 nm C-band amplifier. A hybrid amplifier employing this multicomponent silicate fiber in the first coil and standard aluminosilicate (Al/Si, moderate amount of Al_2O_3) fiber in the power coil resulted in a reduction of the gain ripple from 21% to 10%. The gain ripple of the hybrid amplifier vs. aluminosilicate content is shown in Fig. 18. The optimum split was 30% antimony silicate and 70% aluminosilicate. The performance of the amplifier is shown in Fig. 19.

When the bandwidth increased beyond 32 nm there was little benefit to hybridization. However, the antimony silicate material exhibits much lower ripple at larger bandwidths. Goforth et al.[214] compared the performance of this material to other commercially available fibers. Calculated gain curves are shown in Fig. 20. These particular curves represent optimized inversion for gain flatness in a 48 nm floating band (i.e., the band position was allowed to

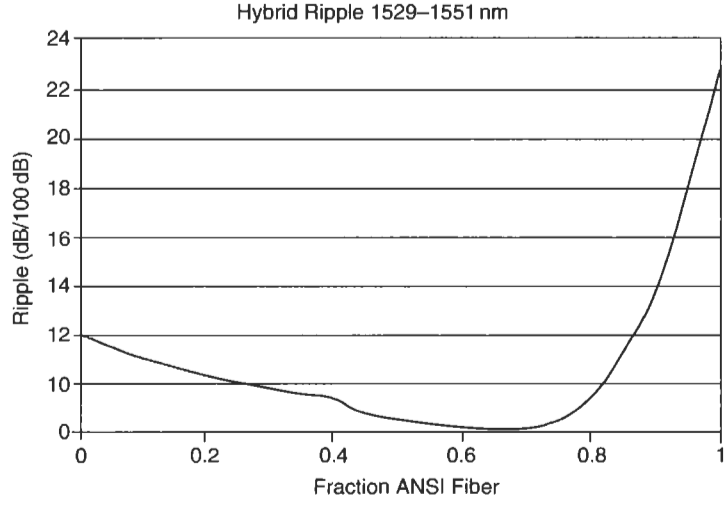


Fig. 18 Gain ripple in an antimony silicate/aluminosilicate EDFA hybrid.[213]

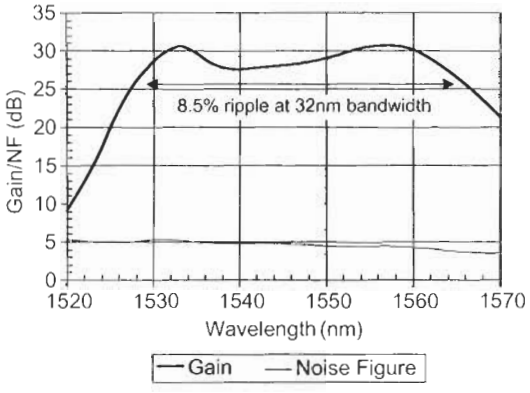


Fig. 19 Gain measurement in a hybrid EDFA comprising 30% antimony silicate and 70% aluminosilicate fiber.[213]

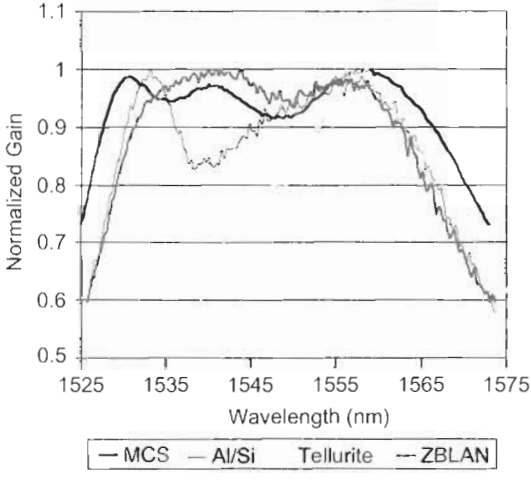


Fig. 20 Normalized gain curves for antimony silicate (MCS) amplifiers compared to other materials.[214]

float to a point of optimum flatness). The antimony-silicate fiber is substantially broader and flatter than aluminosilicate (Al/Si) fiber. The ZBLAN fiber is very flat, but over a much smaller bandwidth than the antimony-silicate fiber. The tellurite fiber, on the other hand, is very broad (well beyond the graph shown), but not particularly flat. Therefore, to achieve the optimal combination of broadness and flatness in the C-band, the antimony-silicate fiber is most appropriate.

Figure 21 illustrates the calculated gain ripple $((\text{Gain}_{\max} - \text{Gain}_{\min})/\text{Gain}_{\min})$ for floating bands at various bandwidths. The improvement in ripple relative

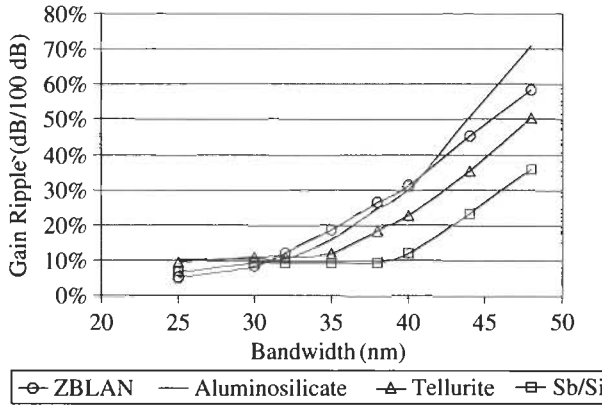


Fig. 21 C-band gain ripple as a function of bandwidth for several gain materials.[214]

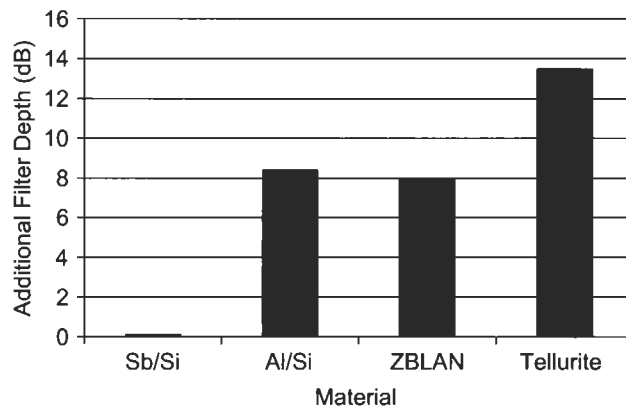


Fig. 22 Additional filter depths at a bandwidth of 48 nm for various materials relative to antimony silicate.[214]

to Al/Si fiber substantially reduces filter depth over floating bands of 32 to 48 nm. This in turn improves efficiency and overall performance, and eases tolerances on filter design. The high erbium solubility of these new materials and the absence of cross-relaxation at high pump power allows high Er dopant levels (3×10^{19} Er ions/cc), or over 60 dB/m peak gain (fully inverted). The ratio of passive loss to peak absorption at 1530 nm is <0.5%, allowing for efficient amplification.

While a 48 nm bandwidth may be achieved with Al/Si fibers, particularly with higher Al_2O_3, they would be very inefficient because of the large filter depth. Assuming an amplifier with 25 dB gain, Fig. 22 highlights the additional filter depth, compared to the antimony-silicate fiber, required for the

various fiber types at their optimum ripple. If the amplifier design required inversions other than the optimum, then the relative filter depth would be substantially greater. Even more of a concern would be trying to reproducibly manufacture the complex filter shapes required for the other fibers; in some cases the filters would be nearly impossible to make.

The gain spectra achieved in an antimony-silicate amplifier is shown in Fig. 23. The gain ripple is approximately 35% at 48 nm. Although this is a fairly deep filter it is readily manufacturable. A gain-equalized amplifier is shown in Fig. 24 where the ripple is reduced to <1 dB. The 48 nm C-band amplifier provides approximately 50% more bandwidth than a conventional C-band amplifier.

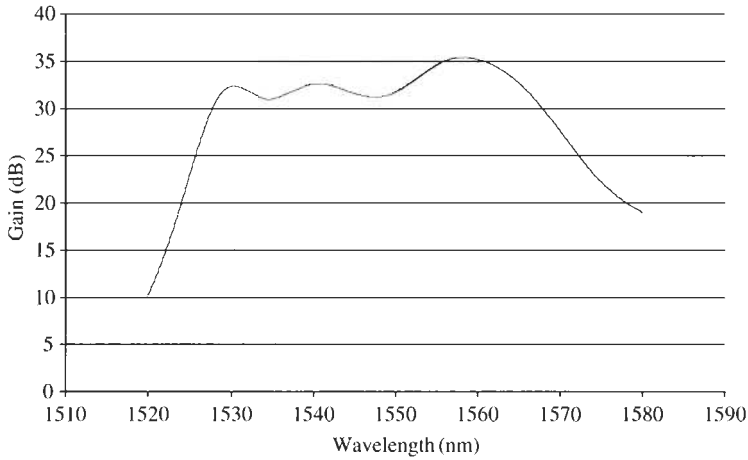


Fig. 23 Unfiltered gain spectrum for C-band antimony-silicate EDFA.[214]

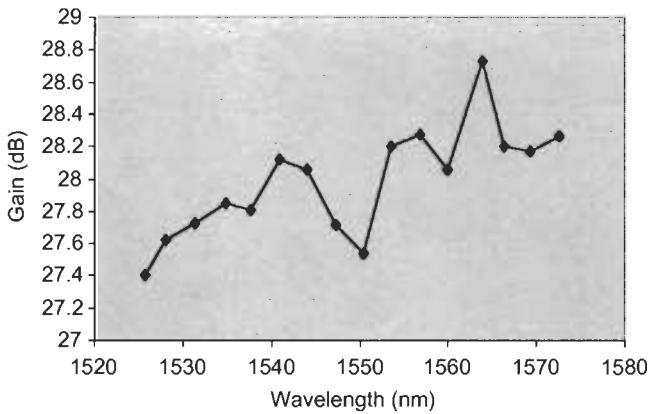


Fig. 24 Filtered gain spectrum of 48 nm antimony silicate amplifier showing gain ripple below 1 dB.

L-band EDFAs

The ultimate limit to C-band gain bandwidth is where the ripple becomes unmanageable with gain-flattening filters. The normal way to increase bandwidth is to use the so-called split band architecture. In this case the amplifier is split into two distinct signal bands each optimized in turn for gain flatness and noise figure. In L-band amplifiers the attributes that are sought differ considerably from C-band. This is because L-band gain is intrinsically much flatter than for a C-band amplifier of similar bandwidth. The problems in the L-band are a low gain coefficient, C-band ASE leading to self-saturation and poor efficiency. The ultimate bandwidth limitation results from signal excited state absorption (ESA).

The low gain coefficient can lead to problems with four-wave mixing in L-band EDFAs. Hence, high concentration fibers are preferred for L-band applications to limit nonlinear impairments. However high concentration can lead to detrimental ion-pair quenching and increased energy migration to the quenching sites. Understanding loss mechanisms in high-concentration fibers becomes much more important than in C-band amplifiers.

The L-band EDFA was initially demonstrated by Massicott et al. in the early 1990s (Fig. 25).[215] There was little interest in this band for several years, however, because of the excellent performance of C-band amplifiers in terms of efficiency and bandwidth. It was only the realization of impending bandwidth exhaust that revived interest in the L-band amplifiers. It is usual to pair an L-band amplifier with a conventional C-band amplifier such as illustrated in Fig. 26 taken from Sun et al.[216]

The bandwidth limitation arises from signal ESA between the 1550 ($I_{13/2}$) level and the $I_{9/2}$ level (Fig. 27). While the peak in this ESA lies around

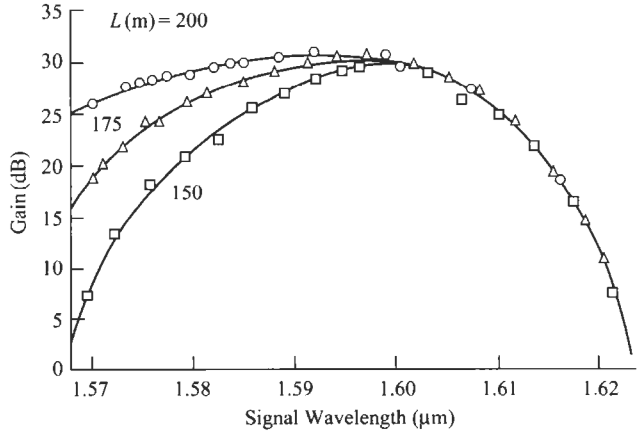


Fig. 25 An early example of L-band amplification. Note the long lengths of erbium fiber needed.[215]

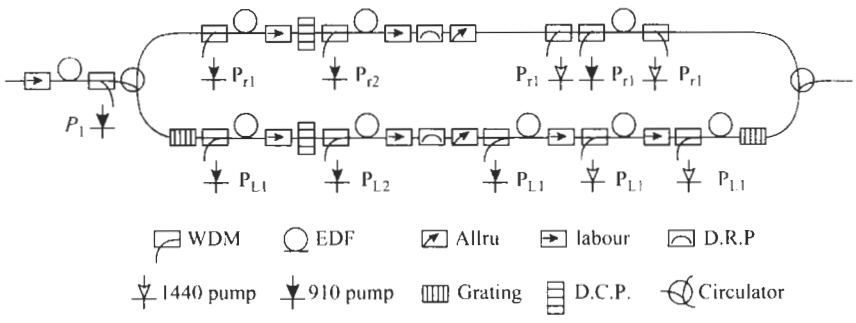


Fig. 26 Schematic diagram of the split-band amplifier architecture.[216]

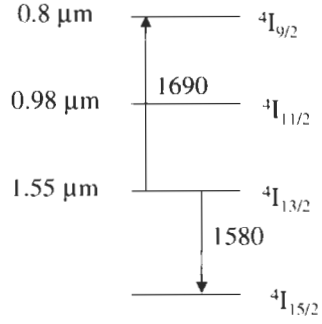


Fig. 27 Energy level diagram illustrating the limitations in bandwidth imposed by signal ESA in L-band EDFAs.

1690 nm, it has a broad tail extending to shorter wavelength. In alumina-doped silica (Al/Si) EDFA fiber the short wavelength tail of SESA substantially impacts L-band gain beyond 1605 nm. In tellurites and antimony silicates, however, operation is possible as far as 1620 nm. This creates 15 nm of additional gain in these materials relative to those based on Al/Si.

Sb-silicate L-band EDFAs

A minor compositional modification to the antimony-silicate material initially developed for extended C-band amplifier was found to shift the short wavelength tail of the signal ESA to longer wavelengths.[217] The result of the shift in ESA is shown in the modeling result in Fig. 28. The curves were calculated using the appropriate Er concentrations, fiber lengths, and inversions, to provide similar gain and broad bandwidth. Choosing the 20 dB gain point (4 dB down from maximum), the figure shows the bandwidths provided by the modified antimony silicate, the original antimony silicate, and type 2Al/Si (54, 49, and 43 nm, respectively). Additional bandwidth can be achieved by using larger filter depths. Alternatively, the bandwidth can be extended by going to lower inversions, at the expense of the short wavelength side.

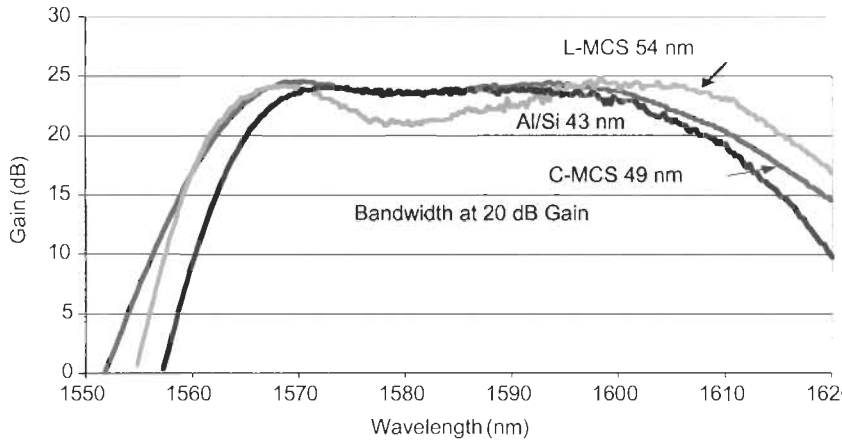


Fig. 28 Modelled L-band spectra for antimony silicate and aluminosilicate EDFAs.[217]

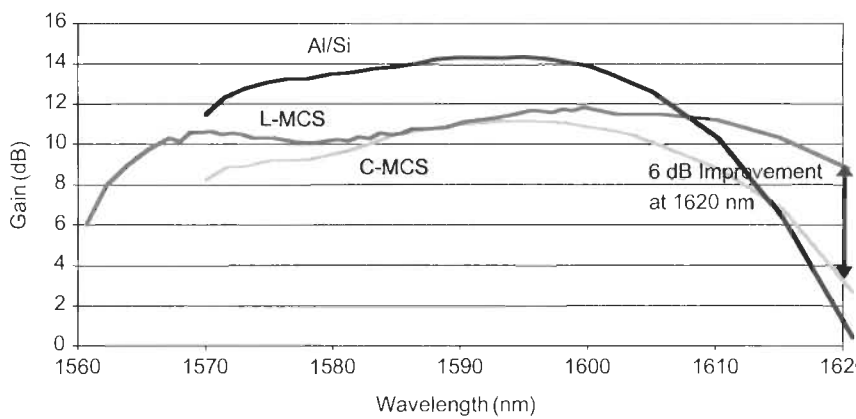


Fig. 29 Experimental gain curves for two antimony silicate compositions illustrating a 6 dB enhancement in the new composition. Also shown is an aluminosilicate gain spectrum.[217]

To demonstrate the gain characteristics of modified antimony-silicate fiber, gain was measured from a simple, single coil using bidirectional 1480 pumps, and monitoring 43 channels. The results are shown in Fig. 29. Both gain measurements were normalized at 1590 nm, and, because the inversions are similar, a direct comparison is possible. All three fibers were measured under different conditions, but this result shows that the modified L-band antimony-silicate fiber provides up to 6 dB improvement at 1620 nm over the C-band antimony-silicate fiber and almost 8 dB improvement above Al/Si, even though that fiber started with a much higher initial gain.

While the C-band composition offers intrinsic value relative to Al/Si (1615 nm maximum gain for Sb/Si vs. 1605 nm maximum for Al/Si), the composition modification further enhances this advantage. The L-band composition extends the useable wavelength range of the L-band to at least 1620 nm.

L-band Tellurite EDFAs

Erbium-doped tellurite amplifiers have similar L-band properties to antimony silicate. The gain cross-section is higher than aluminosilicate and the signal ESA is redshifted. While most of the early work on tellurites reported extreme versions of the super-band concept, the performance in a traditional L-band is excellent. Mori et al.[218] reported a 50 nm band from 1560 nm to 1610 nm with gain variations as low as 10%. For a modest increase in gain ripple the bandwidth could easily be extended to 1620 nm.

Super-band Amplifiers

The super-band concept represents a broadband alternative to the accepted norm of the split band amplifier. One problem with the split band architecture is that provision for both bands must be made up front by deploying a band splitter. This adds cost to the amplifier and degrades performance because of the insertion and splice losses of the additional component. A single-broadband offers a performance advantage and a potential pay-as-you-grow upgrade path. The disadvantage is that super-band amplifiers tend to have much higher gain ripple than either C- or L-band amplifiers. This may not be such a problem in the future if pump power continues to increase in power while falling in price.

Tellurite Super-band Amplifiers

Tellurites are some of the broadest gain materials yet reported for EDFAs. This is illustrated clearly in Fig. 30.[219] The tellurite amplifier gain bandwidth can extend several nm to the blue and red ends when compared with fluorozirconate or silica glass hosts. This makes tellurites an ideal candidate for super-band amplifiers. In an extreme example Mori et al.[220] demonstrated 20 dB of gain over an 83 nm bandwidth (see Fig. 31). The pump power required was a modest 350 mW but the input power was very low at −30 dBm. Yamada et al.[221] demonstrated a 76 nm gain-flattened bandwidth with similar efficiency. More recently, by sacrificing some of the C-band wavelengths, they improved the efficiency while maintaining a 68 nm bandwidth.[222] Further bandwidth reduction and a midstage gain equalizer should make a tellurite super-band more viable.

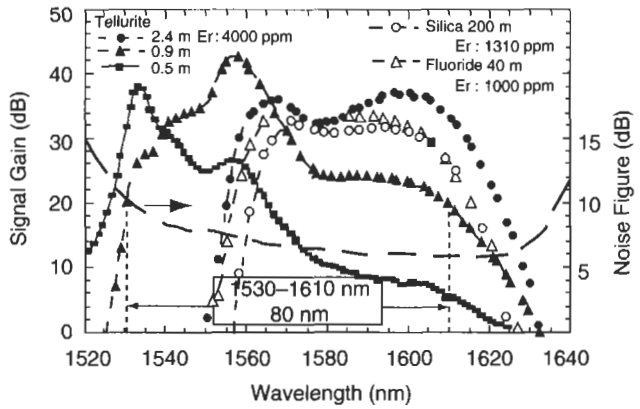


Fig. 30 Gain spectra of tellurite EDFAs at various inversions.[219]

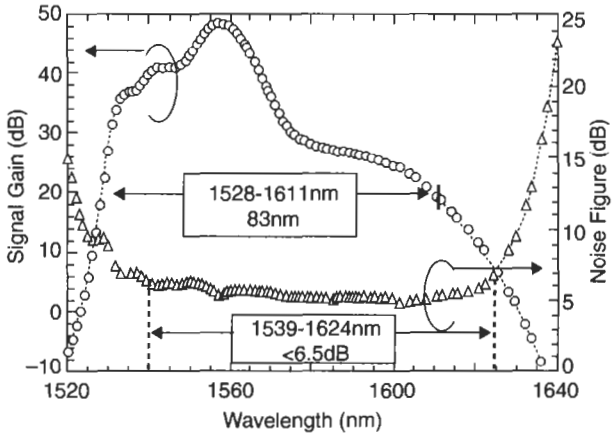


Fig. 31 Tellurite EDFA exhibiting 83 nm bandwidth with gain >20 dB.[220]

Goforth et al.[223] showed the potential for an antimony-silicate-based super-band amplifier. Within a 55 nm band at medium inversion, performance similar to tellurite-based amplifiers was reported.

S-BAND AMPLIFICATION

The S-band represents a wavelength range immediately to the blue side of the low-loss window supported by EDFAs. Once the available bandwidth covered by C- and L-band EDFA systems is utilized this window is recognized as being the most likely band for expansion. Whereas gain-shifted erbium-doped fiber amplifiers are being studied for the longer wavelengths (1570–1620 nm), thulium-doped fiber amplifiers are among the technologies

under consideration for amplification in the shorter wavelength range (1450–1520 nm). This host requires a low phonon energy environment for the active ion if efficient amplification is to be achieved. Therefore a material system other than the workhorse aluminosilicate is absolutely necessary.

This section will review early work on Tm-doped fluorozirconate fiber amplifiers and report on recent efforts to gain-shift fluorozirconate amplifiers to a region the 1480–1520 nm bands rather than the stronger 1450–1490 nm wavelength range in which the early work concentrated. In addition, we will discuss operation in an antimony-silicate host where the Tm ions reside in a local low phonon environment.

Early Fluorozirconate-based Tm Amplifiers

Tm amplifier research has occurred in two waves. Much excellent work was done in the early 1990s shortly after the breakthrough work on silica-based EDFAs. There was little work in this area for several years until it was revived in 1999 by researchers at NTT and NEC in Japan and in the last year by Alcatel and IPG in Europe and Corning in the USA. The reason for the revival is a technological bandwidth exhaust as the limits of the EDFA are reached. Before touching on the recent work it is appropriate to review the pioneering efforts from a decade ago.

The energy level diagram of the Tm ion is shown in Fig. 32. The transition of interest is $^3F_4 \rightarrow {}^3H_4$ in the gain band 1450 to 1520 nm. The spectroscopy is very different from erbium. Notably the transition does not terminate in a manifold of the ground state but on a higher energy level which itself has

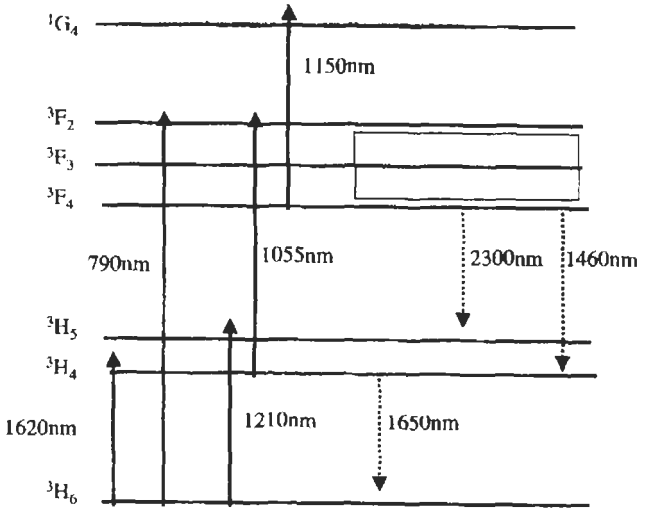


Fig. 32 Principal energy levels of trivalent thulium.

a long metastable lifetime. In addition there are competing transitions to the $^3H_5$ and $^3H_6$ levels. The first competing transition dictates the need for a low-phonon energy host, whereas the latter dictates finding a pumping scheme that maintains a high ground state population. The branching ratios for the 800, 1460, and 2300 nm transitions are approximately 0.89, 0.08, and 0.03. In addition to the branching ratio problem, the long lifetime of the $^3H_4$ level makes the transition self-terminating under normal pumping conditions. Therefore, some means is necessary to depopulate this lower level as well as populating the upper level. Several options were successfully pursued by the early researchers. The host glass of choice has been the fluorozirconate ZBLAN. Here, the lifetimes are approximately 1 ms for the $^3F_4$ level and 10 ms for the $^3H_4$ level.

The most popular and simplest means of excitation was the 1060 nm up-conversion pumping scheme. In this case the 1060 nm wavelength has very weak ground state absorption but a strong absorption from the $^3H_4$ to $^3F_2$ level. Thus, both problems associated with the transition are overcome. The weak ground state absorption ensures that only a small fraction of the total ions are excited, while the strong intermediate absorption quickly recycles the participating ions to the upper level. In addition, the 1060 nm wavelength is readily available from mini-YAG or Nd-doped fiber lasers at high power.

Percival and Williams[224] demonstrated peak gain in excess of 18 dB over a 35 nm bandwidth ranging from 1453 to 1488 nm for a launched pump power of 155 mW. A small signal gain coefficient of 0.25 dB/mW was achieved in a fiber with a numerical aperture of 0.4. Similar results were obtained by Komukai et al.[225] and Rosenblatt et al.[226]

Another pumping scheme involved direct excitation of the $^3F_4$ manifold by a laser diode at 795 nm. This scheme was considered important at the time because it facilitated diode pumping. However two problems remained, ASE at 800 nm and the self-termination problem. One solution was codoping with holmium as reported by Sakamoto et al.[227] This solved the self-termination problem by energy transfer from the $^4H_4$ level of Tm^{3+} to the $^5I_7$ of holmium, but the amplifier still self-saturated by 800 nm ASE at a gain of 18 dB. By introducing distributed filtering at 800 nm the viability of this scheme could be further explored.

A second variant involved colasing of the $^3F_4 \rightarrow {}^3H_6$ transition.[228] This is overall more effective because there is greater flexibility in the choice of pump wavelength and the colasing recycles the power via the ground state of Tm^{3+}.

Gain-shifted Pumping Schemes

After a lull of a few years with little work on thulium amplifiers, a resurgence of interest occurred in the late 1990s. An S-band amplifier was the logical next step in bandwidth extension following a C + L split band EDFA. Interest was particularly strong in Japan, where large quantities of dispersion shifted fiber have been installed. It is well known that such fibers do not support DWDM

in the C-band. A natural upgrade path would be to pair an L-band with an S-band.

Unfortunately, the gain peak obtained from a Tm-doped fluorozirconate fiber amplifier is near 1480 nm, on the blue side of the low-loss window. This has led to the concept of the gain-shifted Tm-doped fiber amplifier operating in the 1480–1520 nm range, i.e., close to the C-band but short of the dispersion zero at 1550 nm.

The first demonstration of a gain-shifted TDFA was reported by Kasamatsu et al.[229] They employed a dual wavelength pumping scheme supplementing the 1060 nm up-conversion pumping technique with a 1560 nm control pump. The supplementary pump had a number of effects.

By adjusting the pump power of 1560 nm relative to 1060 nm the population of the $^3H_4$ level could be controlled. This had the effect of reducing the inversion of $^3F_4$ level relative to the $^3H_4$ level, which in turn shifted the gain spectra to the tails of the transitions much like in an L-band EDFA. Although the fractional inversion of level 3 relative to level 2 is reduced by the 1560 nm pump to achieve the gain shaping, the absolute inversion can be greatly enhanced. The 1560 nm pump depopulates the ground state more efficiently than the 1060 nm pump, thus increasing the gain per unit length by making more Tm^{3+} ions available for amplification. Finally, by partially bleaching the ground state absorption, the ground state absorption of the longer wavelength signals are reduced, thus improving the noise figure in the tails of the gain spectrum.

The configuration of the gain-shifted TDFA is shown in Fig. 33. The first stage was pumped by 330 mW from an Yb fiber laser operating at 1047 nm and by a 42 mW EDFA operating at 1555 nm. The second and third stages were pumped by 29 mW at 1547 nm and 395 mW at 1053 nm. The output characteristics of the amplifier is shown in Fig. 34, where they are also compared to a conventional up-conversion pumped TDFA. A gain-shifted operation with small signal gain of >25 dB and noise figure of approximately 5 dB in the 1475–1510 nm band was obtained.

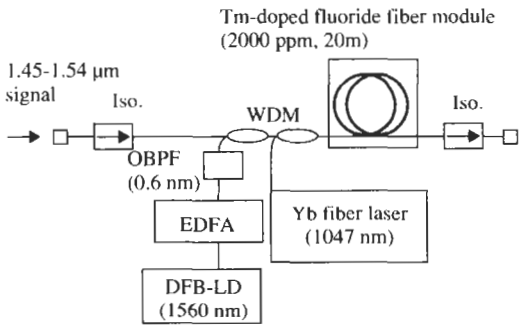


Fig. 33 Schematic diagram of the first reported gain-shifted thulium amplifier.[229]

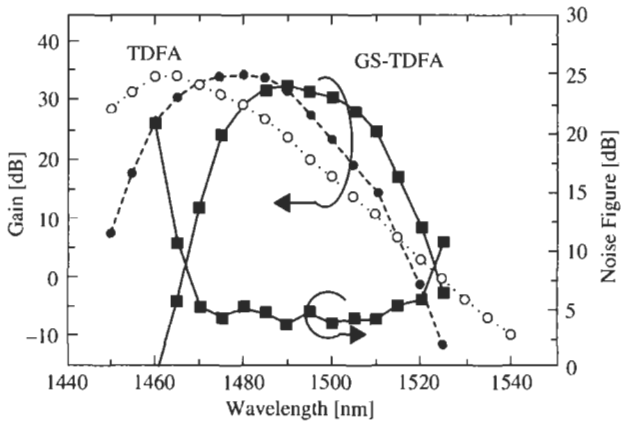


Fig. 34 Comparison between conventional up-conversion-pumped TDFA and gain-shifted TDFA with supplementary pump at 1560 nm.[229]

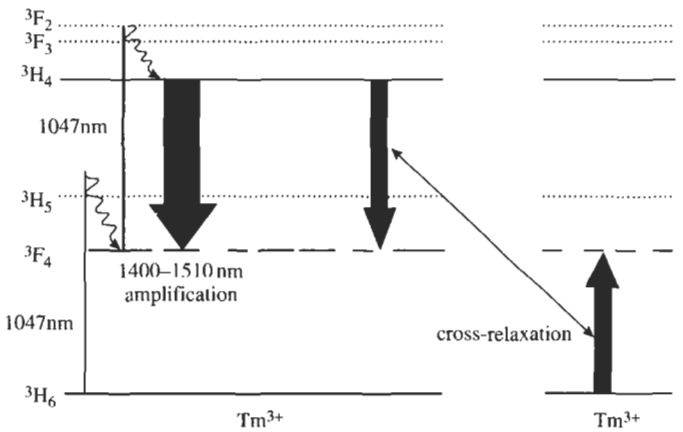


Fig. 35 Illustration of how a cross-relaxation mechanism can seed the population of the $^3F_4$ level.[230]

Another way of achieving a low-inversion amplifier was proposed by Aozasa et al.[230,231] Instead of the auxiliary pump at 1555 nm, they simply used a high concentration fiber and the cross-relaxation process illustrated in Fig. 35. The cross-relaxation effectively seeded the intermediate population in a cascading manner. The concentration dependence of gain spectra is shown

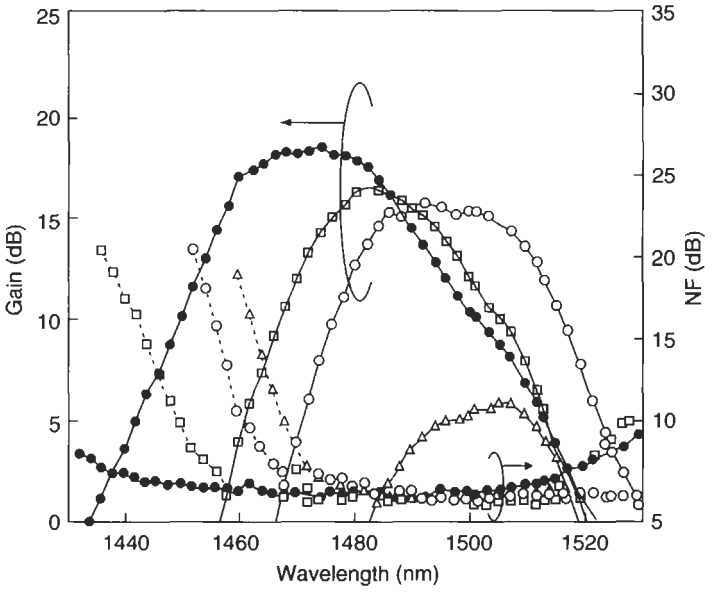


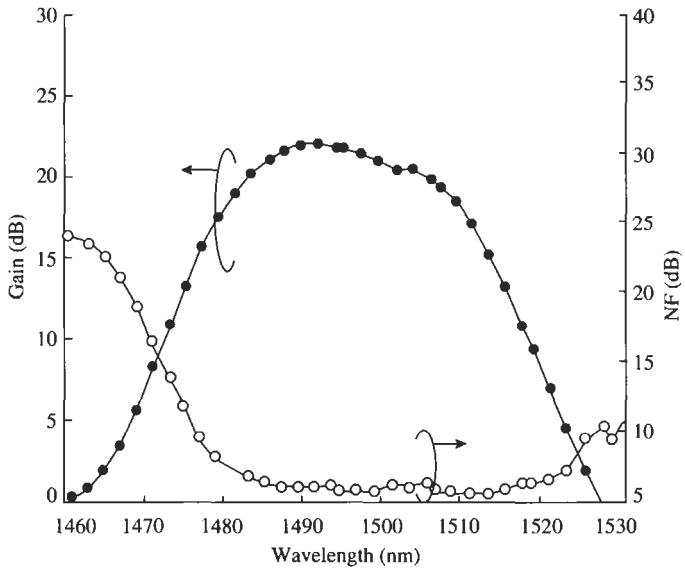
Fig. 36 Dependence of TDFA characteristics on concentration. Optimum concentration for an efficient gain-shifted amplifier is around 6000 ppm.[230]

in Fig. 36, indicating an optimum concentration of around 6000 ppm. In a two-stage amplifier pumped with a total of 1 W at 1047 nm, a small signal gain >30 dB was obtained (Fig. 37). The total output power was a respectable +15 dBm.

Tm-doped Antimony-Silicate Amplifier

The antimony-silicate material developed initially for wide-band EDFAs also performs well as a host for Tm^{3+}. Figure 38 shows a comparison of the fluorescence bandwidths of a typical thulium-doped fluorozirconate fiber (ZBLAN) and in the Tm-doped multicomponent silicate fiber studied here. The FWHM for the two fluorescence bands are 80 and 100 nm, respectively, with most of the broadening occurring on the long wavelength side of the fluorescence peak at 1460 nm. This change in line shape is very important for extending the region of efficient amplification beyond 1500 nm, something that is very difficult to achieve in thulium-doped ZBLAN fibers. The broader bandwidth should enable extending the gain to longer wavelengths without sacrificing efficiency.

As has been stated before, one of the most critical parameters for efficient amplification around 1460 nm in thulium-doped fibers is the fluorescence lifetime for the $^3H_4$ and $^3F_4$ energy levels. As noted above, typical values for these lifetimes in Tm-doped ZBLAN are 1.5 ms and 10 ms respectively. However, in most silicate glasses one would expect considerably shorter lifetimes, in particular for the $^3H_4$ level, which is very sensitive to the glass phonon energy.



Total input signal power = 7 dBm, forward pump power = 325 mW, backward pump power = 175 mW, Fibre length = 13.3 m, Tm^{31} concentration = 6000ppm

Fig. 37 Gain and noise figure at 6000 ppm with 500 mW total pump power.[230]

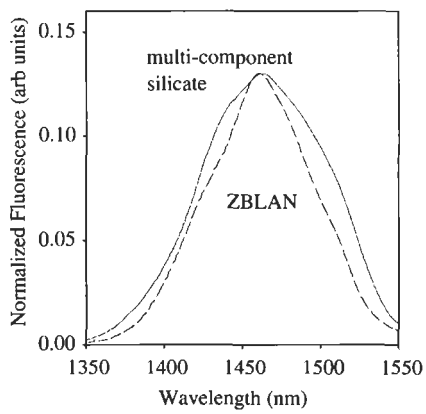


Fig. 38 Fluorescence spectra of antimony-silicate host compared to ZBLAN.

Surprisingly, Samson et al.[232] measured single exponential decay time around 450 μs in multicomponent silicate glasses. This was attributed to a local "low phonon" energy environment favored by rare earth ions in these particular silicate compositions. Indeed, this lifetime is comparable to that measured in germanate and tellurite glasses, glass families considered to have lower phonon energy than silicates.

The pumping scheme for the Sb/Si TDFA is shown in Fig. 39. It is a variant of other up-conversion pumping schemes, utilizing a 1560 nm pump for initial excitation from the ground state, followed by 1405 nm pumping to promote the Tm ions to the metastable level of interest. The amplifier configuration is shown in Fig. 40. Single-stripe laser diodes were employed in all cases. The 9 m-long silicate fiber had a thulium concentration of 0.1 mol% Tm_2O_3 and a numerical aperture of 0.35. The core diameter was 4 μm. The signal wavelength was varied from 1460 nm to 1520 nm through the use of a tunable diode laser (Santec). The signal power was −23 dBm. Light at 1560 nm, produced by tunable laser diode amplified by an EDFA, was counterpropagated with the 1405 nm pump to investigate the effects of dual-wavelength pumping. The small signal gain spectrum for pumping with a fixed 1405 nm power level and different amounts of counter-propagating 1560 nm pump is shown in Fig. 41. The spectrum for 1405 nm pumping alone is also given for comparison. The small signal gain efficiency at a signal wavelength of 1490 nm for variable 1405 nm pumping again at different 1560 nm pump powers is displayed in Fig. 42.

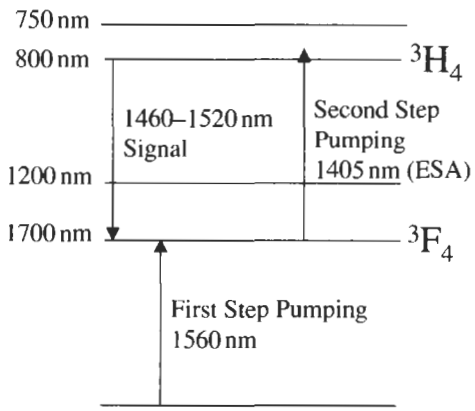


Fig. 39 Dual wavelength pumping scheme using in-band upper-level pump wavelength.[232]

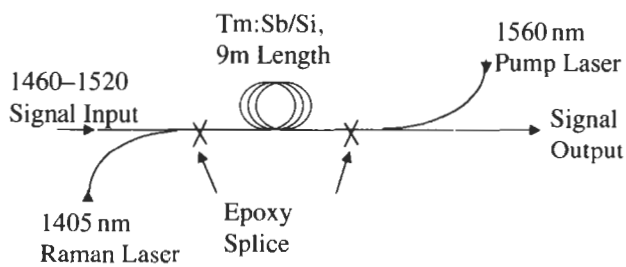


Fig. 40 Schematic diagram of 1405/1560 nm dual-wavelength-pumped TDFA.[232]

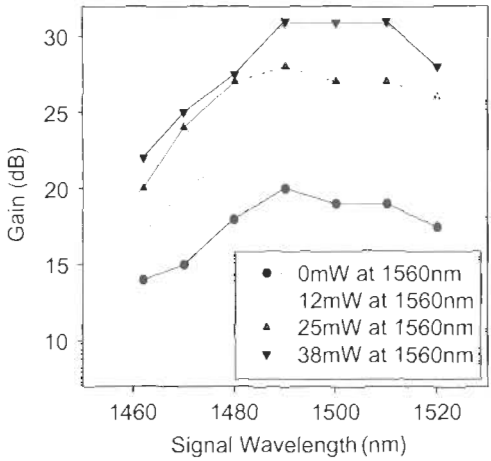


Fig. 41 Gain spectra of antimony silicate amplifier vs 1560 nm pump power.[232]

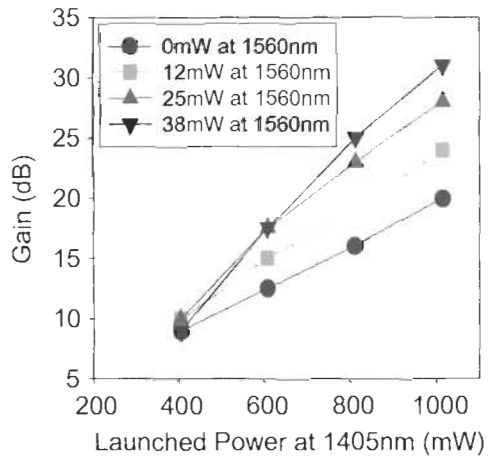


Fig. 42 Gain efficiency for 1405 nm pump for various levels of 1560 nm supplementary power.[232]

Figure 42 indicates that a small amount of 1560 nm pump power (<50 mW) is shown to enhance the gain by 11 dB. This is due to the 1560 nm light populating the first excited state ($^3F_4$ level), from which the 1405 nm pump is efficiently absorbed (1405 nm does not correspond with any strong ground state absorption but does have significant ESA). The gain slope efficiency is improved by more than a factor of two in the presence of the counterpropagating 1560 nm pump. Although the gain coefficient (~0.03 dB/mW) is lower than the figure obtained in Tm-doped ZBLAN fibers, partly as a result of the shorter lifetime inherent in these silicate glasses, this figure can be easily

doubled by reducing the core diameter or increasing the metastable lifetime by composition adjustments. Subsequent measurements on the current fiber yielded slope efficiencies in excess of 0.05 dB/mW through optimization of the fiber length. Increased performance is also expected from reductions in the fiber background loss. It is worth pointing out that the current figure is superior to that obtained from Raman fiber amplifiers operating in the same wavelength regime. Gain exceeding 20 dB has been demonstrated over a 60 nm range with a peak gain of 31 dB, a result made all the more significant since it was obtained in a multicomponent silicate fiber. Power conversion efficiencies > 40% are anticipated for this fiber. A similar pumping scheme was also reported for a Tm-doped ZBLAN fiber amplifier.[233]

Power Scaling of Three-Level Transitions Enabled by High Numerical Aperture Fibers

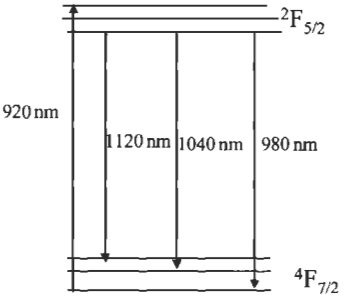
There are occasions when the choice of material is not in itself of particular importance but rather the ease by which the material can be processed that governs the choice. As we have shown in earlier sections, the Al/Ge/Si host is certainly not the best for broadband EDFAs in either C- or L-band. It has become the industry standard simply because fibers can be made by a simple extension of the chemical vapor deposition process. In this section we discuss the problem of power scaling the three-level 980 nm transition of Yb. Although viable in virtually any glass host when pumping with high-power low-brightness broad-area lasers there are clear advantages to going to a nonsilica host. It is the ability to fabricate very high numerical aperture fibers from multicomponent glass that is the key enabling factor in this case.

The 980 nm Transition of Yb

We have seen in earlier sections that the demand for optical bandwidth is driving significant activity to develop broadband erbium-doped fiber amplifiers. As the number of optical channels increases, there is a corresponding need for increased pump power. Amplifiers are conventionally pumped by single-stripe laser diodes, but power scaling of these devices has not kept up with the demand for EDFA pump power. Thus, WDM amplifiers are typically pumped by up to six individual pumps, adding to increased complexity of amplifier design. While semiconductor based master oscillator-power amplifier (MOPA)[234] or flared designs[235] with > 1 W of output power have been reported in the literature, reliability problems have to date hindered their deployment as fiber amplifier pumps. Therefore, attempts to power scale the pump power available for EDFAs have concentrated on using either pump multiplexing techniques,[236] or cascaded Raman oscillators,[237] themselves pumped by double-clad fibers operating in the 1.06–1.12 μm region. While these approaches are certainly viable, they involve extra complexity and have little cost benefit.

High Power 978 nm Fiber Laser

Trivalent ytterbium has the simplest electronic configuration of any rare earth element ion. It is essentially a two-level laser system with a very large stark splittings in the ground and excited state. This is illustrated in Fig. 43. The corresponding absorption and emission spectra are shown in Fig. 44. Yb ions exhibit gain in a narrow (6 nm), band centered at 978 nm and on a broader



Wavelength	Relative Gain	Inversion Required
980nm/3level	1	>50%
1040/quasi 3 level	0.25	>5%
1120/4 level	0.01	>0

Fig. 43 Energy level diagram and principal transitions of trivalent Ytterbium.

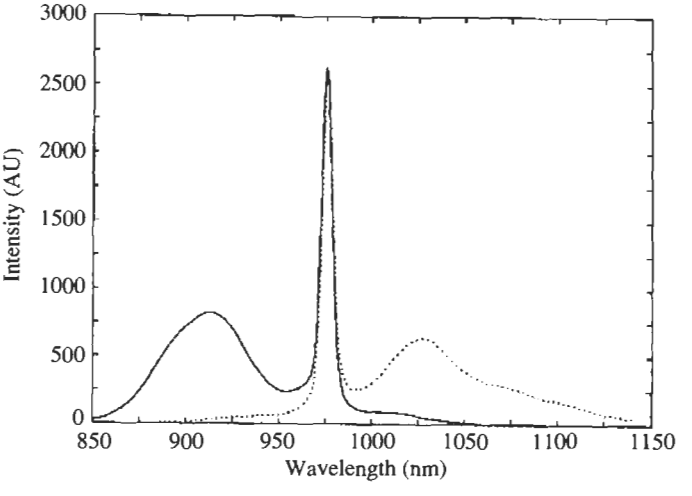


Fig. 44 Typical emission and absorption spectra of Yb doped fiber.

band with a peak at 1030 nm but extending as far as 1140 nm. The 978 nm transition behaves like a classic three-level laser transition while the extreme red wavelength is essentially 4-level. This is a consequence of the large stark splitting in the ground state. The latter transition requires only about 5% inversion for transparency, while the former requires at least 50%.

For many years a source based on the 978 nm transition has been suggested as a pump source for high-power EDFAs. However researchers found it very difficult to find an efficient pumping scheme for power scaling beyond that of a single-stripe diode. The problem lies in the relationship between wavelength dependent the gain and the pump absorption, as clarified below. For example, the gains at 978 nm and 1030 nm, assuming homogeneous broadening, are related by:

$$G_{1030} = 0.25G_{978} + 0.74\alpha_p(\Gamma_s/\Gamma_p) \tag{3}$$

where Γ_s and Γ_p are the overlap factors of the signal pump and pump modes with the dopant profile and α_p is the partially bleached pump absorption. The above relationship is based on published cross-section data for Al/Ge/Si glass, but a similar relation will hold for other hosts.

In a conventional fiber laser where a single-mode semiconductor laser is employed as a pump source, the overlap ratio is close to unity. Therefore the stronger 3-level transition can be selected simply by optimizing the length of the device. Hanna et al.[238] demonstrated this in 1989 for the 978 nm transition of Yb. This work actually preceded the emergence of semiconductor diodes at this wavelength, but once such diodes became available, there was little need for a fiber laser version. Dense WDM has resurrected interest in this transition provided it can be scaled to higher powers in the 500–1000 mW regime.

The usual method for power scaling fiber lasers is the brightness-converting scheme known as cladding pumping,[239] illustrated schematically in Fig. 45. The pump source is typically one or more broad stripe diode lasers. These are high-power multimode diodes that are incompatible with direct coupling into single-mode fibers. They can be used to pump single-mode fiber lasers, however, through use of specially designed double-clad fibers. These fibers typically have a conventional single-mode core doped with rare earth ions surrounded by a large multimode pump waveguide with a diameter of typically 100 μm.

The effect of such a structure is to reduce the pump absorption coefficient in proportion to the cladding waveguide area. In a cladding pumped device the overlap factor for the pump is typically two orders of magnitude less than for the lasing signal. Therefore equation (3) suggests that in the specific case of a double-clad YDFL, high pump absorption will strongly favor gain at 1030 nm over 978 nm. In fact the free-running wavelength of a cladding pumped Yb laser is usually extremely redshifted to 1120 nm.

For a typical single-mode core pumped fiber laser cavity with a round trip loss of 14 dB (one high reflector, one 4% reflecting cleave) a pump absorption of between 6 and 7 dB is possible. However, in a typical cladding pumping

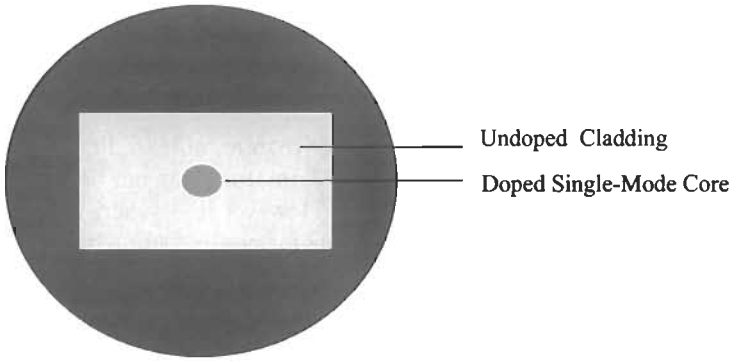


Fig. 45 Double-clad fiber geometry for cladding pumping with broadstripe lasers.

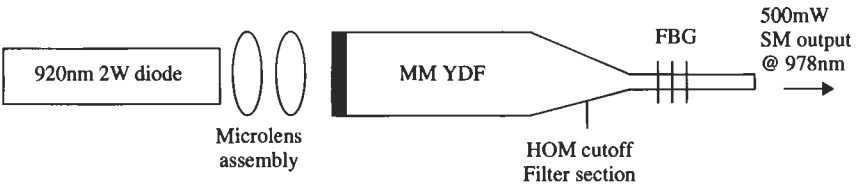


Fig. 46 Schematic diagram of tapered multimode fiber laser oscillator.[240]

structure where the area ratio approaches 100, only a very small percentage of pump light can be absorbed when selecting the 980 nm line. While the use of discriminating dielectric mirrors or fiber gratings can help a little, ASE at the longer wavelength prevents reaching the desired inversion.

We shall return to cladding pumping of this transition, but first we describe the first successful demonstration of high power in this wavelength range. In this case cladding pumping was not used at all. Instead, the ability to make very high numerical aperture fibers was exploited.

The Tapered Multimode Oscillator

Utilizing high-power, low-cost, broad stripe diodes as pump sources for 978 nm generation, the concept illustrated in Fig. 46 (first reported by Minelly et al.[240]) has been exploited. The laser cavity comprises an Yb-doped multimode fiber with a taper-based, mode-selective filter at the fiber end downstream of the pump input. The multimode fiber allows for efficient coupling of the diode laser output into the fiber, while the taper ensures that only the lowest-order mode has round-trip feedback. The core pumping scheme avoids the problem of 1030 nm ASE self-saturation just as in the single-mode case described above. Of course, the fiber laser now has a higher threshold than in the case of a

single mode fiber, but there is no threshold penalty in this technique relative to cladding pumping. The fiber is designed so that

- the threshold power is a small fraction of the available pump power;
- the launch efficiency from a broad stripe diode is >80% with conventional imaging optics;
- 978 nm oscillation can be achieved for minimal pump leakage.

The threshold power of a three-level laser is closely related to the pump power required to bleach a given length of fiber. It is well approximated by:

$$P_t = (h\nu/\sigma_a\tau) \cdot \pi a^2 \times (\alpha_p/4.343) \tag{4}$$

where σ_a is the pump absorption cross-section, τ is the fluorescent lifetime, πa^2 is the area of the core, and α_p is the pump absorption in dB. Clearly, the threshold power will scale with area of the pump waveguide. It is therefore desirable to minimize the area of the cladding waveguide. Unfortunately, this tends to limit the launch efficiency from the diode to fiber. A critical factor in obtaining the best compromise is the numerical aperture of the fiber. Figure 47 illustrates modeled threshold power as a function of numerical aperture for various fiber designs. The assumption is that the fiber etendue (product of N.A. and dimension) is greater than that of the slow axis of the laser diode. It is clear that a circular geometry would only be practical if N.A. close to unity can be achieved. Although this could in principle be possible with compound glasses, there do not exist practical lenses with such N.A. for demagnifying the diode output.

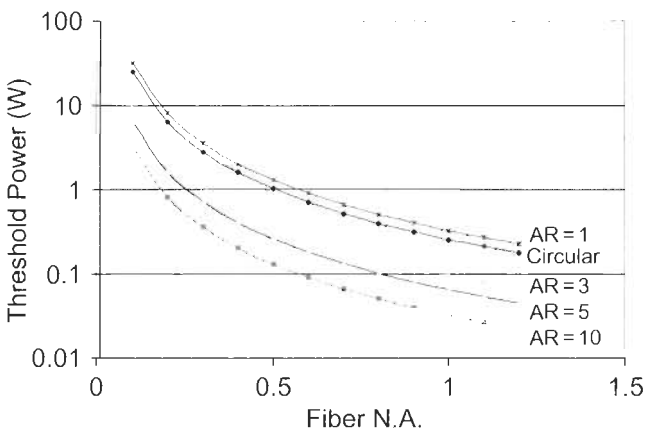


Fig. 47 Theoretical curves illustrating the role of numerical aperture and rectangular geometry in reducing threshold power of YDFL.

This is not necessary, however; lenses with N.A. up to 0.6 are routinely available, and there is in any case a great deal of spot size redundancy in the diffraction limited image plane. By going to a rectangular or elliptical geometry the threshold can be brought down to acceptable levels.

The dimension of the major axis dimension is dictated by geometrical imaging of the broad-stripe diode facet, while that of the minor axis is dictated by lens-induced aberrations. An N.A. of 0.45 was obtained by a special technique in which a multicomponent La/Al/Si glass is clad with pure silica. This allows for a 3 : 1 demagnification of the long axis of the diode facet. The chosen rectangular geometry minimizes threshold power without compromising launch efficiency. The dimensions of the multimode core were $30 \times 10\,\mu m$, and the Yb concentration was 1 wt%.

The laser cavity itself comprised 12 cm of multimode Yb fiber terminated with a taper-based mode-transformer. The taper adiabatically transformed the lowest-order mode of the rectangular structure into a close match to the LP_{01} mode of Corning® CS980 fiber. The cavity was completed by a HT920 nm, HR980 nm dielectric filter at the multimode pump end and a 5% reflectivity, 1 nm broad fiber grating in the output CS980 pigtail. A chirped 920 nm high reflector filters out the small amount of remnant pump which coupled into the single-mode CS980 fiber.

Pump light from a 100 μm broad stripe laser was launched into the Yb fiber laser via a microlens assembly with an efficiency of approximately 80%. The output characteristics of the laser is shown in Fig. 48, indicating a threshold of 350 mW, a slope efficiency of 35%, and a maximum output power of 450 mW. The mode transformer reduces the output power by less than 1.5 dB relative to the fiber laser running multimode. This indicates near optimum mode matching to CS980 fiber and very low mode coupling in the multimode section of the oscillator. The wavelength of 978 nm is ideal for EDFA pumping.

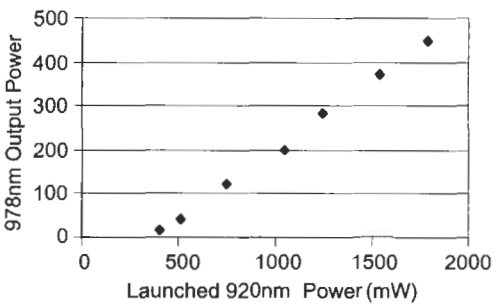


Fig. 48 Output characteristics of tapered multimode YDFL.[240]

Optimized Double-clad Fiber for 980 nm

The tapered oscillator approach achieved the desired result by a fiber design that maximized the laser engineers ability to select the three-level wavelength. However, the multimode fiber itself had no modal discrimination whatsoever, hence the need for the adiabatic taper. This led to some minor stability problems related to amplification of a small amount of fundamental mode power that couples to higher order modes on the way up the taper.

Double-clad fibers can be made to oscillate at 978 nm but with silica materials to date either the pump leakage or threshold has been too high.[241,242] Limiting the area ratio of cladding to core dimensions can greatly improve matters, but when the ratio gets too small, higher-order modes see gain, and the device starts to run in multiple transverse modes. By careful optimization of the ratio, sufficient modal discrimination can be built into the fiber while maintaining the ability to discriminate against 1030 nm radiation. The design constraints are, however, very strict.

Zenteno et al.[243] described an efficient low-threshold double-clad 980 nm Yb fiber laser based on antimony-silicate glass. The step-index core has a N.A. of 0.1 relative to the inner cladding. The latter, in turn, had an effective N.A. of 0.5 relative to the borosilicate glass outer cladding. The core was uniformly doped with 0.45 wt% Yb having a fluorescence decay time of 875 μs. The fiber was drawn by the triple-crucible method. The fiber has an OD of 125 μm, an ellipsoidal inner cladding cross-section with major and minor axis dimensions of 32 × 16 μm^2, and a core diameter $D \approx 11\,\mu m$ (Fig. 49). This fiber has a broad absorption shoulder from 910 to 950 nm, and the quasi-four-level fluorescence emission peaks at 1015 nm (Fig. 50).

The shape and dimensions of the fiber inner cladding were chosen from considerations involving maximization of pump power coupling efficiency, using anamorphic optics without compromising brightness. A 200 × 1 μm^2

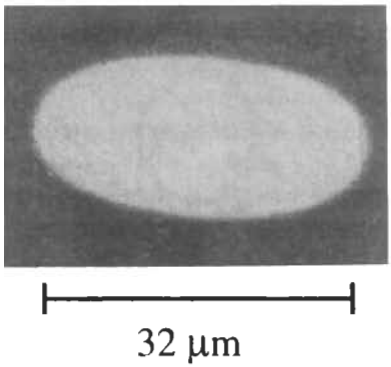


Fig. 49 Photograph of double-clad antimony-silicate Yb fiber laser.[243]

broad-area laser diode with N.A. of 0.1/0.65 in planes parallel/perpendicular to the junction was transformed to a nominally $30 \times 10\,\mu m^2$ spot, and coupled with 75% efficiency into the 0.5 N.A., $32 \times 16\,\mu m^2$ IC. This design preserves high pumping brightness, leading to >50% inversion and fiber laser threshold of only 330 mW.

However, as a result of its small area, the IC higher-order modes (HOM) experience significant gain because of their overlap with the Yb-doped core. This can lead to multimode oscillation, hindering diffraction-limited performance. Figure 51 illustrates this point, where the fundamental and the highest-gain, HOM of a $32 \times 16\,\mu m^2$ IC is computed for core diameter $D = 16\,\mu m$ (a) and $D = 7\,\mu m$ (b). The overlap factor of the fundamental mode with the doped area, which determine the effective modal gain, is $\Gamma_0 = 97\%$ and $\Gamma_0 = 65\%$, respectively. In contrast, the respective HOM has $\Gamma_1 = 96\%$ and $\Gamma_1 = 30\%$. Clearly, as the core size increases, the Γ factor of

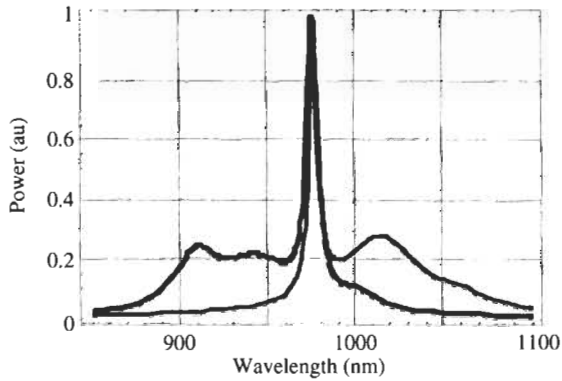


Fig. 50 Absorption and emission spectroscopy of Yb in antimony-silicate glass.[243]

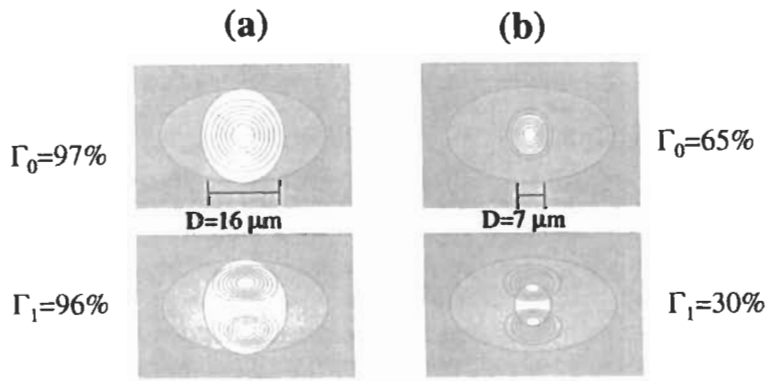


Fig. 51 LP_{01} and LP_{11} modal fields and dopant overlap factor for 16 and 7 μm core diameters within $32 \times 16\,\mu m$ cladding.[243]

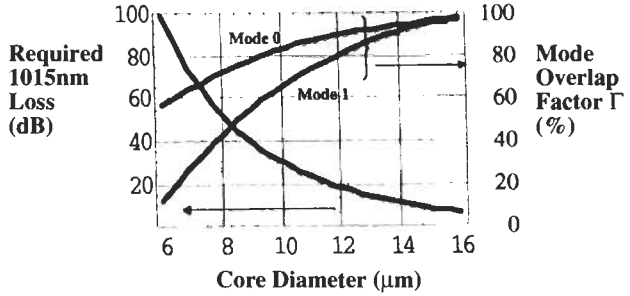


Fig. 52 Required wavelength discrimination against 1015 nm oscillation with a pump absorption of 10 dB.[243]

HOMs increases, approaching the value for the fundamental mode, leading to reduced differential modal gain discrimination.

HOM oscillation is avoided by making the core diameter small enough. Unfortunately, this decreases the pump beam overlap with the Yb-doped area, favoring 1015 nm quasi-four-level oscillation, as discussed in the introduction. To reduce 1015 nm gain to a value below the 978 nm laser threshold gain, a thin-film multilayer dielectric mirror and fiber length control was used. The required 1015 nm loss increases with increasing pump absorption and with decreasing core diameter. Through the well-known relationship linking the gain at three wavelengths in 3-level amplifiers,[244] and making use of the measured absorption and emission cross sections, one can estimate the required 1015 nm loss as a function of core diameter. This is illustrated in Fig. 52, assuming 10 dB of double-pass pump absorption at 915 nm, 13 dB of gain at 978 nm, and fixed inner core area. One can see that the required double-pass loss at 1015 nm is about 80 dB for $D = 7\,\mu m$ and about 5 dB for $D = 16\,\mu m$.

Based on the above considerations, a fiber with $D \approx 11\,\mu m$ was made, representing an optimum balance between achievable spectral loss at 1015 nm using thin films (about 25–30 dB) and sufficient modal gain discrimination (~3 dB) to favor oscillation of only the fundamental mode. Notice that the fiber core V-number is 3.5.

As shown in Fig. 53, over 1 W of output power was obtained with 2.5 W launched. About 10 dB of pump absorption over an uncoiled fiber length of 30 cm was achieved in a double-pass configuration. The input fiber facet M1 was polished at <0.5° with respect to the normal to the fiber axis and had 5% Fresnel reflectivity. This same fiber facet was used to extract the 978 nm laser output via an external dichroic filter D, as shown in Fig. 53. The other fiber facet had a thin film M2 with high reflectivity (>95%) both at 915 and 978 nm and high transmission (>95%) from 1010 to 1050 nm. With respect to launched pump power, threshold was about 330 mW and slope efficiency was 48%. The measured far field of the fiber laser output had a Gaussian profile with full-width $1/e^2 \approx 7°$ as shown in the inset of Fig. 53a, which agrees well with the

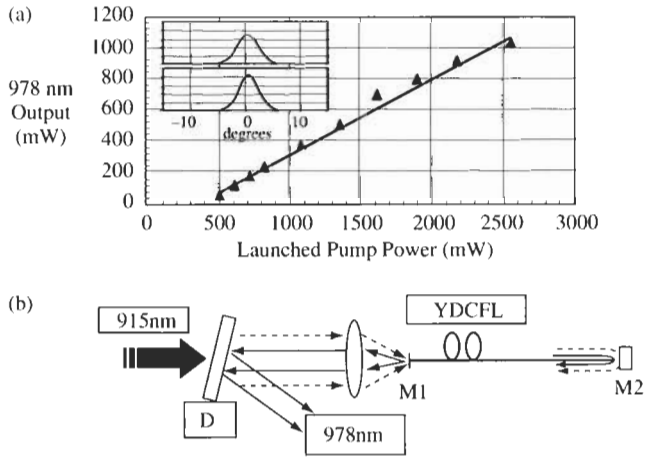


Fig. 53 Output charcteristics and cavity schematic for double-clad antimony silicate YDFL.[244]

computed fundamental MFD of 10 μm. In separate experiments $M^2 < 1.2$ was measured, and the output was coupled to single-mode CS980 fiber with >80% efficiency, further evidence of the diffraction-limited emission of this device.

Pump sources based on the tapered oscillator or optimized double-clad fiber laser are likely enabling technologies for cost reduction in amplifiers.

Summary

In this chapter we have reviewed the main research in glass chemistry as applicable to fiber amplifiers and lasers. Device applications of current interest for dense WDM systems were also discussed. These applications relate principally to bandwidth extension and power scaling.

The role of water and rare earth element clustering was discussed as it pertains to device performance. These problems and the very basic issues surrounding high-purity melting and fiber fabrication were discussed in depth for fluorozirconates. It was shown that the technologies invented to handle fluorozirconates proved useful in the early-stage development of new amplifier materials, particularly tellurites, but antimony silicates as well. Fluoride-based erbium fiber amplifiers have largely been abandoned, but Tm-doped ZBLAN remains a very promising material for S-band amplification. Tellurite and antimony-silicate materials have spectroscopic properties advantageous for broadband EDFAs and possibly for efficient S-band amplifiers as well, though the high optical nonlinearity of tellurites remains a major issue for practical applications. Antimony-silicate glasses have facilitated production of high numerical aperture fibers, enabling high-power cladding-pumped fiber lasers

operating on otherwise difficult three-level transitions. The excellent physical properties of fibers based on Er- and Al-codoped silica and the infrastructure for making fiber were discussed. This illustrated the extraordinary level of performance against which new materials are judged for use in practical devices.

The ultimate deployment of devices based on these new materials is still uncertain. Part of the reason is a bias in the industry towards silica-based materials. Another reason, however, is the steady improvements in Raman technology in recent years, which provide a silica-based approach to bandwidth extension. Finally, a glut of dark fiber offers systems operators the opportunity to light up a new fiber while reusing the bandwidth offered by conventional EDFAs. However, this spare capacity is unlikely to last for long, and many of the applications discussed in this chapter will ultimately be needed. Thus, the issue is less whether new materials are needed but rather the date on which they are deployed.

References

[1] Humbach O, Fabian H, Grzesik U, Haken U, and Heitmann W. Analysis of OH absorption bands in synthetic silica. J. Non-Cryst. Solids 203:19–26 (1996).

[2] Houde-Walter SN, Peters PM, Stebbins JF, and Zeng Q. Hydroxyl-contents and hydroxyl-related concentration quenching in erbium-doped aluminophosphate, aluminosilicate and fluorosilicate glasses. J. Non-Cryst. Solids 286:118–131 (2001).

[3] Ainslie BJ, Craig-Ryan SP, Davey ST, Armitage JR, Atkins CG, Massicott JF, and Wyatt R. Erbium doped fibres for efficient optical amplifiers. IEEE Proceedings 137, Part J:205–208 (1990).

[4] Kagi N, Oyobe A, and Nakamura K. Efficient optical amplifier using a low-concentration erbium-doped fiber. IEEE Photon Technol. Lett. 2:559–561 (1990).

[5] Federighi M and Di Pasquale F. The effect of pair-induced energy transfer on the performance of silica waveguide amplifiers with high Er^{3+}/Yb^{3+} concentrations. IEEE Photon Technol. Lett. 7:303–305 (1995).

[6] Belov AV. Upconversion fluorescence decay at 980 nm and ESA in silica preforms and fibers heavily doped with erbium. European Conference on Optical Communications, ECOC 97, Publication 448:22–25 (1997).

[7] Bremberg D, Helmfrid S, Jaskorzynska B, Swillo M, Philipsen JL, and Palsodottir B. Observation of energy-distribution-dependent homogeneous upconversion in erbium-doped silica glass fibres. Electron. Lett. 35:1189–1191 (1999).

[8] Mylinski P, Szubert C, Bruce AJ, DiGiovanni DJ, and Palsdottir B. Performance of high-concentration erbium-doped fiber amplifiers. IEEE Photon Technol. Lett. 11:973–975 (1999).

[9] Blixt P, Nilsson J, Carlnas T, and Jaskorzynska B. Concentration-dependent upconversion in Er^{3+}-doped fiber amplifiers: Experiments and modeling. IEEE Trans. Photon Technol. Lett. 3:996–998 (1991).

[10] Delevaque E, Georges T, Monerie M, Lamouler P, and Bayon J-F. Modeling of pair-induced quenching in erbium-doped silicate fibers. IEEE Photon Technol. Lett. 5:73–75 (1993).

[11] Delevaque E, Georges T, Monerie M, Lamouler P, and Bayon J-F. Modeling of pair-induced quenching in erbium-doped silicate fibers. IEEE Photon Technol. Lett. 5:73–75 (1993).

[12] Nilsson J, Jaskorzynska B, and Blixt P. Performance reduction and design modification of erbium-doped fiber amplifiers resulting from pair-induced quenching. IEEE Photon Technol. Lett. 5:1427–1429 (1993).

[13] Di Pasquale F and Federighi M. Modelling of uniform and pair-induced upconversion mechanisms in high-concentration erbium-doped silica waveguides. J. Lightwave Technol. 13:1858–1864 (1995).

[14] Myslinski P, Nguyen D, and Chrostowski J. Effects of concentration on the performance of erbium-doped fiber amplifiers. J. Lightwave Technol. 15:112–120 (1997).

[15] Philpsen JL, Broeng J, Bjarklev A, Helmfrid S, Bremberg D, Jaskorzynska B, and Palsdottir B. Observation of strongly nonquadratic homogeneous upconversion in Er^{3+}-doped silica fibers and reevaluation of the degree of clustering. IEEE J. Quantum Electron. 35:1741–1749 (1999).

[16] Poulain M, Poulain M, and Lucas J. Verres fluores au tetrafluorure de zirconium propriétés optiques d'un verre dopé au Nd3+. Mat. Res. Bull. 10:243–246 (1975).

[17] Lucas J, Chanthanasinh M, Poulain M, and Poulain M. Preparation and optical properties of neodymium fluorozirconate glasses. J. Non-Cryst. Solids 27:273–283 (1978).

[18] Poulain M, Chanthanasinh M, and Lucas J. Noveaux verres fluores. Mat. Res. Bull. 12:151–156 (1977).

[19] Lucas J, Chanthanasinh M, Poulain M, and Poulain M. Preparation and optical properties of neodymium fluorozirconate glasses. J. Non-Cryst. Solids 27:273–283 (1978).

[20] Lecoq A and Poulain M. Lanthanum fluorozirconate glasses. J. Non-Cryst. Solids 34:101–110 (1979).

[21] Lecoq A and Poulain M. Fluoride glasses in the ZrF_4–BaF_2–YF_3–AlF_3 quaternary system. J. Non-Cryst. Solids 41:209–217 (1980).

[22] Poulain M. Glass formation in ionic systems. Nature 293:279–280 (1981).

[23] Lucas J. Review: Fluroide glasses. J. Mat. Sci. 24:1–13 (1989).

[24] Lucas J and Adam J-L. Halide glasses and their optical properties. Glastech Ber. 62:422–439 (1989).

[25] Ohsawa K and Shibata T. Preparation and characterization of ZrF_4–BaF_2–LaF_3–NaF–AlF_3 glass optical fibers. J. Lightwave Technol. 2:602–606 (1984).

[26] Fiedler KH, Levin KH, and Tran DC. Raman spectrum of irradiated zirconium fluoride glass. Phys. Chem. Glasses 23:205–206 (1982).

[27] Mitachi S, Miyashita T, and Manabe T. Preparation of fluoride optical fibres for transmission in the mid-infrared. Phys. Chem. Glasses 23:196–201 (1982).

[28] Tran DC, Sigel GH Jr, and Bendow B. Heavy metal fluoride glasses and fibers: A review. J. Lightwave Technol. 2:566–586 (1984).

[29] Sakaguchi S and Takahashi S. Low-loss fluoride optical fibers for midinfrared optical communication. J. Lightwave Technol. 5:1219–1228 (1987).

[30] Wetenkamp L, West GF, and Tobben H. Optical properties of rare earth-doped ZBLAN glasses. J. Non-Cryst. Solids 140: 35–40 (1992).

[31] Freitas JA Jr, Strom U, and Tran DC. Raman scattering and electron spin resonance studies of fluorozirconate glasses. J. Non-Cryst. Solids 81:303–317 (1986).

[32] Rousset JL, Ferrari M, Duval E, Boukenter A, Mai C, Etienne S, and Adam JL. First stages of crystallization in fluorozirconate glasses. J. Non-Cryst. Solids 111:238–244 (1989).

[33] Gur'ev NV, Chikovsky AN, Kolobkov VP, Leskova ML, Petrovskii GT, and Zhmyreva IA. Glass forming ability of rare earths in fluorozirconate glasses. J. Non-Cryst. Solids 170:155–160 (1994).

[34] Qui J, Maeda K, Konishi A, Terai R, and Kadono K. Influence of various divalent metal fluorides on the properties of glasses in the ZrF_4–AlF_3 system. J. Non-Cryst. Solids 184:109–113 (1995).

[35] Lebullenger R, Benjaballah S, Le Deit C, and Poulain M. Systematic substitutions in ZBLA and ZBLAN glasses. J. Non-Cryst. Solids 161:217–221 (1993).

[36] MacFarlane DR, Newman PJ, Zhiping Zhou, and Javorniczky J. Systematic study of refractive index variations with composition in heavy metal fluoride glasses. J. Non-Cryst. Solids 161:182–187 (1993).

[37] Hasz WC, Whang JH, and Moynihan CT. Comparison of physical properties of ZrF_4- and HfF_4-based melts and glasses. J. Non-Cryst. Solids 161:127–132 (1993).

[38] Poignant H, Le Mellot J, Leve A, Le Cabec P, and Riviere D. Effect of some refractive index modifiers on both numerical aperture and mechanical strength of fluorozirconate fibers. J. Non-Cryst. Solids 161:192–197 (1993).

[38a] Kogo T, Onishi M, Kanamori H, and Yokota H. Characterization of fluoride glasses for single-mode optical fibers with large refractive index difference. J. Non-Cryst. Solids 161: 169–172 (1993).

[39] Mitachi S, Teruna Y, Ohishi Y, and Takahashi S. Reduction of impurities in fluoride glass fibers. J. Lightwave Technol. 2:587–592 (1984).

[40] Robinson M. Processing and purification techniques of heavy metal fluoride glass (HMFG). J. Crystal Growth 75:184–194 (1986).

[41] Poulain M, Lebullenger R, and Saad M. Synthesis of high purity fluorides by wet chemistry. J. Non-Cryst. Solids 140:57–61 (1992).

[42] Sommers JA, Fahey JV, Ewing KJ, and Aggarwal ID. The fate of selected impurities in the recrystallization of zirconium oxychloride and in its conversion to zirconium tetrafluoride. J. Mat. Res. 5:392–396 (1990).

[43] Dai G, Huang J, Cheng J, Zhang C, Dong G, and Wang K. A new preparation route for high purity ZrF_4. J. Non-Cryst. Solids 140:229–232 (1992).

[44] Sanghera JS, Hart P, Sachon MG, Ewing KJ, and Aggarwal I. New fluorination reactions of ammonium bifluoride. J. Am. Ceram. Soc. 73:1339–1346 (1990).

[45] Onishi M, Kohgo T, Amemiya K, Nakazato K, Kanamori H, and Yokota H. Thermal and mass analyses of fluorination process with ammonium bifluoride. J. Non-Cryst. Solids 161:10–13 (1993).

[46] Robinson M, Pastor RC, Turk RR, Devor DP, Braunstein M, and Braunstein R. Infrared-transparent glasses derived from the fluorides of zirconium, thorium, and barium. Mat. Res. Bull. 15:735–742 (1980).

[47] Drexhage MG, Moynihand CT, Bendow B, Gboji E, Chung KH, and Boulos M. Influence of processing conditions on IR edge absorption in fluorohafnate and fluorozirconate glasses. Mat. Res. Bull. 16:943–947 (1981).

[48] Nakai T, Mimura Y, Tokiwa H, and Shinbori O. Dehydration of fluoride glasses by NF_3 processing. J. Lightwave Technol. 4:87–89 (1986).

[49] Mailhot AM, Elyamani A, and Riman RE. Reactive atmosphere synthesis of sol-gel heavy metal fluoride glasses. J. Mat. Res. 7:1534–1540 (1992).

[50] Saad M and Poulain M. Fluoride glass synthesis by sol-gel process. J. Non-Cryst. Solids 184:352–355 (1995).

[51] Ballato J, Riman RE, and Snitzer E. Sol-gel synthesis of fluoride optical materials for planar integrated photonic applications. J. Non-Cryst. Solids 213&214:126–136 (1997).

[52] Zhou Z, Newman PJ, Wong DKY, and MacFarlane DR. Electrochemical purification of fluoride melts. J. Non-Cryst. Solids 140:297–300 (1992).

[53] Murtagh MT, Sigel GH Jr, Fajardo JC, Edwards BC, and Epstein RI. Laser-induced fluorescent cooling of rare-earth-doped fluoride glasses. J. Non-Cryst. Solids 253:50–57 (1999).

[54] Baniel P and Belouet C. Gas film levitation: a unique containerless technique for the preparation of fluoride glass rods. J. Non-Cryst. Solids 161:1–6 (1993).

[55] Fujiura K, Nishida Y, Sato H, Sugawara S, Kobayashi K, Terunuma Y, and Takahashi S. Plasma-enhanced chemical vapor deposition of ZrF_4-based fluoride glasses. J. Non-Cryst. Solids 161:14–17 (1993).

[56] Sakaguchi S. Evaluation of heating condition for collapse of fluoride glass tubes. Electron. Comm. Jpn. Part 2, 76:39–45 (1993).

[57] Hasz WC and Moynihan CT. Physical properties of ZrF_4-based glass-forming melts and glasses. J. Non-Cryst. Solids 140:285–292 (1992).

[58] Braglia M, Bruschi C, Cavalli D, Cocito G, Guojun D, Kraus J, and Mosso S. Rheology of fluoride glasses. J. Non-Cryst. Solids 213&214:325–329 (1997).

[59] Zhang G, Jiang J, Poulain M, Delben AS, and Delben JR. Viscosity of fluoride glasses near the fiber drawing temperature region. J. Non-Cryst. Solids 256&257:135–142 (1999).

[60] Hirai S, Ohnishi M, Chigusa Y, Kyoto M, Watanabe M, and Tanaka G. The study of drawing conditions for fluoride glass fiber. Mat. Sci. Forum 67&68:323–328 (1991).

[61] Tran DC, Ginther RJ, and Sigel GH Jr. Fluorozirconate glasses with improved viscosity for fiber drawing. Mat. Res. Bull. 17:1177–1184 (1982).

[62] Mimura Y, Tokiwa H, and Shinbori O. Fabrication of fluoride glass fibres by the improved crucible technique. Electron. Lett. 20:100–101 (1984).

[63] Tokiwa H, Mimura Y, Nakai T, and Shinbori O. Fabrication of long single-mode and multimode fluoride glass fibres by the double-crucible technique. Electron. Lett. 21:1131–1132 (1985).

[64] Mitachi S and Manabe T. Fluoride glass fiber for infrared transmission. Jpn. J. Appl. Phys. 19:L313–L314 (1980).

[65] Mitachi S, Shibata S, and Manabe T. Teflon FEP-clad fluoride glass fibre. Electron. Lett. 17:128–129 (1981).

[66] Mitachi S, Miyashita T, and Kanamori T. Fluoride-glass-cladded optical fibres for mid-infra-red ray transmission. Electron. Lett. 17:591–592 (1981).

[67] Tran DC, Fisher CF, and Sigel GH Jr. Fluoride glass preforms prepared by a rotational casting process. Electron. Lett. 18:657–658 (1982).

[68] Carter SF, Moore MW, Szebesta D, Williams JR, Ranson D, and France PW. Low loss fluoride fibre by reduced pressure casting. Electron. Lett. 26:2115–2117 (1990).

[69] Braglia M, Bruschi C, Chierici E, Dai G, Kraus J, and Mosso S. Fabrication of Er^{3+}-doped fluoride fibers with single-mode geometry based on over-clad drilling. J. Non-Cryst. Solids 256&257:220–225 (1999).

[70] Kanamori T, Terunuma Y, Nishida Y, Hoshino K, Nakagawa K, Ohishi Y, and Sudo S. Fabrication of fluoride single-mode fibers for optical amplifiers. J. Non-Cryst. Solids 213&214:121–125 (1997).

[71] Mitachi S, Miyashita T, and Manabe T. Preparation of fluoride optical fibres for transmission in the mid-infrared. Phys. Chem. Glasses 23:196–201 (1982).

[72] Kortan AR, Kopylov N, Ozdas E, and Pafchek R. A novel method for fabricating optical fiber preforms. J. Non-Cryst. Solids 213&214:90–94 (1997).

[73] Mitachi S, Sakaguchi S, Yonezawa H, Shikano K, Shigematsu T, and Takahashi S. Relationship between scattering loss and oxygen content in fluoride optical fibers. Jpn. J. Appl. Phys. 24:L827–L828 (1985).

[74] France PW, Carter SF, Day CR, and Moore MW. Optical properties and applications, in *Fluoride Glasses* (Alan E. Comyns, ed.), Chapter 5. John Wiley & Sons, NY, 1989, pp. 87–122.

[75] Szebesta D, Davey ST, Williams JR, and Moore MW. OH absorption in the low loss window of ZBLAN(P) glass fiber. J. Non-Cryst. Solids 161:18–22 (1993).

[76] Yasui I, Hagihara H, and Inoue H. The effect of addition of oxides on the crystallization behavior of aluminum fluoride-based glasses. J. Non-Cryst. Solids 140:130–133 (1992).

[77] Ohishi Y, Mitachi S, and Kanamori T. Impurity absorption losses in the infrared region due to 3d transition elements in fluoride glass. Jpn. J. Appl. Phys. 20:L787–L788 (1981).

[78] Aasland S, Grande T, and Julsrud S. Oxidation states of transition metals in ZBLAN. J. Non-Cryst. Solids 161:177–181 (1993).

[79] Jewell JM, Williams GM, Jaganathan J, and Aggarwal ID. Separation of intrinsic and extrinsic optical absorption in a fluoride glass. Appl. Phys. Lett. 59:1–3 (1991).

[80] Ohishi Y, Mitachi S, Shibata S, and Manabe T. Impurity loss due to rare earth elements in a fluoride glass. Jpn. J. Appl. Phys. 20:L191–L193 (1981).

[81] Poulain M and Saad M. Absorption loss due to complex anions in fluorozirconate glasses. J. Lightwave Technol. 2:599–602 (1984).

[82] France PW, Carter SF, and Williams JR. NH_4^+ absorption in fluoride glass infrared fibers. Comm. Am. Ceram. Soc. 67:C-243–C-244 (1984).

[83] Hopgood AA and Rosman G. Effects of thermal history on crystal size distributions and scattering in fluoride fibres. J. Non-Cryst. Solids 140:301–306 (1992).

[84] Sakaguchi S. Evaluation of extrinsic scattering loss due to crystallization in fluoride optical fibers. J. Lightwave Technol. 11:187–191 (1993).

[85] Harbison BB and Aggarwal ID. Hot isostatic pressing of fluoride glasses. J. Non-Cryst. Solids 161:7–9 (1993).

[86] Schneider HW. Strength in fluoride fibers. Mat. Sci. Forum 32–33:561–570 (1988).

[87] Pureza PC, Brower DT, and Aggarwal ID. Fluoride glass fibers with improved mechanical strength. J. Am. Ceram. Soc. 72:1980–1981 (1989).

[88] Delben A, Messadeq Y, Caridade MD, Aegerter MA, and Eiras JA. Mechanical properties of ZBLAN glasses. J. Non-Cryst. Solids 161:165–168 (1993).

[89] Orcel G and Biswas D. Influence of processing parameters on the strength of fluoride glass fibers. J. Am. Ceram. Soc. 74:1373–1377 (1991).

[90] Bellcore. Generic requirements for passive fiber optical component reliability assurance practices, Technical Advisory TA-NWT–001221 (1992).

[91] Fujiura K, Nishida Y, Kanamori T, Terunuma Y, Hoshino K, Nakagawa K, Ohishi Y, and Sudo S. IEEE Photonics Technology Letters 10:946–948 (1998).

[92] Nishida Y, Fujiura K, Hoshino K, Shimizu M, Yamada M, Nakagawa K, and Ohishi Y. Reliability of fluoride fiber module for optical amplifier use. IEEE Photonics Technology Letters 11:1596–1598 (1999).

[93] Brierly MC and France PW. Neodymium-doped fluoro-zirconate fibre laser. Electron. Lett. 23:815–817 (1987).

[94] Miniscalco WJ, Andrews LJ, Thompson BA, Quimby RS, Vacha LJB, and Drexhage MG. 1.3 mm fluoride fibre laser. Electron. Lett. 24:28–29 (1988).

[95] Brierley MC and Millar CA. Amplification and lasing at 1350 nm in a neodymium doped fluorozirconate fibre. Electron. Lett. 24:438–439 (1988).

[96] Brierley MC, France PW, and Millar CA. Lasing at 2.08 μm and 1.38 μm in a holmium doped fluoro-zirconate fibre laser. Electron. Lett. 24:539–540 (1988).

[97] Brierley MC, France PW, Moore MW, and Davey ST. Rare-earth-doped fluorozirconate fiber lasers. Conference on Lasers and Electro-Optics, 1988, Technical Digest Series (CLEO '88), V. 7:TUM29 (1988).

[98] Brierley MC and France PW. Continuous wave lasing at 2.7 μm in an erbium-doped fluorozirconate fibre. Electron. Lett. 24:935–937 (1988).

[99] Hanna DC, Percival RM, Perry IR, Smart RG, and Tropper AC. Efficient operation of an Yb-sensitized Er fibre laser pumped in 0.8 μm region. Electron. Lett. 24:1068–1069 (1988).

[100] Hanna DC, Percival RM, Perry IR, Smart RG, Suni PJ, Townsend JE, and Tropper AC. Continuous-wave oscillation of a monomode ytterbium-doped fibre laser. Electron. Lett. 24:1111–1113 (1988).

[101] Hanna DC, Jauncey IM, Percival RM, Perry IR, Smart RG, Suni PJ, Townsend JE, and Tropper AC. Continuous-wave oscillation of a monomode thulium-doped fibre laser. Electron. Lett. 24:1222–1223 (1988).

[102] Allen R and Esterowitz L. CW diode pumped 2.3 μm fiber laser. Appl. Phys. Lett. 55:721–722 (1989).

[103] Allain JY, Monerie M, and Poignant H. Tunable CW lasing around 0.82, 1.48, 1.88 and 2.35 μm in thulium-doped fluorozirconate fibre. Electron. Lett. 25:1660–1662 (1989).

[104] Carter SF, Szebesta D, Davey ST, Wyatt R, Brierley MC, and France PW. Amplification at 1.3 μm in a Pr^{3+}-doped single-mode fluorozirconate fibre. Electron. Lett. 27:628–629 (1991).

[105] Durteste Y, Monerie M, Allain JY, and Poignant H. Amplification and lasing at 1.3 μm in praseodymium-doped flurozirconate fibres. Electron. Lett. 27:626–628 (1991).

[106] Miyajima Y, Komukai T, and Sugawa T. 1.31–1.36 μm optical amplification in Nd^{3+}-doped fluorozirconate fibre. Electron. Lett. 26:194–195 (1990).

[107] Brierley M, Carter S, and France P. Amplification in the 1300 nm telecommunications window in a Nd-doped fluoride fibre. Electron. Lett. 26:329–330 (1990).

[108] Pederson JE, Brierley MC, and Lobbett RA. Noise characterization of a neodymium-doped fluoride fiber amplifier and its performance in a 2.4 Gb/s system. IEEE Photon. Technol. Lett. 2:750–752 (1990).

[109] Sugawa T, Miyajima Y, and Komukai T. 10 dB gain and high saturation power in a Nd^{3+}-doped fluorozirconate fibre amplifier. Electron. Lett. 26:2042–2044 (1990).

[110] Kogo T, Onishi M, Hirai S, Chigusa Y, Nakazato K, Shigematsu M, Suzuki S, and Watanabe M. A study of the control of Nd^{3+} energy levels in fluorozirconate glasses. J. Non-Cryst. Solids 140:233–237 (1992).

[111] Lobbett R, Wyatt R, Eardley P, Whitley TJ, Smyth P, Szebesta D, Carter SF, Davey ST, Millar CA, and Brierley MC. System characterisation of high gain and high saturated output power, Pr^{3+}-doped fluorozirconate fibre amplifier at 1.3 μm. Electron. Lett. 27:1472–1474 (1991).

[112] Pederson B, Miniscalco WJ, and Quimby RS. Optimization of Pr^{3+}:ZBLAN fiber amplifiers. IEEE Photon. Technol. Lett. 4:4446–4448 (1992).

[113] Yamada M, Shimizu M, Ohishi Y, Temmyo J, Kanamori T, and Sudo S. One-LD-pumped Pr^{3+}-doped fluoride fibre amplifier module with signal gain of 23 dB. Electron. Lett. 29:1950–1951 (1993).

[114] Whitley TJ. A review of recent system demonstrations incorporating 1.3-μm praseodymium-doped fluoride fiber amplifiers. J. Lightwave Technol. 13:744–760 (1995).

[115] Carter JN, Smart RG, Tropper AC, and Hanna D. Thulium-doped fluorozirconate fibre lasers. J. Non-Cryst. Solids 140:10–15 (1992).

[116] Komukai T, Yamamoto T, Sugawa T, and Miyajima Y. Efficient upconversion pumping at 1.064 μm of Tm^{3+}-doped fluoride fibre laser operating around 1.47 μm. Electron. Lett. 28:830–832 (1992).

[117] Allen R, Esterowitz L, and Aggarwal I. An efficient 1.46 μm thulium fiber laser via a cascade process. IEEE Quant. Electron. Lett. 29:303–306 (1993).

[118] Komukai T, Yamamoto T, Sugawa T, and Miyajima Y. Upconversion pumped thulium-doped fluoride fiber amplifier and laser operating at 1.47 μm. IEEE J. Quant. Electron. 31:1880–1889 (1995).

[119] Ronarc'h D, Guibert M, Ibrahim H, Monerie M, Poignant H, and Tromeur A. 30 dB optical net gain at 1.543 μm in Er^{3+}-doped fluoride fibre pumped around 1.48 μm. Electron. Lett. 27:908–909 (1991).

[120] Clesca B, Ronarc'h D, Bayart D, Sorel Y, Hamon L, Guibert M, Beylat JL, Kerdiles JF, and Semenkoff M. Gain flatness comparison between erbium-doped fluoride and silica fiber amplifiers with wavelength-multiplexed signals. IEEE Photon. Technol. Lett. 6:509–512 (1994).

[121] Okada T, Nakata Y, and Yoshimura T. Improved fluorescence characteristics of Er-doped fluoride glass by Ce co-doping. Conference on Lasers and Electro-Optics, 1999, Technical Digest Series (CLEO '99), V. 7:CMC6 (1999).

[122] Sun KH. Fluoride glasses. J. Am. Ceram. Soc. 30:277–281 (1947).

[123] Sun KH. Fluoride Glass, US Patent No. 2,466–509 (1949).

[124] Kitamura N, Hayakawa J, and Yamashita H. Optical properties of fluoroaluminate glasses in the UV region. J. Non-Cryst. Solids 126:155–160 (1990).

[125] Chen H and Gan F. Vibrational spectra and structure of AlF3-YF3 fluoride glasses. J. Non-Cryst. Solids 112:272–275 (1989).

[126] Cardoso AV, Korgul P, and Seddon AB. Identification of crystalline phases found in multicomponent AlF_3-based glasses. J. Non-Cryst. Solids 161:56–59 (1993).

[127] Iqbal T, Shahriari MR, Foy P, Ulbrich R, and Sigel GH Jr. AlF_3-based glass fibres with enhanced optical transmission. Electr. Lett. 27:110–111 (1991).

[128] Iqbal T, Shahriari MR, Merberg G, and Sigel GH Jr. Synthesis, characterization, and potential application of highly chemically durable glasses based on AlF_3. J. Mat. Res. 6:401–406 (1991).

[129] Shahriari MR, Iqbal T, Hajcak P, and Sigel GH Jr. The effect of rare earth ions on the thermal stability of AlF_3-based glasses. J. Non-Cryst. Solids 161:77–80 (1993).

[130] Iqbal T, Shahriari MR, Weitz G, and Sigel GH Jr. New highly stabilized AlF_3-based glasses. J. Non-Cryst. Solids 184:190–193 (1995).

[131] Colaizzi J and Matthewson MJ. Mechanical durability of ZBLAN and aluminum fluoride-based optical fiber. J. Lightwave Technol. 12:1317–1324 (1994).

[132] Naftaly M, Jha A, and Taylor ER. Spectroscopic properties of Nd^{3+} in fluoroaluminate glasses for an efficient 1.3 μm optical amplifier. J. Non-Cryst. Solids 256&257:248–252 (1999).

[133] Zachariasen WH. The atomic arrangement in glass. J. Am. Ceram. Soc. 54:3841–3851 (1932).

[134] Gan F. Optical properties of fluoride glasses: a review. J. Non-Cryst. Solids 184:9–20 (1995).

[135] Fujino S, Muller M, and Morinaga K. Vacuum ultraviolet transmission of aluminium fluoroberyllate glasses. Proc. SPIE 4102:324–331 (2000).

[136] Namikawa H, Arai K, Kumata K, Ishii Y, and Tanaka H. Preparation of Nd-doped SiO_2 glass by plasma torch CVD. Jpn. J. Appl. Phys. Lett. 21:L360–L362 (1982).

[137] Namikawa H, Ishii Y, Kumata K, Arai K, Iida I, and Tsuchiya T. Preparation of Nd-doped SiO_2 glasses by axial injection plasma torch CVD and their fluorescence properties. Jpn. J. Appl. Phys. Lett. 23:L409–L411 (1984).

[138] Arai K, Namikawa H, Kumata K, Honda T, Ishii Y, and Handa T. Aluminum or phosphorus co-doping effects on fluorescence and structural properties of neodymium-doped silica glass. J. Appl. Phys. 59:3430–3436 (1986).

[139] Morkel PR, Farries MC, and Poole SB. Spectral variation of excited state absorption in neodymium doped fibre lasers. Opt. Comm. 67:349–352 (1988).

[140] Arai K, Yamasaki S, Isoya J, and Namikawa H. Electron-spin-echo envelope-modulation study of the distance between Nd^{3+} ions and Al^{3+} ions in the co-doped SiO_2 glasses. J. Non-Cryst. Solids 196:216–220 (1996).

[141] Sen S. Atomic environments of high-field strength Nd and Al cations as dopants and major components in silicate glasses: a Nd L_{III}-edge and Al K-edge X-ray absorption spectroscopic study. J. Non-Cryst. Solids 261:226–236 (2000).

[142] Jin J, Sakida S, Yoko T, and Nogami M. The local structure of Sm-doped aluminosilicate glasses prepared by sol-gel method. J. Non-Cryst. Solids 262:183–190 (2000).

[143] Craig-Ryan SP, Ainslie BJ, and Millar CA. Fabrication of long lengths of low excess loss erbium-doped optical fibre. Electron. Lett. 26:185–186 (1990).

[144] Desurvire E. Erbium doped fiber amplifiers; principles and applications. John Wiley and Sons, NY, pp. 244–255 and references therein (1993).

[145] Risbud SH, Kirkpatrick RJ, Taglialavore AP, and Montez B. Solid-state NMR evidence fof 4-, 5- and 6-fold alunimum sites in roller-quenched SiO_2–Al_2O_3 glasses. J. Am. Ceram. Soc. 70:C10–C12 (1987).

[146] Schmucker M, MacKenzie KJD, Schneider H, and Meinhold R. NMR studies on rapidly solidified SiO_2–Al_2O_3 and SiO_2–Al_2O_3–Na_2O glasses. J. Non-Cryst. Solids 217:99–105 (1997).

[147] Betts RA, Ruhl FF, Kwok TM, and Zheng GF. Optical amplifiers based on phosphorus co-doped rare-earth-doped optical fibres. Int. J. Optoelectron. 6:47–64 (1991).

[148] Kakui M, Kashiwada T, Onishi M, Shigematsu M, and Nishimura M. Optical amplification characteristics around 1.58 μm of silica-based erbium-doped fibers containing phosphorus/alumina as codopants. Optical Amplifiers and their Applications (OAA 98), OSA TOPS V. 25:68–71 (1998).

[149] Hyde JF. Method of making a transparent article of silica. US Patent No. 2,272,342 (1934).

[150] Keck DB, Schultz PC, and Zimar F. Method of forming optical waveguide fibers. US Patent No. 3,737,282 (1973).

[151] MacChesney JB and DiGiovanni DJ. Materials development of optical fiber. J. Am. Ceram. Soc. 73:3537–3556 (1990).

[152] Ainslie BJ, Armitage JR, Craig SP, and Wakefield B. Fabrication and optimisation of the erbium distribution in silica based doped fibres. Fourteenth European Conference on Optical Communcation (ECOC '88), Conf. Publ. 292, V. 1:62–65 (1988).

[153] Ainslie BJ. A review of the fabrication and properties of erbium-doped fibers for optical amplifiers. J. Lightwave Technol. 9:220–227 (1991).

[154] Matejec V, Kasik I, and Pospisilova M. Preparation and optical properties of silica optical fibres with an Al_2O_3-doped core. J. Non-Cryst. Solids 192&193:195–198 (1995).

[155] Matejec V, Hayer M, Pospisilova M, and Kasik I. Preparation of optical cores of silica optica fibers by the sol-gel method. J. Sol-Gel Sci. Technol. 8:889–893 (1997).

[156] Schultz PC. Optical absorption of the transition elements in vitreous silica. J. Am. Ceram. Soc. 57:309–313 (1974).

[157] Pederson B, Miniscalco WJ, and Zemon SA. Evaluation of the 800-nm pump band for erbium-doped fiber amplifiers. J. Lightwave Technol. 10:1041–1049 (1992).

[158] Wyatt R. High-power broadly tunable erbium-doped silica fibre laser. Electron. Lett. 25:1498–1499 (1989).

[159] Massicott JF, Wyatt R, Ainslie BJ, and Craig-Ryan SP. Efficient, high power, high gain, Er^{3+} doped silica fibre amplifier. Electron. Lett. 26:1038–1039 (1990).

[160] Wang JS, Vogel EM, and Snitzer E. Tellurite glass: a new candidate for fiber devices. Opt. Mat. 3:187–203 (1994).

[161] Wang JS, Vogel EM, Snitzer E, Jackel JL, da Silva VL, and Silberberg Y. 1.3 μm emission of neodymium and praseodymium in tellurite-based glasses. J. Non-Cryst. Solids. 178:109–113 (1994).

[162] Man SQ, Pun EYB, and Chung PS. Tellurite glasses for 1.3 μm optical amplifiers. Optics Comm. 168:369–373 (1999).

[163] McDougall J, Hollis DB, and Payne MJP. Spectroscopic properties of Er^{3+} and Tm^{3+} in tellurite glasses. Phys. Chem. Glasses 37:254–255 (1996).

[164] Le Neindre L, Jiang S, Hwang B-C, Luo T, Watson J, and Peyghambarian N. Effect of relative alkali content on absorption linewidth in erbium-doped tellurite glasses. J. Non-Cryst. Solids 255:97–102 (1999).

[165] Jha A, Shen S, and Naftaly M. Structural origin of spectral broadening of 1.5-μm emission in Er^{3+}-doped tellurite glasses. Phys. Rev. B62:6215–6227 (2000).

[166] Naftaly M, Shen S, and Jha A. Tm^{3+}-doped tellurite glass for a broadband amplifier at 1.47 μm. Appl. Opt. 39:4979–4984 (2000).

[167] Shen S, Naftaly M, Jha A, and Wilson SJ. Thulium-doped tellurite glasses for S-band amplification. Optical Fiber Communications Conference and Exhibit, 2001. OFC 2001, V. 2:TuQ6-1 – TuQ6-3 (2001).

[168] Choi YG, Cho DH, and Kim KH. Influence of 4f absorption transitions of Dy^{3+} on the emission spectra of Tm^{3+}-doped tellurite glasses. J. Non-Cryst. Solids 276:1–7 (2000).

[169] Cho DH, Choi YG, and Kim KH. Energy transfer from Tm^{3+} : $^3F_4$ to Dy^{3+} : $^6H_{11/2}$ in oxyfluoride tellurite glasses.

[170] Sakamoto T, Shimizu M, Kanamori T, Terunuma Y, Ohishi Y, Yamada M, and Sudo S. 1.4-mm-band gain characteristics of a Tm-Ho-doped ZBLYAN fiber amplifier pumped in the 0.8-mm band. IEEE Photon. Technol. Lett. 7:983–985 (1995).

[171] Tanabe S, Feng X, and Hanada T. Improved emission of Tm^{3+}-doped glass for a 1.4-μm amplifier by radiative energy transfer between Tm^{3+} and Nd^{3+}. Opt. Lett. 25:817–819 (2000).

[172] Jiang C, Gan F, Zhang J, Deng P, and Huang G. Yb:tellurite laser glass with high emission cross-section. Mat. Lett. 41:209–314 (1999).

[173] Duverger C, Bouazaoui M, and Turrell S. Raman spectroscopic investigations of the effect of doping metal on the structure of binary tellurium-oxide glasses. J. Non-Cryst. Solids 220:169–177 (1997).

[174] Braglia M, Mosso S, Dai G, Billi E, Bonelli L, Baricco M, and Battezzati L. Rheology of tellurite glass. Mat. Res. Bull. 35:2343–2351 (2000).

[175] Pan Z and Morgan SH. Raman spectra and thermal analysis of a new lead-tellurium-germanate glass system. J. Non-Cryst. Solids 210:130–135 (1997).

[176] Pan Z and Morgan SH. Optical transitions of Er^{3+} in lead-tellurium-germanate glasses. J. Lumin. 75:301–308 (1997).

[177] Feng X, Tanabe S, and Hanada T. Hydroxyl groups in erbium-doped germanotellurite glasses. J. Non-Cryst. Solids 281:48–54 (2001).

[178] Sahar MR and Noordin N. Oxychloride glasses based on the TeO_2–ZnO–$ZnCl_2$ system. J. Non-Cryst. Solids 184:137–140 (1995).

[179] Sidebottom DL, Hruschka MA, Potter BG, and Brow RK. Structure and optical properties of rare earth-doped zinc oxyhalide tellurite glasses. J. Non-Cryst. Solids 222:282–289 (1997).

[180] Ding Y, Jiang S, Hwang B-C, Luo T, Peyghambarian N, Himei Y, Ito T, and Miura Y. Spectral properties of erbium-doped lead halotellurite glasses for 1.5 μm broadband amplification. Opt. Mat. 15:123–130 (2000).

[181] Mori A, Ohishi Y, and Sudo S. Erbium-doped tellurite glas fibre laser and amplifier. Electron. Lett. 33:863–864 (1997).

[182] More A, Ohishi Y, Yamada M, Onon H, and Sudo S. Broadband amplification characteristics of tellurite-based EDFAs. European Conference on Optical Communication 1997, ECOC '97. Conference Publication 448, V. 3, 135–138 (1997).

[183] Choi YG, Lim DS, Kim KH, Cho DH, and Lee HK. Enhanced $^4I_{11/2} \rightarrow {}^4I_{13/2}$ transition rate in Er^{3+}/Ce^{3+}-codoped tellurite glasses. Electron. Lett. 35:1765–1766 (1999).

[184] Chryssou CE, Di Pasquale F, and Pitt CW. Er^{3+}-doped channel optical waveguide amplifiers for WDM systems: A comparison of tellurite, alumina and Al/P silicate materials. IEEE J. Selected Topics Quant. Electron. 6:114–121 (2000).

[185] Marhic ME, Morita I, Ho M-C, Akasaka Y, and Kazovsky LG. Large cross-phase modulation and four wave mixing in tellurite EDFAs. Electron. Lett. 35:2045–2047 (1999).

[186] Sakamoto T, Mori A, Hoshino K, Shikano K, and Shimizu M. Suppression of nonlinear effects in tellurite-based EDFAs by fiber parameter modification. Optical Fiber Communication Conference and Exhibit. OFC 2001, V. 2: TuQ6-1 – TuQ6-3 (2001).

[187] Prohaska JD, Garton RT, Wang JS, Wu F, and Snitzer E. Observation of UV induced absorption changes in tellurite and selected doped silica glasses. Lasers and Electro-Optics Society Annual Meeting, 1994. LEOS '94 Conference Proceedings. IEEE, V. 2:133–134 (1994).

[188] Shimizu M, Ohmori Y, and Nakamhara M. Fabrication of antimony oxide-doped silica fibres by the VAD process. Electron. Lett. 21:872–872 (1985).

[189] Shimizu M and Ohmori Y. Antimony oxide-doped silica fibers fabricated by the VAD method. J. Lightwave Technol. LT-5:763–769 (1987).

[190] Susa K, Satoh S, and Matsuyama I. Optical properties of sol-gel derived Sb-doped silica glass. J. Non-Cryst. Solids 221:125–134 (1997).

[191] Miniscalco WJ. Erbium-doped glasses for fiber amplifiers at 1500 μm J. Lightwave Technol. 9:234–250 (1991).

[192] FuKuchi K, Kasamatsu T, Morie M, Ohhira R, and Ito T. 10.92-Tb/S (273 × 40 Gbit/s) triple band/ultra dense WDM optical repeatered transmission experiments, Proc OFC 2001, Anaheim, PD 24 (2001).

[193] Bigo S, Frignac Y, Charlet G, Borne S, Tran P, Simonneau C, Bayert D, Jourdan A, Hamaide JP, Idler W, Dischler R, Vieth G, Gross H, and Poelmann W. 10.2 Tbits/s 256 × 42.7 gbits/s PMD, WDM, transmission over 100 km Tera-Light fiber with 1.28 bits/Hz spectral efficiency, Proc. OFC 2001, Anaheim, PD25 (2001).

[194] Desurvire E. Erbium doped fiber amplifiers; principles and applications. John Wiley and Sons, NY, pp. 480–487 (1993).

[195] Giles CR and Desurvire E. Modelling erbium-doped fiber amplifiers. IEEE J. Lightwave Technology, 9(2):271 (1991).

[196] Mears RJ, Reekie L, Jauncey IM, and Payne DN. Low-noise erbium-doped fibre amplifier operating at 1.54 μm. Electron. Lett., 23(19):1026 (1987).

[197] Desurvire E, Simpson JR, and Becker PC. High-gain erbium-doped traveling wave fiber amplifier. Optics Lett. 12(11):888 (1987).

[198] Zyskind JL, Giles CR, Desurvire E, and Simpson JR. Optimal pump wavelength in the $^4I_{15/2}$ $^{-4}I_{13/2}$ absorption band for efficient Er^{3+-} doped fiber amplifiers. IEEE Photon. Technol. Lett., 1(12):428 (1989).

[199] Ainslie BJ. A review of the fabrication and properties of erbium doped fibers for optical amplifiers. IEEE J. Lightwave Tech. 9(2):220 (1990).

[200] Ainslie BJ, Craig-Ryan SP, Davey ST, Armitage JR, Atkins CG, and Wyatt R. Optical analysis of erbium-doped fibres for efficient lasers and amplifiers. Proc. IOOC '89, paper 20A3-2, p. 22, Kobe, Japan (1989).

[201] Atkins CG, Massicott JF, Armitage JR, Wyatt R, Ainslie BJ, and Craig-Ryan SP. High-gain broad spectral bandwidth erbium-doped fiber amplifier pumped near 1.5 μm. Electron. Lett. 25: 910–911 (1989).

[202] Fevier H, Auge J, Parlier V, Bousselet PH, Dursin A, Marcerou JF, and Jacquier B. Erbium-doped fiber amplifier with outstanding gain characteristics. ECOC '89, 1989, paper TUA5-2, pp. 66–69.

[203] Armitage JR. Spectral dependence of the small-signal gain around 1.5 μm in erbium-doped silica fiber amplifiers. IEEE J. Quantum Electron. 26:423,425 (1990).

[204] Kashiwada T, Nakazato K, Ohnishi M, Kanamori H, and Nishimura M. Spectral gain behavior of Er-doped fiber with extremely high aluminum concentration. OAA '93, 1993, paper MA6, pp. 104–107.

[205] Yoshida S, Kuwano S, and Iwashita K. Gain-flattened EDFA with high Al concentration for multistage repeatered WDM transmission systems. Electron. Lett. 31:1765–1767 (1995).

[206] Inoue K, Kominato T, and Toba H. Tunable gain equalization using a Mach-Zehnder optical filter in a multistage fiber amplifier. IEEE Photon. Technol. Lett. 3:718–720 (1991).

[207] Tachibana N, Laming RI, Morkel PR, and Payne DN. Gain-shaped erbium-doped fiber amplifier (EDFA) with broad spectral bandwidth. OAA, 1990, paper MDI, pp. 44–47.

[208] Betts RA, Frisken SJ, and Wong D. Split-beam Fourier filter and its application in a gain-flattened EDFA. OFC '95, 1995, paper TuP4, pp. 80–81.

[209] Hall DW. Proceedings OAA 1998, invited paper, Vale, 1998.

[210] Bayart D, Clesca B, Hamon L, and Beylat JL. 1.55 μm fluoride-based EDFA with gain-flatness control for multiwavelength applications. Electron. Lett. 30(17):1407–1409 (1994).

[211] Clesca B, Ronarc'h D, Bayart D, Sorel Y, Hamon L, Guibert M, Beylat JL, Kerdiles JF, and Semenkoff M. Gain flatness comparison between erbium-doped fluoride and silica fiber amplifiers with wavelength-multiplexed signals. IEEE Photon. Technol. Lett. 6(4):509–512 (1994).

[212] Yamada M, Ono H, Kanamori T, Sakamoto T, Ohishi Y, and Sudo S. A low-noise and gain-flattened amplifier composed of a silica-based and a fluoride-based Er^{3+}-doped fiber amplifier in a cascade configuration. IEEE Photon. Technol. Lett. 8(5):620–622 (1996).

[213] Ellison AJE, Dickinson JE, Goforth DE, Harris DL, Kohli JT, Minelly JD, Samson BN, and Trentelman JP. Hybrid erbium silicate conventional band fiber amplifier with ultra low gain ripple. Proceeding OAA 1999, Nara, Japan, PD2 (1999).

[214] Goforth DE, Minelly JD, Ellison AJE, Wang DS, Trentelman JP, and Nolan DA. Proc. NFOEC, 2000, paper B1 (2000).

[215] Massicott JF, Armitage JR, Wyatt R, Ainslie BJ, and Craig-Ryan SP. High gain broadband 1.6 μm Er^{3+} doped silica fiber amplifier. Electron. Lett. 26(20):1645 (1990).

[216] Sun Y, Sulhoff JW, Srivasta AK, Abramov A, Strasser TA, Wysocki PF, Pedrazzani JR, Judkins JB, Espindola RP, Wolf C, Zyskind JL, Vengsarker AM, and

Zhou J. A gain flattened ultra wideband EDFA for high capacity WDM optical communications systems. Proceedings ECOC 19198, Madrid, pp. 53–54 (1998).

[217] Ellison AJG, Goforth DE, Samson BN, Minelly JD, Trentelman JP, McEnroe DL, and Tyndell BP. Extending the L-band to 1620 nm Using MCS Fiber. Proceedings OFC 2001, Anaheim (2001), TUA-2-1.

[218] Mori A, Sakamoto T, Shikano K, Kobayashi K, Hoshino K, and Shimizu M. Gain flattened Er3+-doped tellurite fibre amplifier for WDM signals in the 1581–1616 nm wavelength region. Electron. Lett. 36(7):621–622 (2000).

[219] Ohishi Y, Mori A, Yamada M, Ono M, Nishida Y, and Oikawa K. Gain characteristics of tellurite-based erbium-doped fiber amplifiers for 1.5-mm broadband amplification. Optics Lett. 23(4):274–276 (1998).

[220] Mori A, Kobayashi K, Yamada M, Kanamori T, Oikawa K, Nishida Y, and Ohishi Y. Low noise broadband tellurite-based Er3+-doped fiber amplifiers. Electron. Lett. 30(9):887–888 (1998).

[221] Yamada M, Mori A, Kobayashi K, Ono H, Kanamori T, Oikawa K, Nishida Y, and Ohishi Y. Gain-flattened tellurite-based EDFA with a flat amplification bandwidth of 76 nm. IEEE Photonics Tech. Lett. 10(9):1244–1247 (1998).

[222] Yamada M, Mori A, Kobayashi K, Ono H, Kobayashi K, Kanamori T, and Ohishi Y. Broadband and gain-flattened Er-doped tellurite fiber amplifier constructed using a gain equalizer. Electron. Lett. 34(4):370–371 (1998).

[223] Goforth DE, Minelly JD, Ellison AJG, Trentelman JP, and Samson BN. Ultra-Wide Band Erbium Amplifiers Using a Multi-Component Silicate Fiber. Proceedings OAA 2000, Quebec, paper OTuA4 (2000).

[224] Percival RM and Williams JR. Highly efficient 1.064 mm upconversion pumped 1.47 um thulium doped fluoride fiber amplifier. Electron. Lett. 30, 20, pp. 1684–1685 (1994).

[225] Komukai T, Yamamoto T, Sugawa T, and Miyajima Y. 1.47 um band Tm3+ doped fluoride fibre amplifier using a 1.064 mm upconversion pumping scheme. Electron. Lett. 29(1):110–112 (1993).

[226] Rosenblatt GH, Ginther RJ, Stoneman RC, and Esterowitz L. Laser emmision at 1.47 μm from fluorozirconate glass doped with Tm3+ and Tb3+. OSA Proceedings, Tunable Solid State Lasers, 5:373–376 (1989).

[227] Sakamoto T, Shimizu M, Kanamori T, Ohishi Y, Teranuma Y, Yamada M, and Sudo S. Tm-Ho doped ZBLYAN fiber amplifiers for 1.4 μm band, in Proceedings 5th Optoelectronics conference, Makuhari, Japan, Vol. 13B2-2, pp. 38–39 (1994).

[228] Percival RM, Szebesta D, and Davey ST. Highly efficient CW cascade operation of 1.47 and 182 μm transitions in Tm-doped fluoride fiber laser. Electron. Lett. 28: 1866–1868 (1992).

[229] Kasamatsu T, Yano Y, and Sekita H. Novel 1.50 μm gain-shifted thulium doped fiber amplifier by using a dual wavelength pumping of 1.05 and 1.56 μm. Proc. OAA 1999, Nara, Japan, PD1, 1999.

[230] Aozasa S, Sakamoto T, Kanamori T, Hoshino K, and Shimizu M. Gain-shifted thulium-doped fibre amplifiers employing novel high concentration doping technique. Electron. Lett. 36(5):418–419 (2000).

[231] Aozasa S, Masuda H, Ono H, Sak T, Kanamori T, Ohishi Y, and Shimizu M. 1480–1510 nm-band Tm-doped fiber amplifier with a high power conversion efficiency of 42%. Proc. OFC 2001, PD1 (2001).

[232] Samson BN, Traynor NJ, Walton DT, Ellison AJG, Minelly JD, Trentelman JP, and Dickinson JE. Thulium doped silicate amplifier at 1560–1520 nm. OAA 2000, Quebec, PD6 (2000).

[233] Kasamatsu T, Yano Y, and Ono T. Laser diode pumping (1.4 and 1.56 μm) of gain shifted thulium-doped fiber amplifier. Electron. Lett. 36:19 (2000).

[234] Kinter MES, Walpole JN, Chin SR, Wang CA, and Missaggia LJ. High power strained layer amplifiers and lasers with tapered gain regions. IEEE Photon.Tech. Lett. 5:605–608 (1993).

[235] O'Brien S, Scoenfelder A, and Lang RJ. 5W CW diffraction limited InGaAS broad-area flared amplifier. IEEE Photon. Tech. Lett. 9: 1217–1219 (1997).

[236] Koyanagi S et al. Trends in Optics and Photonics, Vol 25: Optical Amplifiers and Their Applications, Optical Society of America, Washington, DC, pp. 151–154 (1998).

[237] Inniss D et al. Ultra high power single-mode fiber lasers. Proceedings CLEO, CPD 31 (1997).

[238] Hanna DC, Percival RM, Perry IR, Smart IR, Suni PJ, and Tropper AC. Yb-doped monomode fiber laser broadly tunable from 1.010 μm to 1.162 μm and 3-level operation at 974 nm. J. Mod. Optics, 37:329–331 (1987).

[239] Po H, Snitzer E, Tumminelli R, Zenteno L, Hakimi F, Cho NM, and Haw T. High-gain power amplifier tandem pumped by GaAlAs phased array. Proc. OFC 1989, Houston, PD7 (1989).

[240] Minelly JD, Zenteno LA, Dejneka MJ, Miller WJ, Kuksenkov DV, Davis MK, Crigler SG, and Bardo ME. High power diode-pumped single-transverse-mode Yb fiber laser at 978 nm. OFC 2000, Proc. OFC 2000, Baltimore, PD2 (2000).

[241] Nilsson J, Alvarez-Chavez JA, Turner PW, Clarkson WA, Renaud CC, and Grudinin AB. Widely tunable high-power diode-pumped double-clad Yb^{3+}-doped fiber laser. Proc. Adv. Sol. State Lasers, WA2 (1999).

[242] Kukov AS, Medvedkov OI, Paramonov VM, Vasilev SA, Dianov EM, Solodovnikov V, Zhilin V, Guryanov AN, Laptev AY, and Umnikov AA. High power Yb-doped double-clad fiber lasers for 0.98–1.04 μm. Proc. OAA 2001, Stressa, OWC2 (2001).

[243] Zenteno LA, Minelly JD, Liu A, Ellison AJG, Crigler SG, Walton DT, Kuksenkov DV, and Dejneka MJ. 1W single-transverse-mode Yb-doped double-clad fiber laser at 978 nm. Electron. Lett. 37(13): 819–820 (2001).

[244] Nilsson J, Minelly JD, Paschotta R, Tropper AC, and Hanna DC. Ring-doped cladding pumped single-mode three-level fiber laser. Optics Lett. 23(5):255–357 (1998).

Chapter 4 | Advances in Erbium-Doped Fiber Amplifiers

Atul K. Srivastava and Yan Sun

Onetta Inc., Sunnyvale, California

I. Introduction

Driven by unprecedented capacity demand for data transmission, the capacity of lightwave communication systems has undergone enormous growth during the last few years. Laboratory demonstration of high-capacity transmission now exceeds 10 Tb/s capacity [1] and commercial systems are capable of delivering multiterabit capacity. The transmission systems incorporate Wavelength Division Multiplexing (WDM) technique, which can offer capacities much higher than single channel alternatives. Erbium-doped fiber amplification (EDFA) is a key technology that enabled the deployment of multichannel WDM systems. In order to enable this growth in capacity the EDFAs have evolved to provide a higher performance and greater functionality. In addition to the higher performance, there has also been progress in understanding of the dynamic behavior of EDFAs, so it is now possible to model the amplifiers and predict their behavior in dynamic transmission systems and the networks.

The growth in transmission capacity of lightwave systems is shown in Fig. 1, which summarizes both the results of demonstrations in laboratory experiments and the commercial systems. These conclusions can be drawn from the figure: firstly, the advances in the capacity of laboratory experiments since 1994 are estimated to be growing exponentially at a rate of 3 dB per year. Further, the capacity of commercial systems has grown such that the gap between a laboratory demonstration and product availability has shrunk from 6 years in 1994 to less than 2 years at present. In addition to capacity there is also a push to increase the reach of the transmission systems and thus significantly reduce the number of costly regenerators.

In addition to growth in capacity of point-to-point WDM systems, the fiberoptic communication infrastructure is evolving toward dynamic networks. As the number of wavelengths in lightwave systems continues to grow and the separation between regenerators increases, there will be a greater need to add/drop wavelengths at intermediate sites. Additionally, optical cross-connects will be needed to manage capacity and provide connectivity between equipments from different vendors. The Multiwavelength Optical Networking (MONET) project provided demonstration of key network elements such as optical add-drop multiplexers (OADMs) and optical cross-connects

OPTICAL FIBER TELECOMMUNICATIONS, VOLUME IVA

ISBN: 0-12-395172-0

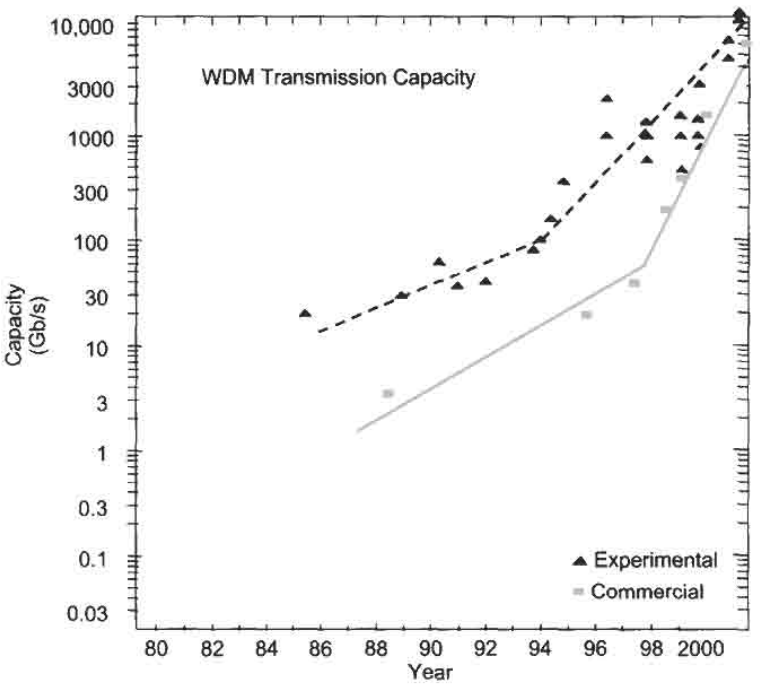


Fig. 1 Transmission capacity growth for lightwave communication systems. The triangles represent laboratory results while the rectangles show progress in commercially deployed systems. The lines are merely guides for the eye.

(OXCs)[2]. These technologies have emerged as key aspects of all optics-based networks. The need for configurable and dynamic networks led to the development of the optical amplifier as an intelligent subsystem capable of adjusting its characteristics in a dynamic way. The dynamics of EDFAs and networks of EDFAs have an important impact on the performance of multichannel optical communication systems and networks with wavelength routing. The fast power transient effect was discovered in EDFA chains constituting networks; this led to the understanding of dynamic behavior of EDFA chains.

This chapter briefly reviews the recent progress in the erbium-doped fiber amplifiers for high-capacity lightwave communication systems and networks. The second section provides the background material, which includes an introduction to the key EDFA parameters that impact the performance of transmission systems and networks. Section III covers some of the key advances in EDFAs, which are crucial for long-distance high-capacity systems. These include wide-bandwidth EDFAs enabled by maturing gain equalization filter technologies, and the realization of L-band EDFAs which has doubled the overall EDFA bandwidth. Related topics of spectral-hole burning, the role of midstage attenuators and nonlinear effects such as cross-phase

modulation (XPM) and four-photon mixing (FPM) on the design and performance of EDFAs in high capacity systems are also covered in this section. Another topic included in this section is that of the Raman amplifier, which has drawn considerable attention in recent years due to its superior noise performance. In particular, a hybrid amplification design consisting of distributed Raman amplification and EDFA (Raman/EDFA), which has led to significant enhancement in system performance, is described. The fourth section is devoted to the advances in the understanding of dynamic phenomena in EDFAs and their impact on the optical networks. After a discussion of the phenomena of fast power transients, several schemes for the gain control in EDFAs are described. Conclusions are presented in the last section.

II. Review of EDFAs and WDM Systems

Since their invention in 1987, significant progress has been made in the understanding and performance of EDFAs. EDFAs have evolved to greater bandwidth, higher output power, lower gain ripple, and increased network friendliness. We focus here on the important system characteristics that are influenced by the amplifier performance. The topics in this section include discussions on optical signal-to-noise ratio (OSNR) and its dependence upon amplifier parameters, system impact of amplifier gain flatness, and amplifier control in WDM networks.

A. OVERVIEW OF EDFAs

In first-generation optical communication systems, optoelectronic regenerators were used between terminals to convert signals from the optical to electrical and then back to the optical domain. Since the first report in 1987 [3,4], the EDFA has revolutionized optical communications. Unlike optoelectronic regenerators, the EDFA does not need high-speed electronic circuitry and is transparent to data rate and format. More importantly, all the optical signal channels can be amplified simultaneously within the EDFA in one fiber, which dramatically reduces cost, thus enabling wavelength-division multiplexing (WDM) technology. Current lightwave systems consist of optical links between two regeneration nodes. The optical links carry multiple wavelengths, each carrying high-bit rate signals to deliver overall high capacity. The EDFAs are used as optical repeaters between fiber spans along the optical link to simultaneously amplify all the WDM channels. The overall performance of the optical link, in terms of capacity and reach, is closely related to the EDFA performance. Some of the key characteristics of the WDM system, which are related to the EDFA performance, are discussed below.

In an EDFA, the erbium-doped fiber (EDF) is usually pumped by 980 nm or 1480 nm semiconductor lasers. In order to obtain low noise figure, the first

stage of a typical EDFA is pumped using 980 nm source, which can create a very high degree of inversion. The power stage in EDFA is pumped at 1480 nm, which provides high quantum conversion efficiency, a measure of the conversion of pump power to signal power. The quantum efficiency of a 980 nm pumped stage is poorer due to greater mismatch of energies of the pump photons as compared to signal photons and significant amount of excited state absorption. A three-level model can be used to describe a 980 nm pumped stage while a two-level model usually suffices for a 1480 nm pumped section [5–7]. A detailed account of the amplifier architecture and its operation is given in the previous edition of this series [8]. A key advantage of EDFAs in high data rate transmission systems is that the spontaneous lifetime of the metastable energy level ($^4I_{13/2}$) is about 10 ms, which is usually much slower than the time corresponding to signal bit rates of practical interest. As a result of the slow dynamics, intersymbol distortion and interchannel crosstalk are negligible.

The gain and loss coefficient spectra at different inversion levels for EDF with Al and Ge codoping are given in Fig. 2. The inversion level at a point in Er fiber is expressed as the fractional percentage $N_2/(N_1 + N_2)$, where N_1 and N_2 are the populations of Er ions in ground and excited states, respectively, and $(N_1 + N_2)$ is the total number of Er ions. Under a homogeneous broadening approximation, the overall gain spectrum of any piece of EDF always matches one of the curves after scaling, and does not depend on the details of pump power, signal power, and saturation level along the fiber. The derivation of a general two-level model that describes spectra and their dependence on EDF length and other parameters is given in [6]. One key parameter in this model is the average fractional upper level density (average inversion level) [7] $\overline{N_2}(t)$

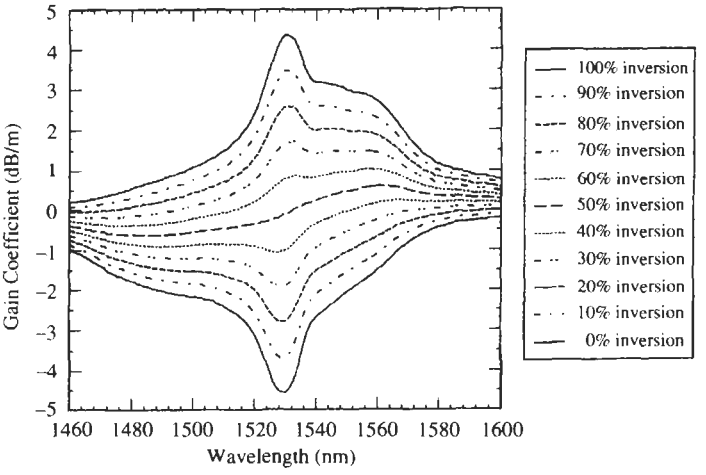


Fig. 2 The gain and loss coefficient spectra at different inversion levels for EDF with Al and Ge codoping.

given by the average of N_2 over the length ℓ of EDF:

$$\overline{N_2}(t) = \frac{1}{\ell}\int_0^{\ell} N_2(z,t)dz$$

In the limit of strong inversion $\overline{N_2}$ takes its maximum value and the gain is highest. When the signal power becomes comparable to pump power, the EDFA is saturated and the level of inversion is reduced. The level of inversion and degree of saturation are closely related and are often used interchangeably. In the limit of low pump and low temperature, $\overline{N_2}(t) = 0$, where the absorption is the strongest. The gain spectra plotted in Fig. 2 are very useful in the study and design of EDFAs, for example in locating the gain peak wavelengths at different inversion levels, finding out inversion levels with wide flat gain range, and understanding relative gain variation at different wavelengths with changing inversion levels. The relative gain spectrum is only a function of the average inversion level for a given type of EDF, while the total integrated gain depends on the EDF length. There are small deviations in the gain spectrum from this model due to inhomogeneity in EDF that gives rise to spectral-hole burning, which is discussed in a subsequent section. Under the homogeneous approximation, however, if there are two EDFAs made of the same type of fiber but of different length, the instantaneous gain spectrum would be similar in shape scaled for the length if the instantaneous length averaged inversion level is the same.

A high inversion level provides a low noise figure, while a low inversion level yields high efficiency in the conversion of photons from pump to signal [6]. To achieve both low noise figure and high efficiency, two or more gain stages are usually used where the input stage is kept at a high inversion level and the output stage is kept at a low inversion level [9,10]. Since the ASE power around the 1530 region can be high enough to cause saturation, an ASE filter can be added in the middle stage to block the ASE in this band [9]. These optical amplifiers were successfully used in early WDM optical networks [11].

For optical amplifiers with two or more gain stages, the overall noise figure is decided mainly by the high-gain input stage, and the output power is basically determined by the strongly saturated output stage [9]. The passive components have minimal impact on noise figure and output power when they are in the midsection of the amplifier. The noise figure of a two-stage EDFA is given by

$$NF = NF_1 + NF_2/L_1G_1$$

where NF_1 and NF_2 are the noise figures of the two stages and L_1 and G_1 are the midstage loss and gain of the first stage, respectively. It can be seen from the above formula that the overall noise figure of a two-stage amplifier is primarily determined by the noise figure of the first stage, since the first stage gain is usually designed to be much larger than middle-stage loss. For example, in a typical case, with first-stage gain $G_1 = 100$ and middle-stage loss $L_1 = 0.1$,

the overall noise figure has only 10 percent contribution from the second stage noise figure.

Besides long-haul communication systems and networks, EDFAs also find important applications in metro area networks and CATV distribution systems. In metro networks [12] the fiber loss is small due to a short distance between nodes, but the total loss can still be high for networks with a large number of OADM sites. Optical amplifiers can be used here to compensate for the loss associated with the OADM or DMUX/MUX, in addition to the transmission fiber. For CATV broadcast systems, amplifiers can be used to increase the signal power for splitting into many branches. Low-cost optical amplifiers are needed for these applications.

B. OPTICAL SIGNAL-TO-NOISE RATIO

In an optically amplified system, channel power reaching the receiver at the end of the link is optically degraded by the accumulated amplified spontaneous emission (ASE) noise from the optical amplifiers in the chain. At the front end of the receiver, ASE noise is converted to electrical noise, primarily through signal-ASE beating, leading to bit-error rate (BER) flooring. System performance depends on the optical signal-to-noise ratio (OSNR) of each of the optical channels. OSNR therefore becomes the most important design parameter for an optically amplified system. Other optical parameters to be considered in system design are channel power divergence, which is generated primarily due to the spectral gain nonuniformity in EDFAs (described in the next section), and maximum channel power relative to the threshold levels of optical nonlinearities such as self-phase modulation, cross-phase modulation and four-photon mixing [13].

Although optical amplifiers are conventionally classified into power, inline and preamplifiers, state-of-the-art WDM systems require all three types of amplifiers to have low noise, high output power, and a uniform gain spectrum. We will not distinguish these three types of amplifiers in the discussion presented in this section. The nominal OSNR for a 1.55 μm WDM system with N optical transmission spans can be given by the following formula [8]:

$$OSNR_{nom} = 58 + P_{out} - 10\log_{10}(N_{ch}) - L_{sp} - NF - 10\log_{10}(N)$$

where $OSNR$ is normalized to 0.1 nm bandwidth, P_{out} is the optical amplifier output power in dBm, N_{ch} is the number of WDM channels, L_{sp} is the fiber span loss in dB, and NF is the amplifier noise figure in dB. For simplicity, it has been assumed here that both optical gain and noise figure are uniform for all channels.

The above equation shows how various system parameters contribute to OSNR; for example, the OSNR can be increased by 1 dB, by increasing the amplifier output power by 1 dB, or decreasing noise figure by 1 dB, or reducing the span loss by 1 dB. This equation indicates that we can make tradeoffs

between number of channels and number of spans in designing a system. However, the tradeoff may not be straightforward in a practical system because of the mutual dependence of some of the parameters. Other system requirements also impose additional constraints; for example, optical nonlinearities place an upper limit on channel power, and this limit depends on number of spans, fiber type, and data rate.

The simple formula above highlights the importance of two key amplifier parameters: noise figure and output power. While it provides valuable guidelines for amplifier and system design, simulating the OSNR evolution in a chain of amplifiers is necessary when designing a practical WDM system. The amplifier simulation is usually based on a mathematical model of amplifier performance. Amplifier modeling is a critical part of the end-to-end system transmission performance simulation that incorporates various linear and nonlinear transmission penalties.

C. AMPLIFIER GAIN FLATNESS

Amplifier gain flatness is another critical parameter for WDM system design. As the WDM channels traverse multiple EDFAs in a transmission system, the spectral gain nonuniformity compounds to create a divergence in channel powers. The worst WDM channel, the channel that consistently experiences the lowest amplifier gain, will have an OSNR value lower than the nominal value. The power deficit, which can be viewed as a form of penalty given rise by amplifier gain nonuniformity, is a complicated function of individual amplifier gain shape [14], and correlation of the shapes of the amplifiers in the chain. The gain flatness is a parameter that can have significant impact on the end-of-system OSNR. The penalty is especially severe for a long amplifier chain, as in the case of long-haul and ultralong-haul applications.

The gain flatness affects system performance in multiple ways; flat-gain amplifiers are essential to getting the system OSNR margin for routed channels and minimizing power divergence to allow practical implementation of networking on the optical layer. Wide bandwidth can enable either large channel spacing as a countermeasure of filter bandwidth narrowing effect or more optical channels for more flexibility routing of traffic. Amplifier gain flatness is critical to maintaining system performance under varied channel loading conditions caused by either network reconfiguration or partial failure.

Figure 3 shows how the OSNR penalty increases as a nonlinear function of the number of transmission spans for three cases: a ripple (flatness) of 1.0, 1.4, and 1.8 dB. The variations in signal strength may exceed the system margin and begin to increase the bit-error rate (BER) if the SNR penalty exceeds 5 dB. When the gain spectrum starts with flatness of 1.8 dB, the OSNR penalty degrades by more than 5 dB after only 8 transmission spans. This OSNR penalty limits the reach of WDM line systems and requires signal regenerators at intervals of approximately 500 kilometers. These expensive

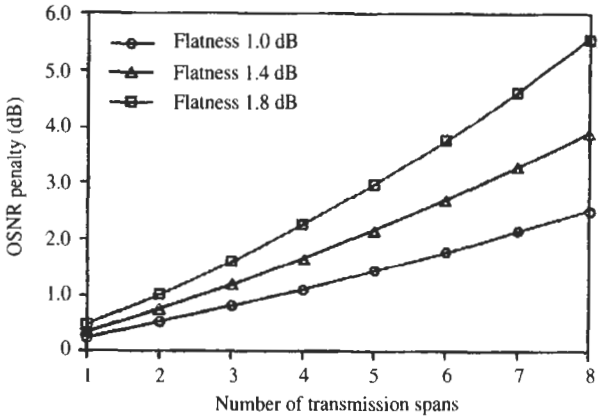


Fig. 3 Optical signal-to-noise ratio penalty vs. number of spans for different EDFA gain ripple [14].

devices convert the signals from the optical domain to the electrical domain, typically reshaping, retiming and reamplifying the signal before triggering lasers to convert the signal back from the electrical domain to the optical domain. In ultralong-haul networks, carriers would like to increase the spacing between regenerators to several thousand kilometers—in which case the signal would have to pass through up to 50 amplifiers without electrical regeneration. These networks require EDFAs with excellent gain uniformity.

The impact of gain nonuniformity, which gives rise to channel power divergence in a chain of amplifiers, is, however, not limited to the OSNR penalty. While the weak channels see an OSNR penalty that limits the system performance, the strong channels continue to grow in power that may reach the nonlinear threshold, also limiting system performance. Additionally, large power divergence increases the total crosstalk from other WDM channels at the optical demultiplexer output. It is thus imperative to design and engineer optical amplifiers with the best gain flatness for WDM networking applications. State-of-the-art optical amplifiers usually incorporate a gain equalization filter to provide uniform gain spectrum, as discussed in Section III. To minimize the residual gain nonuniformity requires careful design, modeling, and engineering of the amplifiers, in particular, the gain equalization filters.

The gain equalization filters are optimized to flatten the gain spectrum of a fully loaded EDFA. But if a carrier wants to operate the system with fewer channels—for example, to reconfigure it dynamically—then in absence of gain control the lower input power can decrease the EDFAs gain uniformity, thereby impairing the effectiveness of the GEF and increasing ripple in the network. Furthermore, as described later, spectral hole burning gives rise to channel loading dependent changes in the gain spectrum of the EDFA by creating a dip in the region of the active channels. Spectral hole burning can

create a gain spectrum for which the GEF was not optimized, making gain flattening very difficult. For all these reasons, future ultra long-haul, dynamically reconfigurable networks will require EDFAs with dynamic gain equalization. The gain spectrum of EDFAs will be equalized by the use of a dynamically controlled filter having variable spectral loss characteristics. The dynamic gain equalizer can be controlled in a feedback loop in conjunction with an optical channel monitor to provide uniform channel powers or OSNR.

D. AMPLIFIER CONTROL

In an amplified system, optical amplifiers may not always operate at the gain value at which its performance, especially gain flatness, is optimized. Many factors contribute to this suboptimal operating condition. Among them is the fact that the span loss can vary at system installation and be maintained in the system's lifetime only to a finite range with respect to the value required by the amplifiers for optimal performance. As a result, amplifier gain will be tilted, and this tilt can have significant impact on system performance in ways similar to gain nonuniformity.

Gain tilt can, if not corrected, result in OSNR penalty and increased power divergence. Control of optical amplifier tilt is often necessary to extend the operational range of the amplifiers and compensate for loss tilt in the system due to, for example, fiber loss variation in the signal band. Control of amplifier gain tilt can be achieved by varying an internal optical attenuator [15,16]. Implementation of such a tilt control function requires a feedback signal that is derived from, for example, measured amplifier gain or channel power spectrum, and an algorithm that coordinates the measurement and adjustment functions. By changing the loss of the attenuator, the average inversion level [6] of the erbium-doped fiber can be adjusted, which affects the gain tilt in the EDFA gain spectrum.

Another important amplifier control function is amplifier power adjustment. In a WDM system, there is a need to adjust the total amplifier output as a function of number of equipped channels. The total output power must be adjusted such that while the per-channel power is high enough to ensure sufficient OSNR at the end of the chain, it is low enough not to exceed the nonlinear threshold. In addition, per-channel power must be maintained within the receiver dynamic range as the system channel loading is changed. Such power adjustment has traditionally been achieved through a combination of channel monitoring and software-based pump power adjustment.

Recent advances in WDM optical networking have called for a power control fast enough to minimize channel power excursion when a large number of channels are changed due to, for example, catastrophic partial system failure. Various techniques, as detailed in Section IV, have been demonstrated to stabilize amplifier gain, thereby achieving the goal of maintaining per-channel power. In addition to amplifier dynamics control, practical implementation in

a system also requires a receiver design that can accommodate power change on a very short time scale.

III. EDFAs for High Capacity Systems

The performance characteristics of EDFAs have evolved significantly in recent years to accommodate the capacity requirements of lightwave systems. In order to support a greater number of WDM channels the EDFA bandwidth and output power requirements have increased proportionately. The bandwidth was nearly doubled by the development of L-band EDFA. Higher channel powers and denser packing of channels also led to the realization of the importance of effects such as spectral hole-burning (SHB) and nonlinearities in EDF. These effects can degrade the signal channel performance and therefore must to be taken into account in the design of EDFAs. Higher performance requirements led the growth of EDFAs from a simple gain block to a multifunctional element consisting of multiple stages. Additional features such as ASE filters and gain equalization filters, gain tilt control using variable optical attenuators, and dispersion compensation are incorporated in the midsection of the amplifier.

A. WIDEBAND GAIN-EQUALIZED AMPLIFIERS

First-generation WDM systems utilized 8–10 nm of spectrum to transmit 8–16 channels between 1540 and 1560 nm, where the gain of EDFAs is quite uniform. Here, the amplifier is operated at an inversion level of 70–80% (Fig. 2). Another amplifier based on erbium-doped fluoride fiber (EDFFA) consisting of fluorozirconate was shown to have greater (24 nm) bandwidth [17,18]. Unlike silica-based EDFAs, however, fluoride-based fiber is not a field-tested technology and there are concerns about its long-term reliability. Availability of gain equalization filters (GEFs) provides a way to increase the usable bandwidth in silica-based EDFAs. Several technologies have been studied to fabricate GEFs, including thin-film filters, long-period gratings [19], short-period gratings [20], silica waveguide structure [21], fused fibers, and acoustic filters [22].

Demonstration of a gain-equalized EDFA with 40 nm bandwidth [23] and a transmission experiment utilizing 35 nm gain-flattened EDFAs [15] clearly established EDFAs as the choice for high-capacity optical communication systems. In the system experiment, transmission of 32 WDM 10 Gb/s channels with 100 GHz spacing over 640 km was demonstrated [15]. The design of EDFAs used in this work is shown in Fig. 4 and consists of two stages in a midstage pumping configuration [9]. The amplifier design gave the first system demonstration of two features crucial for broadband, long-haul systems and optical networks. First, the gain spectrum of EDFAs was equalized to provide flat gain over the wide optical bandwidth of the erbium gain spectrum.

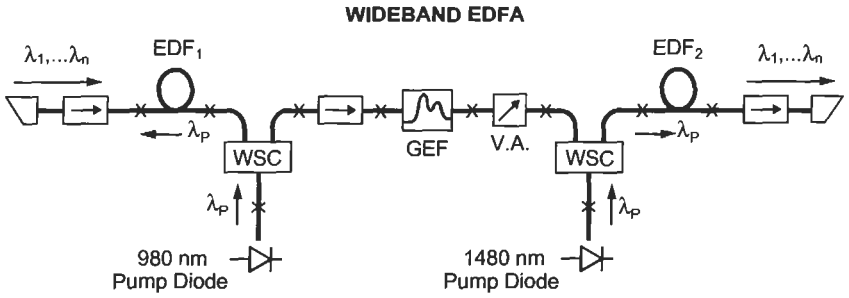


Fig. 4 A two-stage EDFA design with a midstage pumping configuration.

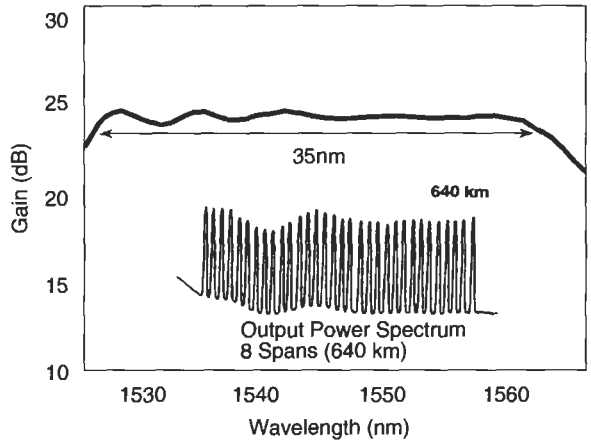


Fig. 5 Gain spectrum from the EDFA of Fig. 3. The gain is flattened with a long-period fiber grating equalization filter. The inset shows the channel power spectrum after transmission through 8 × 80 km spans and 8 EDFAs.

The gain spectrum of the amplifier using a long-period fiber grating equalizer filter is shown in Fig. 5. The gain spectrum shows a peak-to-peak gain variation over a 34 nm bandwidth of less than 0.6 dB or 2.5% of the gain. The inset shows output channel spectrum after 8 EDFAs. Channel power variation of less than 5 dB was recorded after 8 EDFAs. Second, with a midamplifier attenuator [16] the gain-flattened EDFA can be operated with this broad optical bandwidth in systems with a wide range of span losses. The attenuator can be adjusted to permit broadband, flat-gain operation for a wide range of gains, which is necessary to accommodate variations in span losses commonly encountered in practical transmission systems and multiwavelength optical networks. A discussion on the role of attenuators is given later this section.

Spectral Hole Burning in EDFAs and Its System Impact

In lightwave transmission applications EDFAs are operated in saturation mode. The gain saturation in EDFAs is predominantly homogeneous, which means that in a multichannel WDM system, once the gain of one of the channels is known, the gain of other channels can be calculated directly. This result comes from the homogeneous property of the EDFA model. While the gain spectrum of EDFAs is predominantly homogeneous, a small amount of inhomogeneity has been observed [24–26]. The inhomogeneous broadening gives rise to spectral hole burning (SHB) in the gain spectra of optical amplifiers. Using difference measurement technique, the SHB in EDFAs has been measured at room temperature [27]. The result of SHB measurement for different saturation levels is shown in Fig. 6. The figure shows the existence of a spectral hole having FWHM of 8 nm. The depth of the hole increases linearly at a rate of 0.027 dB per 1 dB increase in gain compression relative to small signal gain. For 10 dB gain compression a dip of 0.28 dB in the gain spectra due to SHB is observed. The SHB is strongly dependent upon the wavelength and has been shown to be four times larger at 1532 nm than at 1551 nm [28]. The dependence of the spectral hole width on the saturating wavelength is shown in Fig. 7. The FWHM of the hole increases as the saturating wavelength is increased.

The SHB effect impacts the gain shape of the long-haul optical transmission systems. The effect manifests itself such that each WDM channel in the system reduces the gain of the neighboring channels within the spectral hole

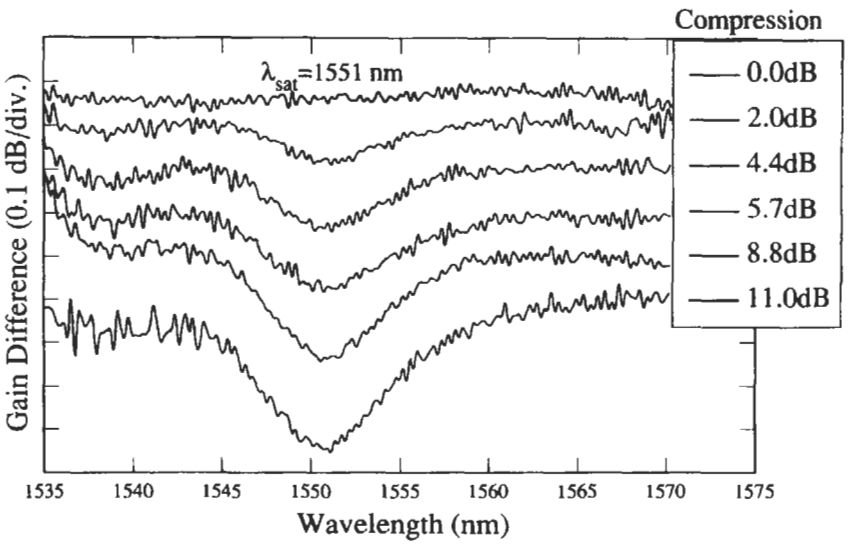


Fig. 6 SHB for different saturation levels [27].

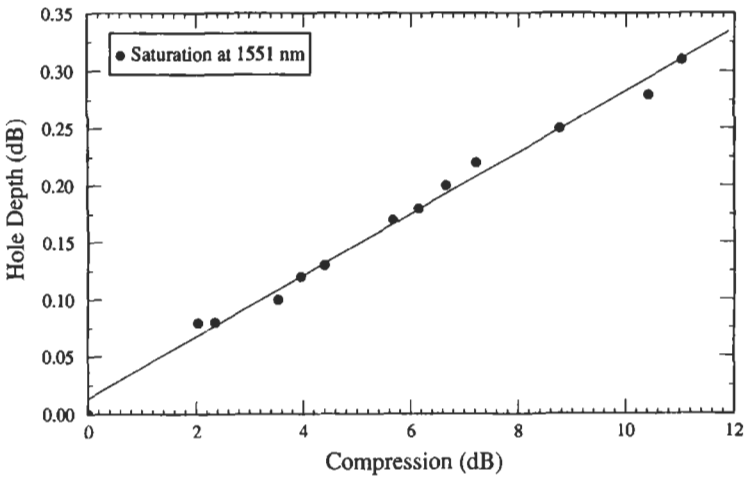


Fig. 7 Dependence of the spectral hole width on the saturating wavelength [27].

width but does not significantly affect channels far removed in wavelength. While characterizing the gain spectra of the amplifiers with full channel loading, it is therefore important that multiwavelength input signal with channel separation less than the SHB width be employed. The SHB effect observed in an individual amplifier is small, but in a long chain of amplifiers, such as in a long-haul or submarine system, it can add up to a significant observable change in the overall spectrum. The importance of SHB was noted in long-haul transmissions over 9300 km [29]. The SHB impacts a WDM system in a positive way since it helps in the mitigation of channel power divergence and should be included in the system design.

Midstage Attenuators for Dynamic Range and Tilt Control

In order to support the growth in the number of channels in WDM transmission systems, optical amplifiers with wider bandwidth are required. The gain of amplifiers must be very uniform over the entire WDM transmission bandwidth for the channels to be transmitted without impairments due to either nonlinear effects in fiber or due to poor OSNR at the receiver. The power spectrum tilt in wideband systems can arise from several reasons, such as EDFA gain tilt, spectral loss in transmission fiber, dispersion compensation fiber or other passive components, variation in input signal power due to uneven fiber span loss, and Raman effect.

In a transmission system, the wide bandwidth of the amplifiers has to be maintained while accommodating the variations in the losses of fiber spans deployed in the field. The midstage attenuator provides a control of the gain flatness of the amplifier over a wide range of variations in span loss. This is

achieved by maintaining the average inversion level of EDF constant by changing the attenuator loss. The use of the attenuators, however, raises the concern about the increase of EDFA noise figure and therefore the end-of-system OSNR degradation. The effect of an attenuator in the midsection of a two-stage EDFA on gain flatness and end-of-system OSNR has been investigated in an 8 × 80 km transmission system. In the experiment, 18 WDM channels with 200 GHz separation were transmitted through the chain of amplifiers and end-of-system channel spectrum was measured. The result of changes in the channel spectrum when the span loss changes from 24 to 6 dB is shown in Fig. 8. The gain flatness can be maintained by increasing the midstage attenuator loss. The end-of-system SNR (Fig. 9), however, shows an increase initially with the decrease in span loss, since the first stage of EDFA receives a larger signal, and less ASE is generated. At smaller span loss the larger midstage attenuator loss causes SNR degradation. The midstage attenuator provides a dynamic range of 12 dB over which the SNR is not degraded.

The gain tilt control is very important for the operation of WDM systems and networks. The spectrum of WDM channels after transmission through fiber could acquire positive linear tilt due to the Raman effect, which will lead to the transfer of power from shorter to longer wavelength channels. Alternatively, it may be desirable to have a negative tilt in the spectrum of channels at the output of the amplifier in order to compensate for the Raman effect in the transmission fiber. Both of these conditions can be achieved by controlling the gain tilt in the EDFA. The EDFA gain spectrum acquires a tilt when the average inversion level of the erbium fiber is changed, as is evident from Fig. 2. By adding a wavelength-independent variable loss element such as an attenuator, the average inversion level of the amplifier can be controlled,

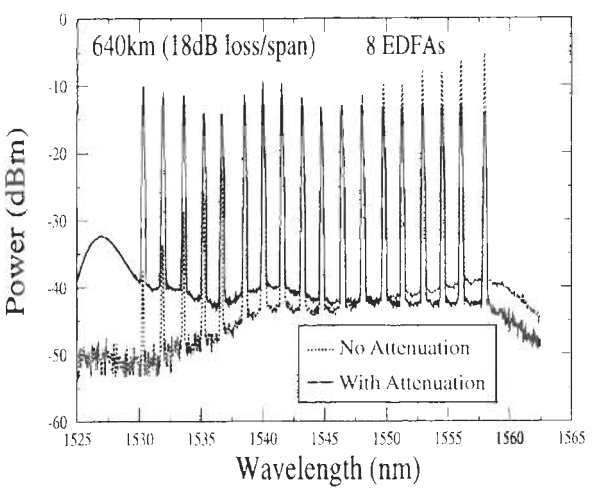


Fig. 8 Spectrum variation as the span loss changes from 24 to 6 dB.

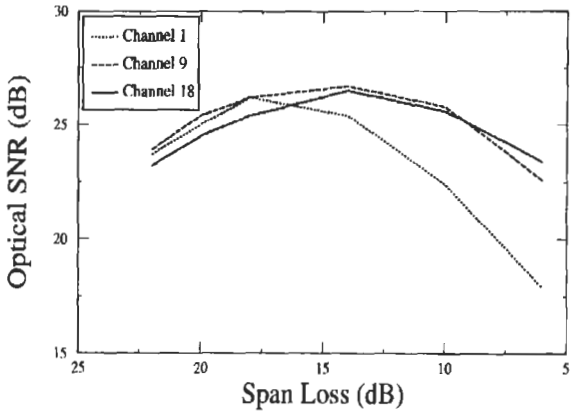


Fig. 9 End of system SNR for different span losses.

in turn regulating the tilt in the gain spectrum. It is therefore possible to mitigate both the positive and negative spectral slope by increasing or decreasing the midstage attenuator loss, respectively. The increase in midstage attenuator loss may lead to reduction in the amplifier output power and increase in the noise figure.

B. L-BAND EDFAs

In order to expand the optical bandwidth usage per fiber, the WDM systems have to be expanded beyond the conventional band or C-band (1525–1565 nm). The realization of EDFAs in the longer wavelength region [14,30] or L-band (1570–1610 nm) has doubled the usable transmission bandwidth. In addition to capacity, L-band EDFAs enable WDM system operation over different types of dispersion-shifted fiber (DSF) having low dispersion in the C-band. This is very significant since the deployment of nonzero dispersion-shifted fiber (NZDSF) now exceeds that of SMF. Dense WDM transmission over DSF/NZDSF was previously not feasible in the 1550 nm region due to the low values of dispersion, which result in unacceptably high levels of four-wave mixing. The dispersion in fiber increases at longer wavelength, and consequently the levels of FWM are reduced in the L-band.

In the long wavelength region, the EDF has nearly 0.2 dB/m gain coefficient for the inversion level between 20 and 30%. Even though at this inversion level the gain is much smaller than the gain at the highest peak in the C-band at 1532 nm, the gain shape in the L-band is much more uniform as compared to that in the C-band. This means that much less gain filtering is required in L-band EDFA. A comparison of the EDF gain spectra of C- and L-bands is shown in Fig. 10. The inversion level in EDF is optimized to provide the most bandwidth in both the cases. In order to achieve low-noise operation, high

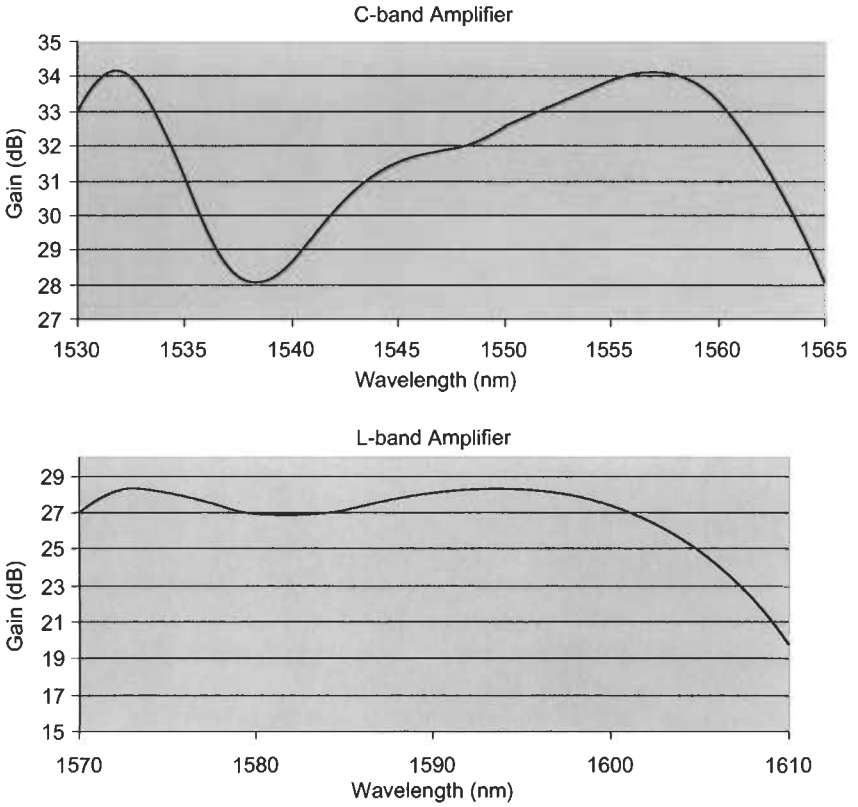


Fig. 10 A comparison of C- and L-band gain spectra, top, inversion 65%; bottom, inversion 38%.

gain in the first stage of EDFA is required. Due to a small gain coefficient, the EDF lengths exceeding 5 times that in a C-band EDFA are therefore necessary. Unfortunately, longer EDF length leads to larger background loss, which is detrimental to noise figure. Recently, EDF optimized for low-inversion L-band operation has been developed which can provide gain coefficient as high as 0.6 dB/m [31]. In new EDFs, larger overlap between the mode-field of the signals and the ion-doped core area is needed to increase the absorption without producing concentration-quenching effect. The new EDF is designed to have a longer cutoff wavelength around 1450 nm and has 2–3 times greater power overlap. The greater cutoff wavelength leads to smaller mode field diameter, enabling better than 90 percent optical power confinement in the core area. The new fiber has other advantages such as lower background loss, increased tolerance to fiber bend, and higher pump efficiency [31]. In this type of fiber power efficiency as high as 60 percent has been measured, and 5 mm bend radius did not generate additional loss.

Several L-band transmission experiments at 10 Gb/s and higher rates in recent years [32–36] have demonstrated the feasibility of L-band EDFAs and WDM transmission over DSF/NZDSF. The first 64 × 10 Gb/s WDM transmission [35] with 50 GHz channel spacing over DSF verified that nonlinear effects such as four-photon mixing can be controlled in the L-band. Subsequently a 1 Tb/s capacity transmission experiment with 25 GHz channel spacing over NZDSF [36] showed the feasibility of L-band ultradense WDM systems.

Nonlinearities in EDFAs

EDFAs operating in the C-band regime are designed around high erbium-fiber gains in the 1525–1565 nm range and are typically shorter than 50 m. These lengths are generally too short to generate intra-amplifier nonlinear effects that might lead to WDM system impairments. However, as discussed previously, EDF provides much lower gain per unit length in the L-band window, typically requiring up to five times longer lengths than those in the C-band. Long erbium fiber length, combined with high internal optical intensities, naturally increases nonlinear distortion of optical signals. Recent work has suggested that two nonlinear effects, cross-phase modulation and four-photon mixing, produced in EDF in WDM systems may be comparable to those produced from transmission over fiber. The result of these studies is summarized below.

There are several factors that affect the importance of combined XPM from EDFAs. Although a fiber transmission span is more than two orders of magnitude longer than an amplifier, most of the effect from XPM is produced in only a short section (a few kilometres) at the start of the span over a distance where the walkoff between channels is small [37]. Secondly, the nonlinearity of EDF is larger than that of conventional transmission fiber due to its smaller effective area. Lastly, in amplifiers the XPM increases with the total number of channels, irrespective of spacing, since the amplifier length is much shorter than the walkoff length, whereas in transmission fiber, the XPM is determined by only a few immediate neighboring channels, since the fiber span is much longer than the walkoff length [38]. Measurement of XPM in L-band EDFAs was carried out using two WDM channels; the results showed that XPM has the potential of becoming the dominant nonlinear crosstalk mechanism in L-band WDM systems using standard single-mode fiber [39]. In another experiment, a comparison of XPM arising from L-band EDFA and that from two types of transmission fibers (SMF and DSF) was compared. In this measurement 20 WDM channels spaced by 100 GHz were used; the results showed that the XPM produced in the L-band amplifier is negligible compared to that from DSF and a factor of nine lower than that from conventional fiber. The study [40] concluded that even though higher levels of XPM may be possible, for WDM systems with more channels or with different EDFAs it is not necessarily a problem for WDM systems.

Measurement and analysis of four-wave mixing (FWM) in L-EDFAs and its impact on DWDM link performance has been studied using both time-domain and spectral measurements [41]. Significant magnitude of FWM in L-EDFA was observed with worst observed case results in 1.5 dB eye penalty for 50 GHz spaced WDM channels and FWM-generated harmonics 25 dB below the signal level. The FWM level generated in L-EDFA is strongly dependent upon EDF parameters: dispersion, effective area, and length. The other important characteristic is the gain evolution function governed by L-EDFA topology [42].

C. ULTRAWIDEBAND WITH A SPLIT-BAND ARCHITECTURE

In order to increase the overall bandwidth of EDFAs beyond the C-band, a novel architecture combining the C- and L-bands has been demonstrated. Since the gain drops sharply on both sides of the C-band at a 70–80% inversion, it is not practical to further increase the bandwidth with a GEF. However, a flat gain region between 1565 and 1615 (L-band) can be obtained at a lower inversion level (30 to 40%) [43,44]. The principle of combined C- and L-band amplifiers with very wide bandwidth is shown in Fig. 11.

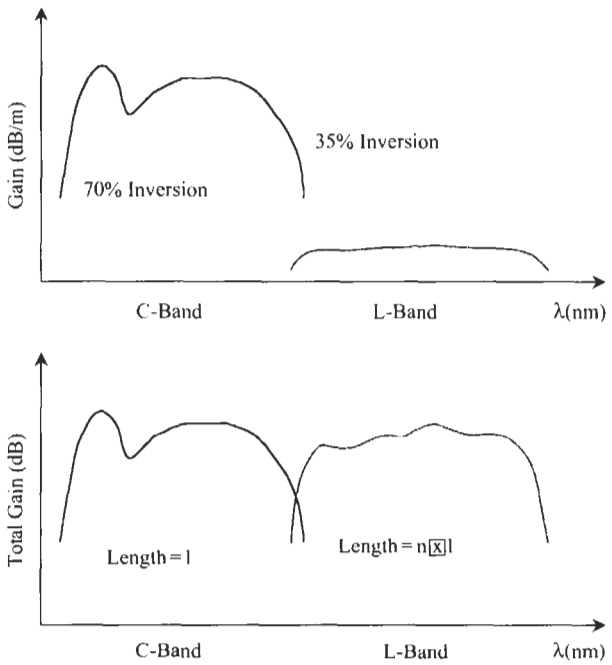


Fig. 11 Principle of combined C- and L-band ultrawideband amplifiers [14].

After the initial demonstration of principle [44–46], much progress has been made on the understanding and design of ultrawideband optical amplifiers with a split-band structure. A recent design is shown in Fig. 12(a), utilized fiber grating filters to split and combine the C- and L-band signals. The two arms of the amplifier could be optimized separately to provide flat gain operation in the two bands. The total bandwidth of the combined EDFA was 84.3 nm covering 1526–1612 nm with less than 2 nm guard band in between, as shown in Figure 12(b) [47] with total output power 25 dBm. The amplifier noise figure was less than 6 dB over the whole bandwidth. Besides wide bandwidth and low noise figure, this amplifier also provides power tilt control, which is realized by the variable attenuation, and dispersion compensation, which is realized by

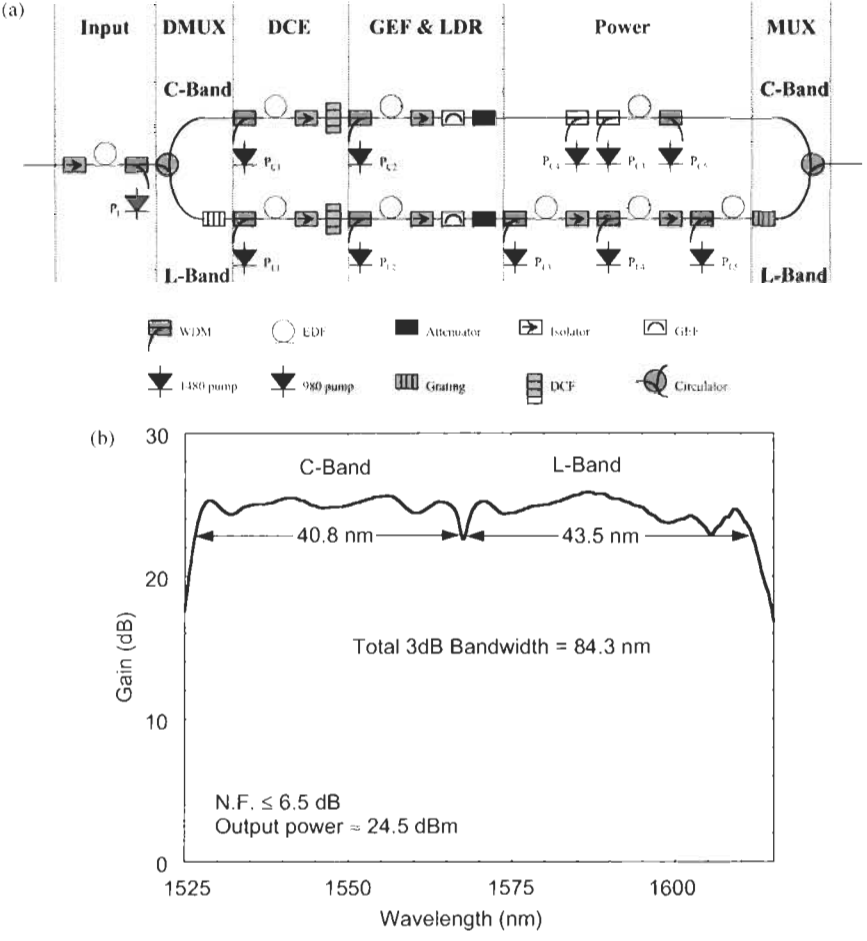


Fig. 12 (a) Schematic of ultrawideband amplifier. See also Plate 2. (b) Gain spectrum of ultrawideband amplifier [49].

the dispersion compensation element after the second gain stages. Dispersion compensation is needed for high-speed WDM channels and can be done with dispersion compensating fiber. Recently, dispersion compensation using fiber gratings has also been reported [48]. Since there is a significant mismatch in the dispersion slopes in transmission and dispersion compensation fibers, the split-band architecture provides an opportunity for more accurate dispersion compensation in C- and L-bands separately. Another advantage is significantly reduced crosstalk between the two bands due to double rejection at the splitting and combining stages. With ultra wideband optical amplifier, the first long-distance WDM transmission at 1 terabit per second was demonstrated in early 1998 [49,50].

D. RAMAN/EDFAs

The Raman effect in silica fiber has been intensively investigated in recent years. Stimulated Raman scattering transfers energy from the pump light to the signal via the excitation of vibrational modes in the constituent material. Measurement of Raman gain coefficient in silica fiber [51] has revealed that a significant amount of gain can be obtained at moderately high pump powers. The Raman gain is given by $G_r = \exp(P_p/2L_{eff}A_{eff})$, where P_p, L_{eff}, and A_{eff} are the pump power, effective length, and effective area, respectively. The Raman gain peak is offset by one Stokes shift in wavelength from the pump signal; thus Raman amplifiers can be implemented at any wavelength by selecting a suitable pump signal wavelength. The Stokes shift for silica fibers is approximately 13 THz, which corresponds to nearly 100 nm at 1550 nm. Raman gain spectrum is fairly uniform and has a 3 dB bandwidth of about 5 THz , corresponding to 40 nm in the C-band.

Low-noise Raman amplification can be applied to enhance the system margin in WDM transmission systems. Unlike the EDFA, which requires a certain pump power to maintain high inversion level for low-noise operation, the Raman amplifier can be inverted regardless of the pump level, since the absorption of the signal photon to the upper virtual state is extremely small. The other significant advantage of Raman amplifiers is the ability to provide distributed gain in transmission fiber. The Raman gain, distributed over tens of kilometers of transmission fiber, effectively reduces the loss L of the fiber span and results in superior end-of-system OSNR, see Section IV A. Demand for higher capacity and longer reach systems coupled with the availability of high-power pumps in the 1450 and 1480 nm wavelength regions have enabled the application of Raman amplifiers to WDM transmission systems. In addition, large deployment of smaller-core NZDSF in terrestrial networks has made it possible to obtain significant Raman gain with standard pump diodes. Several WDM transmission experiments have been reported, which have used distributed Raman amplification in the transmission fiber to enhance system OSNR. The Raman pumping is normally implemented in a

counterpropagating configuration in order to avoid noise transfer from the Raman pump to the WDM signals. The counterpropagating pump configuration also efficiently suppresses any signal-pump-signal crosstalk that may occur if the Raman pump is depleted by the WDM channels.

The gain coefficient in the Raman amplifier is quite small and as a result gain media consisting of tens of kilometers of fiber are needed. Since the power conversion efficiency in Raman amplifiers is ~10%, which is several times smaller compared to that in EDFAs (>60%), the most attractive design for low noise figure amplifiers is a hybrid configuration consisting of a Raman preamplifier stage followed by an EDFA power stage. In such a design, the gain of the Raman stage is kept below 16 dB in order to minimize intersymbol interference arising from amplified double reflections of the signal from either discrete reflection points or from double Rayleigh scattering [52]. Noise figure improvement of 3–4 dB in such a hybrid design over the EDFA counterpart has been demonstrated [52,53].

The enhanced margin derived from superior noise performance of Raman/EDFAs can be utilized in several ways, such as to increase the separation between amplifiers, to increase the overall reach of the transmission system, and to increase the spectral efficiency of transmission by reducing the channel separation or increasing the bit rate per channel. For example, a decrease in channel spacing to obtain higher spectral efficiency requires a reduction in launch power to avoid increased penalties from fiber nonlinearities and must be accompanied by a noise figure reduction to maintain optical signal-to-noise ratio (OSNR). Likewise, an increase in line rate requires a reduction in span noise figure to increase the OSNR at the receiver accordingly. Published results qualitatively confirm the outlined relation between channel spacing and noise figure for WDM systems limited by four-wave mixing [54,55]. In addition to superior noise performance, Raman amplification can provide gain in spectral regions beyond the C- and L-bands. As mentioned earlier, the Raman gain curve is intrinsically quite uniform, and broader gain spectra are naturally achievable in Raman amplifiers with gain-flattening filters. Alternatively, wide gain spectra may be obtained by the use of multiwavelength pump sources [56,57], a technique applicable only to Raman amplifiers.

Three transmission experiments are described below which show the benefits of incorporating Raman/EDFAs in WDM systems. The first experiment demonstrates that by the use of distributed Raman gain in a multispan system, the system margin can be enhanced by 4–5 dB. The second experiment shows that terabit capacity ultradense WDM transmission with 25 GHz channel spacing is made possible by the use of Raman gain, since the launched power per channel can be lowered and nonlinear effects can be minimized. Finally a transmission experiment at 40 Gb/s line rate covering C- and L-bands with overall capacity of 3.2 Tb/s is demonstrated by the use of Raman/EDFAs. More than 6 dB reduction in span noise obtained by distributed Raman amplification is expected to become essential in WDM transmission systems operating at

40 Gb/s. The improvement makes up for the 6 dB higher OSNR requirements for 40 Gb/s signals compared to a 10 Gb/s signal. Thus, the amplifier spacing used in today's 10 Gb/s WDM systems can be accommodated at 40 Gb/s by incorporating distributed Raman amplification.

A schematic of the first experiment [53] is shown in Fig. 13. 40 WDM channels modulated at 10 Gb/s in the L-band with 50 GHz channel spacing were transmitted. Raman gain was provided at the input of each inline amplifier by incorporating two polarization-multiplexed 1480 nm pump lasers with a wavelength selective coupler (shown in the lower section of Fig. 13). The DSF spans acted as the gain medium with the signal and pump propagating in the opposite directions. The total Raman pump power in the fiber was 23.5 dBm, which resulted in a peak gain of 12 dB at 1585 nm. The total power launched into fiber spans was ~15 dBm or −1 dBm/ch. The transmission span consisted of five 120 km lengths of dispersion-shifted fiber. The eye diagrams of all channels after transmission through 600 km DSF are open and exhibit little distortion. All channels achieved error rates below 10^{-9}. The power penalty was between 2 and 4 dB for 10^{-9} BER. Without the use of Raman amplification, span lengths were restricted to 100 km length for similar BER performance. Thus the addition of Raman gain allows nearly 4 dB of additional span loss for error-free transmission.

In the second experiment, error-free transmission of 25 GHz spaced 100 WDM channels at 10 Gb/s over 400 km of NZDS fiber was reported [36].

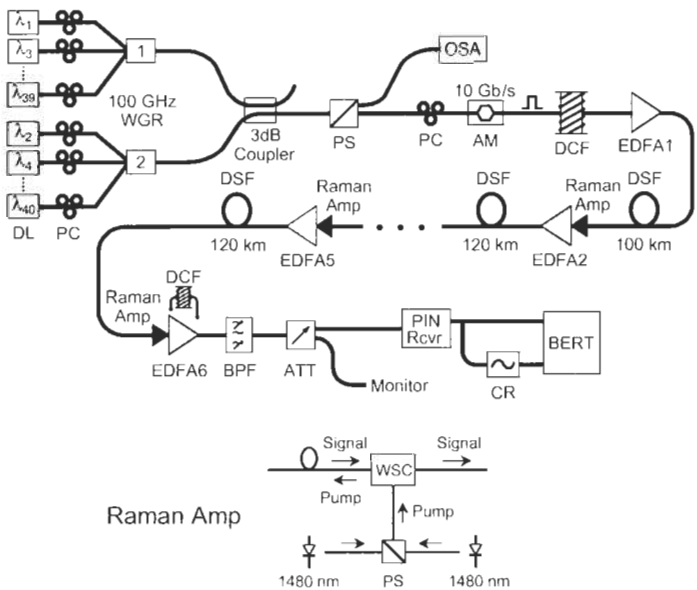


Fig. 13 A schematic of multispan Raman gain-enhanced transmission experiment [53].

High spectral efficiency of 0.4 b/s/Hz was achieved by the use of distributed Raman gain. Raman pump power in the fiber was 22–23 dBm, which resulted in a peak gain of 10 dB at 1585 nm. Three fiber spans of positive NZDS fiber having zero dispersion wavelength in the range 1508 to 1527 nm and lengths 125, 132, and 140 km, respectively, were used. The channel spectrum at the beginning and at the end of system is shown in Fig. 14. The eye diagrams of all the channels were open and, as expected, exhibited little distortion due to non-linear effects. All channels achieved error rates below 10^{-9}. The power penalty at 10^{-9} BER was between 2.1 and 5.3 dB. The penalty can be attributed to the variation of OSNR due to low-power lasers and to gain nonuniformity. Calculated value of OSNR 22 dB is in good agreement with the measured 21–25 dB.

Distributed Raman amplification was employed in two recent 40 Gb/s WDM transmission experiments, both achieving a spectral efficiency of 0.4 bit/s/Hz. In both experiments, distributed Raman amplification allowed the launch power into 100 km spans of TrueWave fiber to be as low as −1 dBm/channel while maintaining sufficient optical signal-to-noise ratio at the output of the system for error-free performance of the 40 Gb/s WDM channels. In the first experiment, 40 WDM channels were transmitted in the C-band over four 100 km spans of TrueWave fiber, using four hybrid Raman/erbium inline amplifiers [58]. A single Raman pump wavelength was sufficient to provide effective noise figures below 0 dB over the entire C-band, as shown

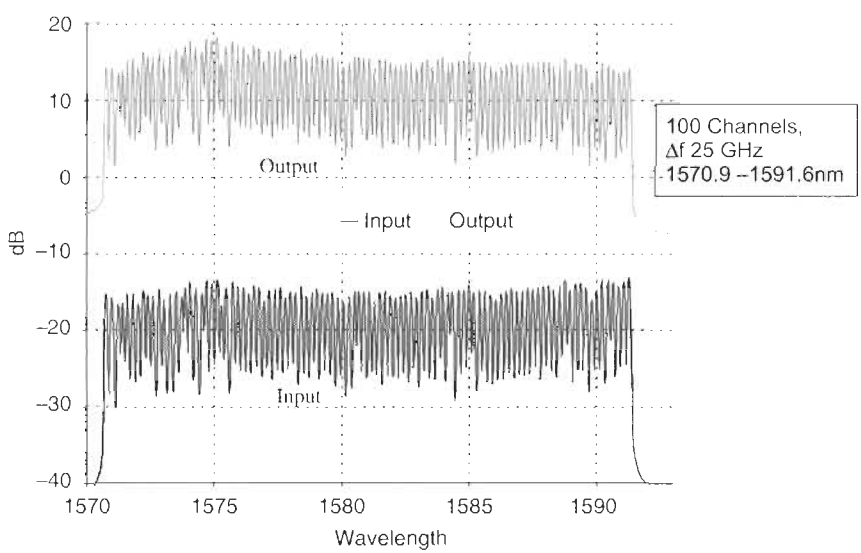


Fig. 14 Channel spectra of ultradense (25 GHz spaced) terabit capacity (100 × 10 Gb/s) WDM transmission [36].

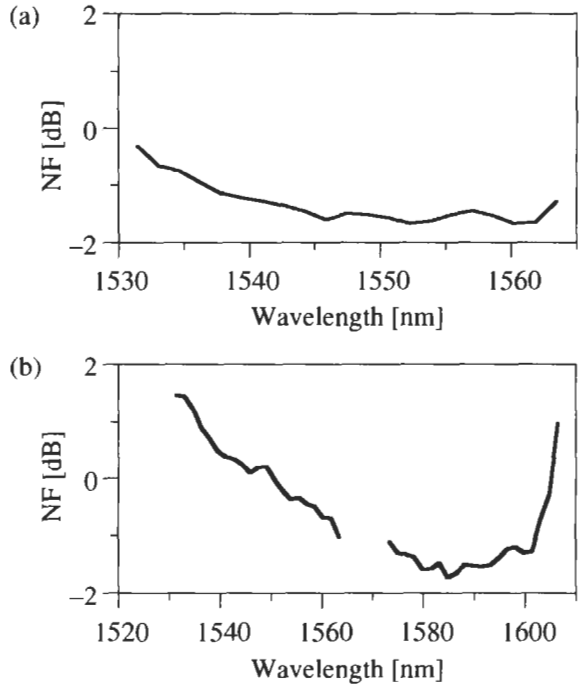


Fig. 15 (a) Equivalent noise figure of C-band Raman EDFA [58]. (b) Equivalent noise figure of a combined C- and L-band Raman/EDFA [59].

in Fig. 15(a). For dual C- and L-band systems, at least two pump wavelengths are required to obtain adequate noise figures over the combined C- and L-bands. In the second experiment, 3.28 Tb/s were transmitted over three 100-km spans of nonzero dispersion-shifted fiber [59]. Two Raman pumps were used to achieve noise figures ranging from +1.5 dB in the lower end of the C-band to −1.7 dB in the L-band, as shown in Fig. 15(b). The two experiments illustrate one important complication that arises from using multiple Raman pumps in ultra broadband systems, namely that Raman pumps at lower wavelengths will be depleted by pumps at higher wavelengths. Preemphasis of the pumps must be used, and the noise figure of the WDM channels being pumped by the shorter wavelength Raman pump generally will be higher compared to the noise figure of the WDM channels being pumped by the longer wavelength Raman pump.

IV. EDFAs for Dynamic WDM Networks

Erbium-doped fiber amplifiers are employed in the current multiwavelength optical networks to compensate for the loss of fiber spans and network

elements. The amplifiers are normally operated in a saturated mode in these applications. In the event of either network reconfiguration or a failure, the number of WDM signals traversing the amplifiers would change and the power of surviving channels would increase or decrease due to the cross-saturation effect in the amplifiers. The quality of surviving channels can be severely affected through four mechanisms when channel loading changes. First, nonlinear optical effects in transmission fibers will occur if the power excursions are large enough when signal channels are lost. Self-phase modulation (SPM) has been observed to affect the performance of the surviving channels [60]. Second, when channels are added, the optical power at the receiver can be reduced during the transient period, which would cause eye closure. If the optical signal power at the receiver is lowered by more than the system margin, i.e., if the signal power is lower than the receiver sensitivity, bit-error rate would be severely degraded. Third, optical SNR maybe degraded due to the change of inversion level and therefore the change in gain spectrum during the transient period. Fourth, the received power at the receiver varies, requiring that the threshold of the receiver be optimized at high speed, which can be a problem for certain receivers.

A. GAIN DYNAMICS OF SINGLE EDFAs

The speed of gain dynamics in a single EDFA is in general much faster than the spontaneous lifetime (10 ms) [61] because of the gain saturation effect. The time constant of gain recovery on single-stage amplifiers was measured to be between 110 and 340 μs [62]. The time constant of gain dynamics is a function of the saturation caused by the pump power and the signal power. Present-day WDM systems with 40–100 channels require high-power EDFAs in which the saturation factor becomes higher leading to shorter transient time constants. In a recent report, the characteristic transient times were reported to be tens of microseconds in a two-stage EDFA [61]. The transient behavior of surviving channel power for the cases of one, four, and seven dropped channels, in an eight-channel system, is shown in Fig. 16. In the case of seven dropped channels the transient time constant is nearly 52 μs. As can be seen, the transient becomes faster as the number of dropped channels decreases. The time constant decreases to 29 μs when only one out of 8 channels is dropped. The rate equations [63] for the photons and the populations of the upper ($^4I_{13/2}$) and lower ($^4I_{15/2}$) states can be used to derive the following approximate formula for the power transient behavior [64]:

$$P(t) = P(\infty)\,[P(0)/P(\infty)]^{\exp(-t/\tau_e)}$$

where $P(0)$ and $P(\infty)$ are the optical powers at time $t = 0$ and $t = \infty$, respectively. The characteristic time τ_e is the effective decay time of the upper level averaged over the fiber length. It is used as a fitting parameter to obtain best fit with the experimental data. The experimental data (Fig. 16) are in good

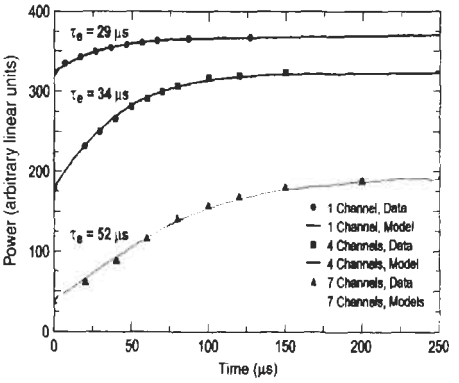


Fig. 16 Measured and calculated surviving power transients for the cases of 1, 4, 7 channels dropped out of 8 WDM channels.

agreement with the model for the transient response. The model has been used to calculate the fractional power excursions in decibels of the surviving channels for the cases of one, four, and seven dropped channels. The times required to limit the power excursion to 1 dB are 18 and 8 μs, respectively, when four or seven channels are dropped. As EDFAs advance further to support larger numbers of WDM channels in lightwave networks, the transient times may fall below 10 μs. Dynamic gain control of the EDFAs with faster response times will be necessary to control the signal power transients.

A model of EDFA dynamics is needed to understand the transient behavior in large systems or networks. Recently a simple model has been developed for characterizing the dynamic gain of an EDFA. The time-dependent gain is described by a single ordinary differential equation for an EDFA with an arbitrary number of signal channels with arbitrary power levels and propagation directions. Most previous EDFA models are represented by sets of coupled partial differential equations [65,66], which can be solved only through iterative, computationally intensive numerical calculations, especially for multichannel WDM systems with counterpropagating pump or signals. The mathematical details of the model are provided in [67]. Here, the simulation results from the model are compared with the measured time-dependent power excursions of surviving channels when one or more input channels to an EDFA are dropped. The structure of the two-stage EDFA used in the experiment [61] and simulation is shown in the inset of Fig. 17. The experimentally measured power of the surviving channel when 1, 4, and 7 out of 8 WDM channels are dropped are plotted. It is seen from the figure that the simulation results agree reasonably well with the experimental data without any fitting parameters. The exception is a 0.9 dB difference at large t for the 7 channel drop case. This discrepancy is believed to arise from pump-excited state absorption at high pump intensity. The model can be very useful in the study of power transients in amplified optical networks.

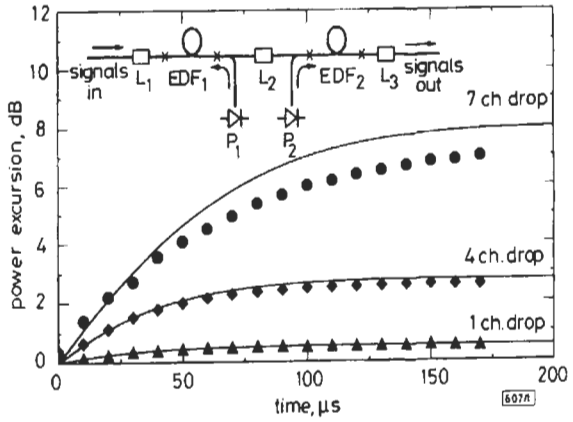


Fig. 17 Comparison between theory and experiment for output power excursions of surviving channels from a two-stage amplifier when 1, 4, 7 of input channels are dropped.

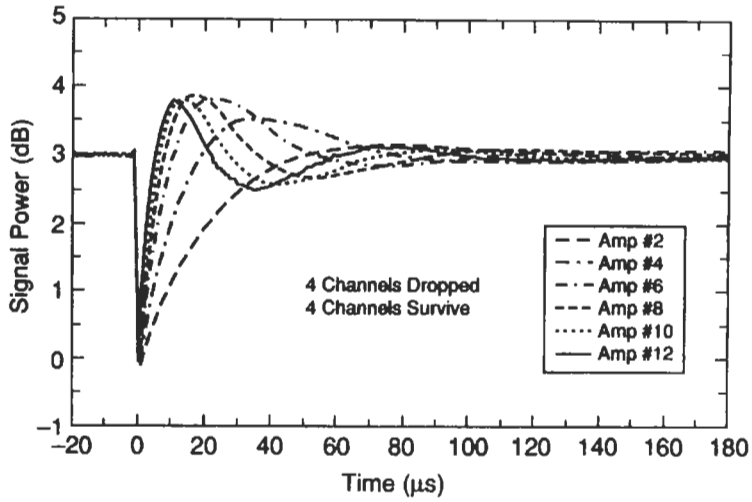


Fig. 18 Measured output power as a function of time after 0, 2, 4, 6, 8, 10, and 12 EDFAs (at time $t = 0$, 4 out of 8 WDM channels are dropped).

B. FAST POWER TRANSIENTS IN EDFA CHAINS

In a recent work, the phenomenon of fast power transients in EDFA chains was reported [68,69]. The effect of dropped channels on surviving power in an amplifier chain is illustrated in Fig. 18. When 4 out of 8 WDM channels are suddenly lost, the output power of each EDFA in the chain drops by 3 dB, and the power in each surviving channel then increases toward double

the original channel power to conserve the saturated amplifier output power. Even though the gain dynamics of an individual EDFA is unchanged, the increase in channel power at the end of the system becomes faster for longer amplifier chains. Fast power transients result from the effects of the collective behavior in chains of amplifiers. The output of the first EDFA attenuated by the fiber span loss acts as the input to the second EDFA. Since both the output of the first EDFA and the gain of the second EDFA increase with time, the output power of the second amplifier increases at a faster rate. This cascading effect results in faster and faster transients as the number of amplifiers increase in the chain. To prevent performance penalties in a large-scale WDM optical network, surviving channel power excursions must be limited to certain values depending on the system margin. Taking the MONET network as an example, the power swing should be within 0.5 dB when channels are added and 2 dB when channels are dropped [70]. In a chain consisting of 10 amplifiers, the response times required in order to limit the power excursions to 0.5 dB and 2 dB would be 0.85 and 3.75 μs, respectively. The response times are inversely proportional to the number of EDFAs in the transmission system.

The time response of EDFAs can be divided into three regions—the initial perturbation region, the intermediate oscillation region, and the final steady-state region. In the initial perturbation region, the gain of the EDFA increases linearly with time, and the system gain and output power increases at a rate proportional to the number of EDFAs. The time delays for a channel power excursion of 2 dB (Fig. 18) and the inverse of time delays, i.e., the power transient slope in the perturbation region, are plotted in Fig. 19. Assuming that the amplifiers operate under identical conditions, the rate of change of gain at each EDFA is the same and is proportional to the total lost signal power.

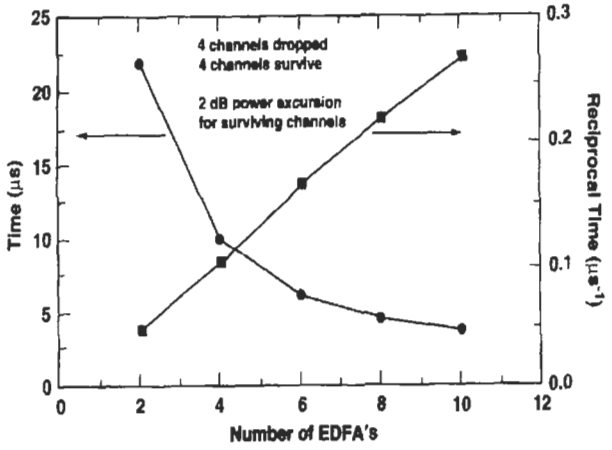


Fig. 19 Delay and reciprocal of delay for surviving channel power excursion to reach 2 dB after the loss of 4 out of 8 WDM channels.

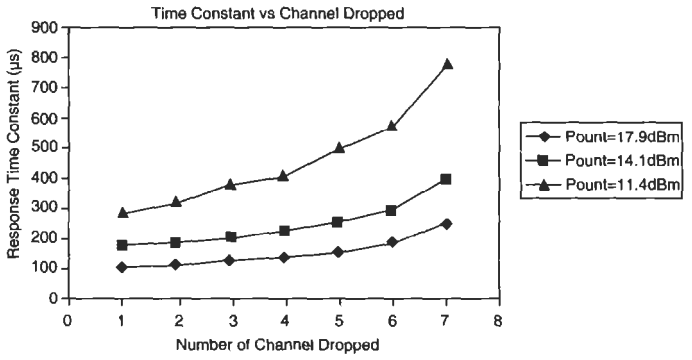


Fig. 20 Response time constant vs. the number of dropped channels under different saturation conditions in a two-stage L-band EDFA.

The slope plotted in Fig. 19, therefore, increases linearly with the number of EDFAs in the chain. These experimental results have been confirmed by modeling and numerical simulation from a dynamic model [6,67].

In the intermediate region, an overshoot spike can be observed after 2 EDFAs in Fig. 18. The first overshoot peak is the maximum power excursion, since the oscillation peaks that follow are smaller than the first one. From the results of both experimental measurements and numerical simulation on a system with N EDFAs, the time to reach the peak is found to be inversely proportional to N, and the slope to the peak is proportional to $N - 1$ [71]. This indicates that the overshoot peaks are bounded by a value determined by the dropped signal power and the operating condition of the EDFAs. These properties in the perturbation and oscillation regions can be used to predict the power excursions in large optical networks.

A study of dynamic behavior of L-band EDFA has been carried out recently [72]. In this work, transient response of the surviving channels in a two-stage L-band EDFA under different channel loading conditions was reported. The observed dynamic behavior in the L-band is similar to that in the C-band. However, the response time is very different. The response time constants as a function of the number of dropped channels under different saturation conditions is shown in Fig. 20. The time constants are about 105 μs and 260 μs when one and seven channels out of eight channels are dropped and the amplifier is well saturated. These values are about four to five times larger than that observed in a C-band EDFA. The difference can be explained by the different intrinsic saturation power in these two bands.

C. CHANNEL PROTECTION SCHEMES

As discussed earlier, channels in optical networks will suffer error bursts caused by signal power transients resulting from a line failure or a network

reconfiguration. Such error bursts in surviving channels represent a service impairment that is absent in electronically switched networks and is unacceptable to service providers. The speed of power transients resulting from channel loading, and therefore the speed required to protect against such error bursts, is proportional to the number of amplifiers in the network, and for large networks can be extremely fast. Several schemes to protect against the fast power transients in amplified networks have been demonstrated in recent years.

Pump Control

The gain of an EDFA can be controlled by adjusting its pump current. Early reported work addressed pump control on time scales of the spontaneous lifetime of EDFAs [65]. One of the studies demonstrated low-frequency feed forward compensation with a low-frequency control loop [62]. After the discovery of fast power transients, pump control on short time scales [73] was demonstrated to limit the power excursion of surviving channels. In the experiment, automatic pump control in a two-stage EDFA operating on a time scale of microseconds was demonstrated. The changes in the surviving channel power in the worst case of 7-channel drop/add in an 8-channel WDM system are shown in Fig. 21. In the absence of gain control, the change in surviving channel signal power exceeds 6 dB. When the pump control on both stages is active, the power excursion is less than 0.5 dB both for drop and add conditions. The control circuit acts to correct the pump power within 7–8 μs, and this effectively limits the surviving channel power excursion.

Link Control

The pump control scheme described above would require protection at every amplifier in the network. Another technique makes use of a control channel in the transmission band to control the gain of amplifiers. Earlier work demonstrated gain compensation in an EDFA at low frequencies (<1 kHz) using an

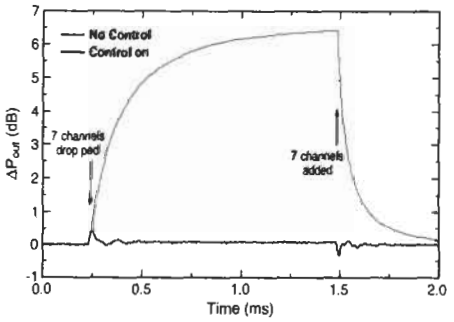


Fig. 21 Surviving channel power transient in two-stage EDFA when 7 out of 8 channels are dropped and added.

idle compensation signal [74]. Recently, link control, which provides surviving channel protection against fast transients, has been demonstrated [60]. The scheme, as illustrated in Fig. 22, protects surviving channels on a link-by-link basis. A control channel is added before the first optical amplifier in a link (commonly the output amplifier of a network element). The control channel is stripped off at the next network element (commonly after its input amplifier) to prevent improper loading of downstream links. The power of the control channel is adjusted to hold constant the total power of the signal channels and the control channel at the input of the first amplifier. This will maintain constant loading of all EDFAs in the link.

The experimental demonstration of link control surviving channel protection is illustrated in Fig. 23. The experiment is set up with 7 signal channels and 1 control channel. A fast feedback circuit with a response time of 4 μs is used to adjust the line control channel power to maintain constant total power. The signal and control channels are transmitted through seven amplified spans of fiber and the bit-error rate performance of one of the signal channels is monitored. The measured results are summarized in Fig. 24. When 5 out of 7 signal channels are added/dropped at a rate of 1 kHz, the surviving channel suffers a power penalty exceeding 2 dB and a severe BER floor. An even worse BER floor is observed when 5 out of 6 channels are added/dropped, resulting from cross-saturation induced by change in channel loading. With the fast link control in operation, the power excursions are mitigated, BER penalties are reduced to a few tenths of a dB, and error floors disappear.

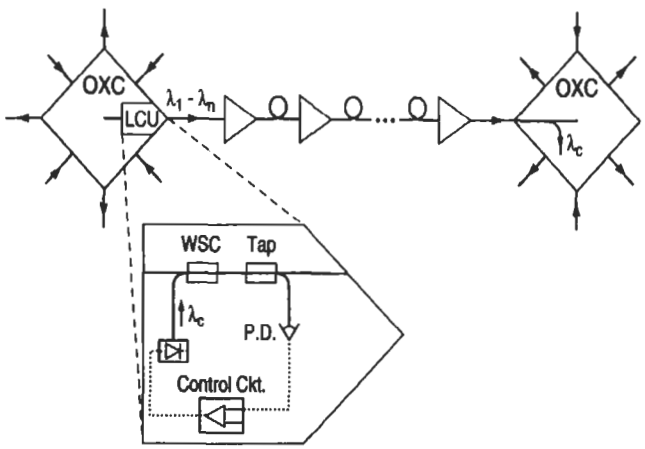


Fig. 22 Schematic representation of link control for surviving channel protection in optical networks.

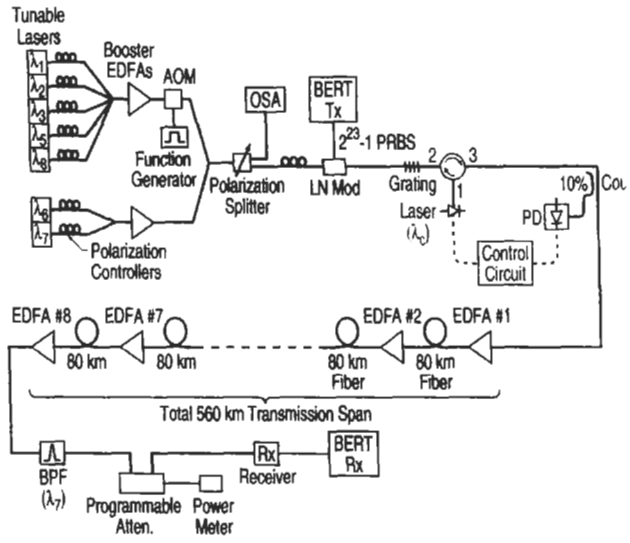


Fig. 23 Schematic of experimental setup for link control demonstration.

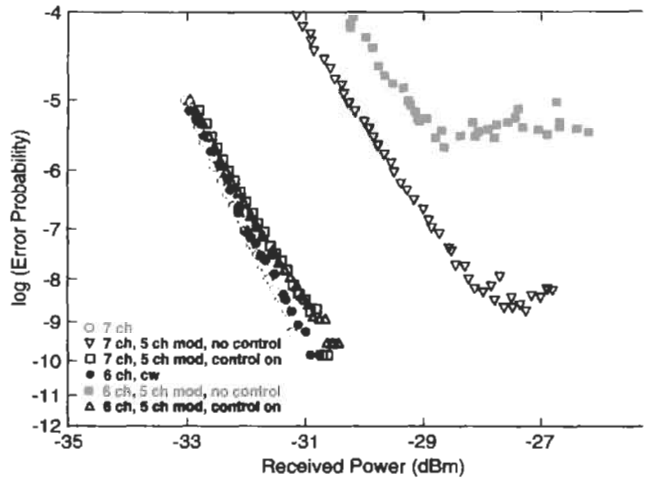


Fig. 24 Bit-error rates measured for transmission of 6 and 7 channels without modulation and with modulation of channels 1, 2, 3, 5, and 8 on/off at 1 KHz with and without link control.

Laser Control

Laser automatic gain control has been extensively studied [75]. A new scheme for link control based on laser gain control has been proposed recently [76]. In this work, a compensating signal in the first amplifier is generated using an optical feedback laser loop and then propagates down the link. Stabilization is

reached within a few tens of microseconds, and output power excursion after 6 EDFAs is reduced by more than an order of magnitude to a few tenths of a dB. For laser gain control, the speed is limited by laser relaxation oscillations [77], which are generally on the order of tens of microseconds or slower. Inhomogeneous broadening of EDFAs and the resulting spectral hole burning can cause gain variations at the signal wavelength, which will limit the extent of control from this technique. The same is true for the link control scheme.

V. Conclusions

The successful introduction of commercial WDM systems, enabled by practical EDFAs, has in turn fueled the development of high-power, wide-bandwidth, low-noise, gain-flattened optical amplifiers. Availability of such high-performance optical amplifiers and other advanced optical technologies, as well as the market demand of more bandwidth at lower cost, have made optical networking an attractive solution for advanced networks. Optical networking utilizes the WDM wavelengths not only to transport large capacity but also to route and switch different channels. Compared to point-to-point systems, optical networking applications make some of the requirements on optical amplifiers more demanding: gain flatness, wide bandwidth, and dynamic gain control.

Considerable progress has been made in the optical amplifiers in recent years. The bandwidth of amplifiers has increased nearly seven times and flat gain amplifiers with 84 nm bandwidth have been demonstrated. This has been made possible by addition of the L-band branch. With the advent of these amplifiers, commercial terabit lightwave systems are realizable. Progress has also been made in the understanding of gain dynamics of amplifiers. Several control schemes have been successfully demonstrated to mitigate the signal impairment due to fast power transients in a chain of amplifiers and will be implemented in lightwave network design. Terrestrial lightwave systems have been increasing in transmission capacity. The amplifiers have evolved significantly in performance to enable the currently available terabit capacity system soon be followed by systems having multiterabit and higher capacity on a single fiber.

Acknowledgments

Much of the work discussed in this chapter was done at Bell Laboratories and Optical Networking Group of Lucent Technologies. The authors appreciate greatly the close collaborations with many colleagues, among others, J. Sulhoff and C. Wolf. Special thanks go to J. L. Zyskind, now with Sycamore Networks, who helped the authors and made significant contribution on erbium-doped

fiber amplifiers research at Crawford Hill Laboratory. The authors also acknowledge the help and support from colleagues at Onetta. In particular, discussions with Chia-Chi Wang contributed significantly to this chapter.

References

1. K. Fukuchi, et al., "10.92 Tb/s (273 × 40 Gb/s) Triple-Band/Ultra-Dense WDM Optical Repeatered Transmission Experiment," PD24; and S. Bigo, et al., "10.2 Tb/s (256 × 42.7 Gb/s PDM/WDM) Transmission over 100 km Tera-Light Fiber with 1.28 bit/s/Hz Spectral Efficiency," PD25, *Proc. Optical Fiber Communications Conf.*, Anaheim, CA, March, 2001.
2. S. R. Johnson and V. L. Nichols, "Advanced Optical Networking—Lucent's MONET Network Elements," *Bell Labs. Tech. J.*, Vol. 4, Jan–Mar 1999, pp. 145–162.
3. R. J. Mears, et al., "Low noise erbium-doped fiber amplifier operating at 1.54 μm," *Electronics Letters*, Vol. 23, September 10, 1987, pp. 1026–1028.
4. E. Desurvire, J. R. Simpson, and P. C. Becker, "High gain erbium-doped traveling wave fiber amplifier," *Optics Letters*, Vol. 12, November 1987, pp. 888–890.
5. C. R. Giles and E. Desurvire, "Modeling erbium-doped fiber amplifiers," *J. Lightwave Technol.*, Vol. 9, February, 1991, pp. 271–283.
6. Y. Sun, J. L. Zyskind, and A. K. Srivastava, "Average Inversion Level, Modeling, and Physics of Erbium-Doped Fiber Amplifiers," *IEEE J. Selected Areas in Quantum Elect.*, Vol. 3, No. 4, August, 1997, pp. 991–1007.
7. P. F. Wysocki, J. R. Simpson, and D. Lee, "Prediction of Gain Peak Wavelength for Er-Doped Fiber Amplifiers and Amplifier Chains," *IEEE J. Photon. Technol. Lett.*, Vol. 6, No. 9, 1994, pp. 1098–1100.
8. J. L. Zyskind, J. A. Nagel, and H. D. Kidorf, "Erbium-Doped Fiber Amplifiers," *Optical Fiber Telecommunications*, Vol. IIIB, ed. by I. P. Kaminow and T. L. Koch, Academic Press, 1997, pp. 13–68.
9. R. G. Smart, J. L. Zyskind, and D. J. DiGiovanni, "Two-stage erbium-doped fiber amplifiers suitable for use in long-haul soliton systems," *Electron. Lett.*, Vol. 30, No. 1, January, 1994, pp. 50–52.
10. J.-M. P. Delavaux and J. A. Nagel, "Multi-stage erbium-doped fiber amplifier designs," *J. Lightwave Technol.*, Vol. 13, No. 5, May, 1995, pp. 703–720.
11. A. R. Chraplyvy, J.-M. Delavaux, R. M. Derosier, G. A. Ferguson, D. A. Fishman, C. R. Giles, J. A. Nagel, B. M. Nyman, J. W. Sulhoff, R. E. Tench, R. W. Tkach, and J. L. Zyskind, "1420-km Transmission of Sixteen 2.5-Gb/s Channels using Silica-Fiber-Based EDFA Repeaters," *Photonic Technol. Lett.*, Vol. 6, No. 11, 1994, pp. 1371–1373.
12. Y. Chen, M. T. Fatehi, H. J. La Roche, J. Z. Larsen and B. L. Nelson, "Metro Optical Networking," *Bell Labs. Tech. J.*, Vol. 4, Jan–Mar, 1999, pp. 163–186.
13. F. Forgieri, R. W. Tkach, and A. R. Chraplyvy, "Fiber Nonlinearities and their Impact on Transmission Systems," *Optical Fiber Communications*, Vol. IIIA, ed. by I. P. Kaminow and T. L. Koch, Academic Press, 1997, pp. 196–264.

14. Y. Sun, A. K. Srivastava, J. Zhou, and J. W. Sulhoff, "Optical Fiber Amplifiers for WDM Optical Networks," *Bell Labs. Tech. J.*, Jan–Mar, 1999, pp. 187–205.
15. Y. Sun, J. B. Judkins, A. K. Srivastava, L. Garrett, J. L. Zyskind, J. W. Sulhoff, C. Wolf, R. M. Derosier, A. H. Gnauck, R. W. Tkach, J. Zhou, R. P. Espindola, A. M. Vengsarkar, and A. R. Chraplyvy, "Transmission of 32-WDM 10-Gb/s Channels Over 640 km Using Broad Band, Gain-Flattened Erbium-Doped Silica Fiber Amplifiers," *IEEE Photon. Tech. Lett.*, Vol. 9, No. 12, December 1997, pp. 1652–1654.
16. S. Kinoshita, Y. Sugaya, H. Onaka, M. Takeda, C. Ohshima, and T. Chikama, "Low-noise and wide dynamic range erbium-doped fiber amplifers with automatic level control for WDM transmission systems," *Proc. Opt. Amplifiers and Their Appl.*, Monterey, CA, July 11–13, 1996, pp. 211–214.
17. S. Artigaud, et al., "Transmission of 16 × 10 Gb/s channels spanning 24 nm over 531 km of conventional single-mode fiber using 7 in-line fluoride-based EDFAs," OFC '96 Postdeadline Paper PD27, 1996.
18. M. Fukutoku, et al., "25 nm Bandwidth Optical Gain Equalization for 32-Channel WDM Transmission with a Lattice Type Optical Circuit," *Optical Amplifiers and their Applications*, 1996, p. 66.
19. A. M. Vengsarkar, et al., "Long-period fiber gratings based gain equalizers," *Optics Letters*, Vol. 21, 1996, p. 336.
20. R. Kashyap, R. Wyatt, and P. F. McKee, "Wavelength Flattened Saturated Erbium Amplifier using Multiple Side-tap Bragg Gratings," *Elect. Lett.*, Vol. 29, 1993, pp. 1025–1026.
21. Y. P. Li, et al., "Waveguide EDFA Gain Equalization Filter," *Elect. Lett.*, Vol. 31, 1995, pp. 2005–2006.
22. S. H. Yun, B. W. Lee, H. K. Kim, and B. Y. Kim, "Dynamic Erbium-Doped Fiber Amplifier with Automatic Gain Flattening," OFC '99 post-deadline paper PD28, 1999.
23. P. Wysocki, J. B. Judkins, R. P. Espindola, M. Andrejco, and A. M. Vengsarkar, "Broadband Erbium-Doped Fiber Amplifier Flattened Beyond 40 nm Using Long Period Grating Filter," *Photonic Technol. Lett.*, Vol. 9, 1997, pp. 1343–1345.
24. M. Tachibana, R. I. Laming, P. R. Morkel, and D. N. Payne, "Spectral Gain Cross Saturation and Hole Burning in Wideband Erbium Doped Fiber Amplifier," Topical Meeting of Optical Amplifiers and Applications, Optical Society of America, Washington, DC, 1991, p. 104.
25. V. J. Mazurczyk, "Spectral Response of a Single EDFA Measured to an Accuracy of 0.01," Optical Fiber Conference, San Jose, California, 1994, p. 271.
26. H. Chou and J. Stimple, "Inhomogeneous Gain Saturation of Erbium-Doped Fiber Amplifiers," Topical Meeting on Optical Amplifiers and Their Applications," Vol. 18, OSA. Washington, DC, 1995, p. 92.
27. A. K. Srivastava, J. L. Zyskind, J. M. Sulhoff, J. D. Evankow, Jr., and M. A. Mills, "Room Temperature Spectral-Hole Burning in Erbium-Doped Fiber Amplifiers," OFC '96 Technical Digest, San Jose, CA, 1996, pp. 33–34.

28. J. W. Sulhoff, A. K. Srivastava, C. Wolf, Y. Sun, and J. L. Zyskind, "Spectral-Hole Burning in Erbium-Doped Silica and Fluoride Fibers," *IEEE Photon. Tech. Lett.*, Vol. 9, No. 12, December 1997, pp. 1578–1579.
29. N. S. Bergano, C. R. Davidson, M. A. Mills, P. C. Corbett, S. G. Evangelides, B. Pedersen, R. Menges, J. L. Zyskind, J. W. Sulhoff, A. K. Srivastava, C. Wolf, and J. Judkins, "Long Haul WDM Transmission Using Optimum Channel Modulation: A 16 Gb/s (32 × 5 Gb/s) 9300 km Demonstration," OFC '97 Technical Digest, Postdeadline Papers, PD 16, Dallas, TX, February 22–27, 1997.
30. H. Ono, M. Yamada, and Y. Ohishi, "Gain-flattened Er-doped fiber amplifier for a WDM signal in 1.57–1.60 μm wavelength region," *IEEE Photonics Tech. Lett.*, Vol. 9, 1997, pp. 596–598.
31. S. Ishikawa, M. Kakui, M. Nishimura, and Y. Makio, "High Gain Per Unit Length Silica-Based Erbium Doped Fiber for 1580 nm Band Amplification," *Proc. Optical Amplifiers and Their Applications*, Vail, CO, July 27–29, 1998, pp. 111–114.
32. T. Sakamoto, J. Kani, S. Aisawa, K. Oda, M. Fukui, H. Ono, M. Yamada, and K. Oguchi, "1580 nm Band Equally-Spaced 8 × 10 Gb/s WDM Channel Transmission over 360 km (3×120 km) of Dispersion-Shifted Fiber Avoiding FWM Impairment," *Proc. Optical Amplifiers and Their Applications*, Victoria, Canada, July 21–23, 1997, pp. 36–39.
33. S. Aisawa, T. Sakamoto, J. Kani, S. Norimatsu, M. Yamada, and K. Oguchi, "Equally Spaced 8 × 10 Gb/s Optical Duobinary WDM Transmission in 1580-nm Band over 500 km of Dispersion-Shifted Fiber without Dispersion Compensation," OFC 1998, February 22–27, San Jose, CA, pp. 411–412.
34. Y. Yano, M. Yamashita, T. Suzuki, K. Kudo, M. Yamaguchi, and K. Emura, "640 Gb/s WDM Transmision over 400 km of Dispersion Shifted Fiber using 1.58 μm Band and Initial Chirp Optimization," ECOC '98, Madrid, Spain, September 20–24, pp. 261–264.
35. A. K. Srivastava, Y. Sun, J. L. Zyskind, J. W. Sulhoff, C. Wolf, J. B. Judkins, J. Zhou, M. Zirngibl, R. P. Espindola, A. M. Vengsarkar, Y. P. Li, and A. R. Chraplyvy, "Error free transmission of 64 WDM 10 Gb/s channels over 520 km of TrueWave™ fiber," Proc. ECOC '98, Madrid, Spain, September 20–24, 1998, p. 265.
36. A. K. Srivastava, S. Radic, C. Wolf, J. C. Centanni, J. W. Sulhoff, K. Kantor, and Y. Sun, Ultra-dense Terabit Capacity Transmission in L-Band, Postdeadline paper, OFC '00 Paper PD-27, Baltimore, MD, March 7–10, 2000.
37. M. Shtaif, M. Eiselt, R. W. Tkach, R. G. Stolen, and A. H. Gnauck, "Cross-talk in WDM Systems Caused by Cross-Phase Modulation in Erbium-Doped Fiber Amplifiers," *IEEE Photonics Technology Letters*, Vol. 10, 1998, pp. 1796–1798.
38. D. R. Marcuse, A. R. Chraplyvy, and R. W. Tkach, "Dependence of Cross-Phase Modulation on Channel Number in Fiber WDM Systems," *J. Lightwave Technol.*, Vol. 12, 1994, pp. 885–890.
39. M. Eiselt, M. Shtaif, R. W. Tkach, F. A. Flood, S. Ten, and D. Butler, "Cross-Phase Modulation in L-Band EDFA," *IEEE Photonics Technology Letters*, Vol. 11, No. 12, December 1999, pp. 1578–1580.

40. G. J. Pendock, S. Y. Park, A. K. Srivastava, S. Radic, J. W. Sulhoff, C. L. Wolf, K. Kantor, and Y. Sun, "The Contribution to Cross-Phase Modulation in L-Band WDM Systems from Erbium-Doped Fiber Amplifiers," *IEEE Photonics Technology Letters*, Vol. 11, No. 12, December 1999, pp. 1575–1577.
41. S. Radic, G. Pendock, A. K. Srivastava, P. Wysocki, and A. R. Chraplyvy, "Signal Impairment due to Four-Wave Mixing in L-Band EDFAs," Proc. ECOC, Postdeadline Paper, Nice, France, September 26–30, 1999, pp. 22–23.
42. S. Radic, G. Pendock, A. Srivastava, P. Wysocki, and A. Chraplyvy, "FWM in Optical Links using Quasi-Distributed Optical Amplifiers," *J. Lightwave Technol.*, Vol. 19, 2001, p. 636.
43. J. Massicott, J. R. Armitage, R. Wyatt, B. J. Ainslie, and S. P. Craig-Ryan, "High Gain, Broadband, 1.6 μm Er3+ Doped Silica Fiber Amplifier," *Electron. Lett.*, Vol. 20, 1990, pp. 1645–1646.
44. Y. Sun, J. W. Sulhoff, A. K. Srivastava, J. L. Zyskind, C. Wolf, T. A. Strasser, J. R. Pedrazzani, J. B. Judkins, R. P. Espindola, A. M. Vengsarkar, and J. Zhou, "Ultra Wide Band Erbium-Doped Silica Fiber Amplifier with 80 nm of Bandwidth," Optical Amplifiers and their Applications, Postdeadline paper PD2, Victoria B.C., Canada, July 1997.
45. M. Yamada, et al., "Broadband and Gain-Flattened Amplifier Composed of a 1.55 μm band and a 1.58 μm band Er^{3+}-doped Fiber Amplifier in a Parallel Configuration," *Electron. Lett.*, Vol. 33, No. 8, April 1997, pp. 710–711.
46. Y. Sun, J. W. Sulhoff, A. K. Srivastava, J. L. Zyskind, C. Wolf, T. A. Strasser, J. R. Pedrazzani, J. B. Judkins, R. P. Espindola, A. M. Vengsarkar, and J. Zhou, "An 80 nm Ultra Wide Band EDFA Low Noise Figure and High Output Power," ECOC '97, Postdeadline paper TH3C, Edinburgh, UK, September 22, 1997, pp. 69–72.
47. Y. Sun, J. W. Sulhoff, A. K. Srivastava, A. Abramov, T. A. Strasser, P. F. Wysocki, J. R. Pedrazzani, J. B. Judkins, R. P. Espindola, C. Wolf, J. L. Zyskind, A. M. Vengsarkar, and J. Zhou, "A Gain-Flattened Ultra Wide Band EDFA for High Capacity WDM Optical Communications Systems," Proc. ECOC '98, Madrid, Spain, September 20–24, 1998, p. 53.
48. L. D. Garrett, A. H. Gnauck, F. Forghieri, V. Gusmeroli, and D. Scarano, "8 × 20 Gb/s–315 km, 8 × 10 Gb/s–480 km WDM Transmission over Conventional Fiber Using Multiple Broadband Fiber Gratings," OFC '98 Technical Digest, Postdeadline Paper PD18, San Jose, CA, February 22–27, 1998.
49. A. K. Srivastava, Y. Sun, J. W. Sulhoff, C. Wolf, M. Zirngibl, R. Monnard, A. R. Chraplyvy, A. A. Abramov, R. P. Espindola, T. A. Strasser, J. R. Pedrazzani, A. M. Vengsarkar, J. L. Zyskind, J. Zhou, D. A. Ferrand, P. F. Wysocki, J. B. Judkins, and Y. P. Li, "1 Tb/s transmission of 100 WDM 10 Gb/s channels over 400 km of TrueWave™ fiber," OFC '98 Technical Digest, Postdeadline Papers PD10, San Jose, CA, February 22–27, 1998.
50. S. Aisawa, T. Sakamoto, M. Fukui, J. Kani, M. Jinno, and K. Oguchi, "Ultra-wide Band, Long Distance WDM Transmission Demonstration: 1 Tb/s (50 × 20 Gb/s), 600 km Transmission Using 1550 and 1580 nm Wavelength Bands," OFC

'98 Technical Digest, Postdeadline Paper PD11, San Jose, CA, February 22–27, 1998.

51. P. B. Hansen, L. Eskildsen, A. J. Stentz, T. A. Strasser, J. Judkins, J. J. Demarco, R. Pedrazzani, and D. J. DiGiovanni, "Rayleigh Scattering Limitations in Distributed Raman Preamplifiers," *IEEE Photon Tech. Lett.*, Vol. 10, No. 1, January 1998, pp. 159–161.
52. L. D. Garrett, M. Eiselt, R. W. Tkach, V. Domini, R. Waarts, D. Gilner, and D. Mehuys, "Field Demonstration of Distributed Raman Amplification with 3.8 dB Q-Improvement for 5 × 120-km Transmission," *IEEE Photon Tech. Lett.*, Vol. 13, No. 2, February 2001, pp. 157–159.
53. A. K. Srivastava, Y. Sun, L. Zhang, J. W. Sulhoff, and C. Wolf., "System Margin Enhancement with Raman Gain in Multi-Span WDM Transmission," OFC '99 Paper FC2, San Diego, CA, February 1999, pp. 53–55.
54. H. Suzuki, J. Kani, H. Masuda, N. Takachio, K. Iwatsuki, Y. Tada, and M. Sumida, "25 GHz Spaced 1Tb/s (100 × 10 Gb/s) Super Dense WDM Transmission in the C-Band over a Dispersion Shifted Cable Employing Distributed Raman Amplification," ECOC '99, PD2-4, Nice, France, September 26–30.
55. Y. Yano, et al., ECOC '96, Paper ThB.3.1, Oslo, Norway, 1996.
56. K. Rottwitt and H. D. Kidorff, "A 92 nm Bandwidth Raman Amplifier," OFC '98 Technical Digest, Postdeadline Paper PD6, San Jose, CA, February 22–27, 1998.
57. Y. Emori and S. Namiki, "100 nm Bandwidth Flat Gain Raman Amplifiers Pumped and Gain Equalized by 12-Wavelength Channel WDM High Power Laser Diodes," OFC '99 postdeadline paper PD19, San Diego, CA, February 1999.
58. T. N. Nielsen, A. J. Stentz, P. B. Hansen, Z. J. Chen, D. S. Vengsarkar, T. A. Strasser, K. Rottwitt, J. H. Park, S. Stultz, S. Cabot, K. S. Feder, P. S. Westbrook, and S. G. Kosinski, "1.6 Tb/s (40 × 40 Gb/s) Transmission over 4 × 100 km nonzero-dispersion fiber using hybrid Raman/Erbium-doped inline amplifiers," postdeadline paper, ECOC '99, PD2-2, Nice, France, September 26–30.
59. T. N. Nielsen, A. J. Stentz, K. Rottwitt, D. S. Vengsarkar, Z. J. Chen, P. B. Hansen, J. H. Park, K. S. Feder, T. A. Strasser, S. Cabot, S. Stultz, D. W. Peckham, L. Hsu, C. K. Kan, A. F. Judy, J. Sulhoff, S. Y. Park, L. E. Nelson and L. Gruner-Nielsen, "3.28 Tb/s (82 × 40 Gb/s) Transmission over 3 × 100 km nonzero-dispersion fiber using dual C- and L-band hybrid Raman/Erbium-doped inline amplifiers," OFC '00 postdeadline paper PD-23, Baltimore, MD, March 7–10.
60. A. K. Srivastava, J. L. Zyskind, Y. Sun, J. Ellson, G. Newsome, R. W. Tkach, A. R. Chraplyvy, J. W. Sulhoff, T. A. Strasser, C. Wolf, and J. R. Pedrazzani, "Fast-Link Control Protection of Surviving Channels in Multiwavelength Optical Networks," *IEEE Photonics Technology Letters*, Vol. 9, No. 11, December 1997, pp. 1667–1669.
61. A. K. Srivastava, Y. Sun, J. L. Zyskind, and J. W. Sulhoff, "EDFA Transient Response to Channel Loss in WDM Transmission System," *IEEE Photonics Technology Letters*, Vol. 9, No. 3, March 1997, pp. 386–388.
62. C. R. Giles, E. Desurvire, and J. R. Simpson, "Transient Gain and Cross-talk in Erbium-doped Fiber Amplifier," *Optics Letters*, Vol. 14, No. 16, 1989, pp. 880–882.

63. E. Desurvire, "Analysis of Transient Gain Saturation and Recovery in Erbium-doped Fiber Amplifiers," *IEEE Photonics Technology Letters*, Vol. 1, 1989, pp. 196–199.
64. Y. Sun, J. L. Zyskind, A. K. Srivastava, and L. Zhang, "Analytical Formula for the Transient Response of Erbium-Doped Fiber Amplifiers," *Applied Optics*, Vol. 38, No. 9, March 20, 1999, pp. 1682–1685.
65. E. Desurvire, *Erbium-doped Fiber Amplifiers—Principles and Applications*, John Wiley, 1994, pp. 469–480.
66. P. R. Morkel and R. I. Laming, "Theoretical Modeling of Erbium-Doped Fiber Amplifiers with Excited State Absorption," *Opt. Lett.*, Vol. 14, 1989, pp. 1062–1064.
67. Y. Sun, G. Luo, J. L. Zyskind, A. A. M. Saleh, A. K. Srivastava, and J. W. Sulhoff, "Model for Gain Dynamics in Erbium-doped Fiber Amplifiers," *Electronics Letters*, Vol. 32, No. 16, August 1, 1996, pp. 1490–1491.
68. J. L. Zyskind, Y. Sun, A. K. Srivastava, J. W. Sulhoff, A. L. Lucero, C. Wolf, and R. W. Tkach, "Fast Power Transients in Optically Amplified Multiwavelength Optical Networks," *Proc. Optical Fiber Communications Conference*, 1996, postdeadline paper PD31.
69. Y. Sun, J. L. Zyskind, A. K. Srivastava, J. W. Sulhoff, C. Wolf, and R. W. Tkach, "Fast Power Transients in WDM Optical Networks with Cascaded EDFAs," *Electron. Letters*, Vol. 33, No. 4, February 13, 1997, pp. 313–314.
70. MONET 2nd Quarterly Report, Payable Milestone # 2, 1995.
71. Y. Sun and A. K. Srivastava, "Dynamic Effects in Optically Amplified Networks," *Proc. Optical Amplifiers and their Applications*," Victoria, Canada, July 21–23, 1997, pp. 44–47.
72. S. J. Shieh, J. W. Sulhoff, K. Kantor, Y. Sun, and A. K. Srivastava, "Dynamic Behavior in L-Band EDFA," OFC '00 Paper PD27, Baltimore, MD, March 7–10, 2000.
73. A. K. Srivastava, Y. Sun, J. L. Zyskind, J. W. Sulhoff, C. Wolf, and R. W. Tkach, "Fast Gain Control in Erbium-doped Fiber Amplifiers," *Proc. Optical Amplifiers and their Applications*, 1996, postdeadline paper PDP4.
74. E. Desurvire, M. Zirngibl, H. M. Presby, and D. DiGiovanni, "Dynamic Gain Compensation in Saturated Erbium-Doped Fiber Amplifiers," *IEEE Photonic Technology Letters*, Vol. 3, May 1991, pp. 453–455.
75. M. Zirngibl, "Gain Control in Erbium-Doped Fiber Amplifier by an all-Optical Feedback Loop," *Electron. Letters*, Vol. 27, 1991, pp. 560–561.
76. J. L. Jackel and D. Richards, "All-Optical Stabilization of Cascaded Multi-Channel Erbium-doped Fiber Amplifiers with Changing Numbers of Channels", OFC '97 Technical Digest, Dallas, TX, February 22–27, 1997, pp. 84–85.
77. G. Luo, J. L. Zyskind, Y. Sun, A. K. Srivastava, J. W. Sulhoff, and M. A. Ali, "Relaxation Oscillations and Spectral Hole Burning in Laser Automatic Gain Control of EDFAs," OFC '97 Technical Digest, Dallas, TX, February 22–27, 1997, pp. 130–131.

Chapter 5 | Raman Amplification in Lightwave Communication Systems

Karsten Rottwitt

Ørsted Laboratory, Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

Andrew J. Stentz

Photuris, Inc., Piscataway, New Jersey

Introduction

Light propagating through an optical fiber interacts with the molecules of the fiber. In one process, light is scattered by the molecules, generating a nonpropagating vibration of the molecules. This process is known as the Raman effect. It was first discovered in 1928 by the Indian physicist Sir C. V. Raman who was awarded the Nobel Prize in 1930 for his work. Since 1928, Raman scattering has been used in many applications. For example, Raman spectroscopy has been used extensively in the study of molecular structures (Willard, 1988).

The application of Raman scattering in optical fibers is a promising technology for extending the capabilities of lightwave communication systems. This potential has spurred renewed interest in a venerable research topic. R. H. Stolen, E. P. Ippen, and A. R. Tynes (Stolen, 1972) were among the first to demonstrate Raman scattering in optical fibers in the early 1970s. Work on Raman amplification in optical fibers in the 1970s was scarce due to the lack of low-loss fiber and appropriate pump sources. In the late 1980s, with the commercialization of erbium-doped fiber amplifiers (EDFAs), the work on Raman amplification was largely abandoned. Extremely high signal gains and low noise figures can be achieved in EDFAs with modest pump powers. (For instance, in an EDFA a signal gain of a few dB per mW of pump power is obtainable, whereas only 0.06 dB of gain per mW of pump power in a Raman amplifier is achievable in typical optical transmission fibers at present.) Over the last several years, we have witnessed a tremendous rebirth of the interest in Raman amplification in optical fibers. Practical, efficient, high-power pump sources have diminished the disadvantage of the relatively poor efficiency of the Raman process compared to the erbium amplification process, and Raman amplifiers do offer several very attractive advantages over EDFAs.

One very important feature of Raman amplifiers is their capability to provide gain at any signal wavelength, as opposed to optical amplifiers based on rare-earth ions. This feature makes the Raman amplifier particularly attractive

OPTICAL FIBER TELECOMMUNICATIONS, VOLUME IVA

ISBN: 0-12-395172-0

for future lightwave systems, where there is a strong desire to continually increase the number of signal wavelengths by extending the bandwidth of transmission in dense wavelength division multiplexed systems. In addition, the wavelength window over which amplification is achievable is expandable simply by combining multiple pump wavelengths (Rottwitt, 1998).

In addition to being able to provide gain at any wavelength, the Raman amplifier offers improved noise performance. Since the amplification occurs over a significant segment of the transmission fiber itself, typically 20 km, the signal does not drop as much as it would in a conventional system based on lumped amplification. Thus, the signal-to-noise ratio in this distributed amplifier is improved over lumped amplification.

Finally, the Raman amplifier is very simple. Raman scattering of light is inherent to all silica-based optical fibers. Thus, the Raman amplifier is a very simple way to obtain cancellation of the intrinsic fiber loss. All that is needed to turn the transmission fiber into an amplifier by itself is to launch pump light simultaneously with the signal light.

The purpose of this chapter is to describe the basics of Raman scattering in optical fibers, as well as the relevant parameters, benefits, key components, and applications in communication systems. Section 1 addresses the basic properties of the Raman process in an optical fiber. The gain and noise performance may be predicted based on knowledge of the fiber and its Raman gain coefficient. In Section 2, we develop a simple model that yields some very useful results. The noise properties are critical and are treated in Section 3. We focus on noise originating from amplified spontaneous emission (ASE) and compare this quantity to the noise created by other optical amplification methods. In real optical communication systems, there are other phenomena that are just as important as noise properties from spontaneous emission. In Section 4, we describe some of these, including pump depletion, crosstalk, and Rayleigh scattering. The key enabler for utilizing Raman gain is the pump laser. In Section 5, we describe various configurations of pump lasers. The benefits of Raman gain, briefly mentioned above, include noise properties, bandwidth over which gain may be achieved, and simplicity. In Section 6, we will describe some of the applications published in the literature including recently published world record transmission experiments.

1. Raman Scattering in Optical Fibers

In this section we describe the Raman process in optical fibers from an applications point of view. In Section 1.1, we focus on the fundamentals of the Raman process, and in Section 1.2, we discuss the dependence of the Raman gain coefficient on material composition and waveguide properties.

1.1 THE RAMAN PROCESS

In a quantum mechanical picture, the Raman effect is a process in which a photon of frequency ν_p is scattered by a molecule in the fiber. In the most frequent event, known as Stokes scattering, the frequency of the photon is reduced to a lower frequency ν_s and energy is transferred from the original photon to the molecule in the form of kinetic energy, inducing stretching, bending, or rocking of the molecular bonds (Shibata, 1981). Quanta of these mechanical oscillations, known as optical phonons, have frequencies that are specific to the material and are strongly damped. The frequency difference between ν_p and ν_s is the Stokes frequency.

The probability of a Raman scattering event is proportional to the number of photons in the pump wave per cross-sectional area and the Raman cross-section. The frequency spectrum of the Raman cross-section is largely a function of material properties, since the Raman process is strongly related to the vibrational modes of the molecules of the material. In crystalline materials, the Raman scattered light has a narrow bandwidth centered about a specific frequency that is offset from the frequency of the launched light. This frequency difference is characteristic of the sample under study. Silica, currently the main constituent of any practical optical fiber, is amorphous, and therefore the Raman spectrum is relatively broad, in excess of 20 THz.

In 1978 F. L. Galeener et al. published Raman cross-sections of vitreous, or bulk, glass formers that are typically used when making high-purity optical fibers (Galeener, 1978). Figure 5.1 displays the intensity of the Raman

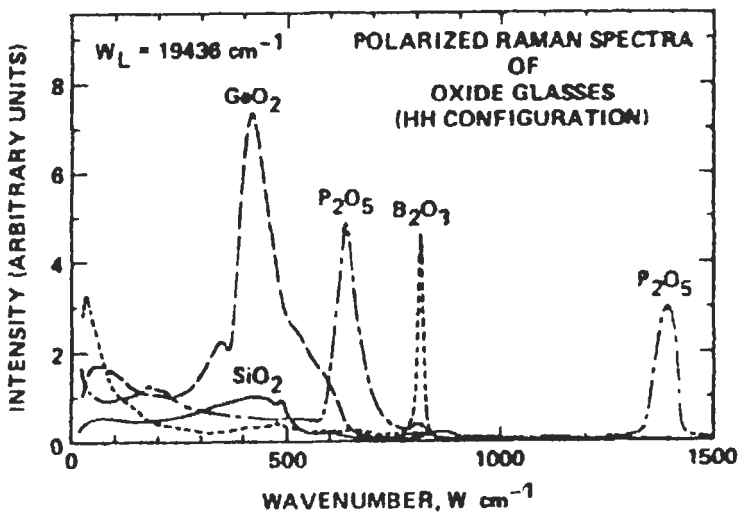


Fig. 5.1 Relative Raman cross-sections of several oxide glasses versus the Stokes shift in wavenumbers, $W = 1/\lambda_p - 1/\lambda_s$. (Reprinted with permission from Galeener, 1978.)

scattered light relative to that scattered by vitreous silica. Vitreous silica, the main constituent of optical fibers, has a main peak in its Raman cross-section at 440 cm^{-1}. Vitreous germanium, typically used to raise the refractive index of the fiber core, has a pronounced peak at 420 cm^{-1}, vitreous boron peaks at 808 cm^{-1}, and finally, vitreous phosphorus has two peaks, one at 640 cm^{-1}, and one at 1390 cm^{-1}.

The probability of a Raman scattering event is always nonzero but increases with increasing intensity of the pump wave and for those frequency differences where the material possesses a large Raman cross-section.

The aforementioned Stokes scattering process may occur even when the molecules of the fiber are completely at rest at zero Kelvin. The opposite process, where energy is transferred from the molecules to the photons, may also occur and is known as anti-Stokes scattering. In this process, the scattered photon has a higher frequency than the launched photon. The anti-Stokes is strongly temperature-dependent and does not happen at zero Kelvin.

The process by which a single photon scatters, generates an optical phonon, and attains a new frequency is referred to as spontaneous Raman scattering. The Raman process may also be stimulated. In this case, a photon of frequency ν_s stimulates the scattering of a photon of frequency ν_p, creating a new photon with identical frequency, phase, and direction as the incident photon of frequency ν_s. Since stimulated Raman scattering may be used to amplify communication signals, it is of great interest in optical communication systems. However, when analyzing an amplifier based on the Raman process, both stimulated and spontaneous scattering need to be considered, as well as the Stokes and anti-Stokes processes. We will return to this topic in Section 2.

1.2 RAMAN GAIN COEFFICIENT IN A FIBER

The Raman process may also be treated with a classical electromagnetic description. In single-mode optical fibers, the growth of stimulated Raman scattered signal intensity is proportional to the product of the pump and signal intensities and the gain coefficient, γ_r, such that

$$\frac{dI_s}{dz} = \gamma_r I_p I_s \tag{1}$$

where I_s and I_p are the intensities of the Stokes-shifted and pump waves, respectively. In general, the intensities of the pump and signal have radial, temporal, and frequency dependencies. In order to generate stimulated emission, it is strictly necessary that the Stokes and pump waves overlap spatially and temporally.

In single-mode optical fibers, one most often measures the total optical power traversing a section of fiber. The power P is found by integrating the

intensity over a cross-section of the fiber such that

$$P = \int_{\theta=0}^{2\pi} \int_{r=0}^{\infty} I(r,\theta) r \, dr \, d\theta \tag{2}$$

By integrating Eq. 1 over the cross-section of the fiber, we find that the growth in stimulated Raman scattered power is given by

$$\frac{dP_s}{dz} = P_p P_s \frac{\int_{r=0}^{\infty} \gamma_r(r) I_s(r) I_p(r) r \, dr}{2\pi \int_{r=0}^{\infty} I_s(r) r \, dr \int_{r=0}^{\infty} I_p(r) r \, dr} = P_p P_s g_r \tag{3}$$

Assuming that γ_r is invariant with radius, the integral in Eq. 3 reduces to a spatial overlap A_{eff}^R between the signal and pump modes, and the Raman gain coefficient g_r may be expressed as

$$g_r = \gamma_r \frac{\int_{r=0}^{\infty} I_s(r) I_p(r) r \, dr}{2\pi \int_{r=0}^{\infty} I_s(r) \, dr \int_{r=0}^{\infty} I_p(r) \, dr} = \frac{\gamma_r}{A_{eff}^R} \tag{4}$$

This approach was described by W. P. Urquhart et al. (1985). If in addition the spatial modes of the pump and Stokes waves are assumed to be identical, Eq. 3 may be approximated with

$$\frac{dP_s}{dz} = P_p P_s \frac{\gamma_r}{A_{eff}} \tag{5}$$

where the effective core area, A_{eff}, is similar to the effective core area of self-phase modulation (Agrawal, 1995). In the remainder of this chapter, we will use the Raman gain coefficient g_r of Eq. 3, which is directly obtainable from measurements of the Raman gain, the launched pump power, the intrinsic fiber attenuation, and the fiber length (see Section 2).

The Raman gain coefficient has not only the waveguide dependence discussed above but also depends on the fiber material. It is the material composition of the fiber that predominantly determines the spectral shape of the gain coefficient. When germanium, boron, or phosphorus is used to modify the refractive index of silica glass, the spectrum of the Raman cross-section of the binary glass is different from the spectra for the vitreous samples in Fig. 5.1.

In general the spectrum of the Raman cross-section of a binary glass depend on the concentration of the added glass modifier. N. Shibata et al. (Shibata, 1981) showed that binary GeO_2-SiO_2 glass (19 mol% GeO_2) has a pronounced peak at 425 cm^{-1}. A B_2O_3 doped glass containing 10 mol% B_2O_3 has a peak at 450 cm^{-1}. This Stokes shift is 358 cm^{-1} lower than the peak in Fig. 5.1.

A highly doped P_2O_5-SiO_2 glass (30 mol% P_2O_5) has two strong peaks in its Raman cross-section, one at 520 cm^{-1}, and one at 1320 cm^{-1}. Both of these Stokes shifts are also lower than the peaks in Fig. 5.1.

Since typical transmission fibers are germanosilicate fibers, we will henceforth focus exclusively on this type of fiber. In such fibers, the amplitude and spectral shape of the Raman gain coefficient varies with the germanium concentration. Galeener et al. (1978) showed that the Raman cross-section in vitreous germanium at its peak is nine times greater than that of vitreous silica. Figure 5.2 illustrates the Raman gain coefficient measured in four different fibers. The pure silica fiber contains no germanium, whereas the other three fibers have different germanium concentrations and waveguide properties. For a more fundamental review of the origin of Raman scattering and its spectral shape in vitreous and binary glasses, see Galeener et al. (1978, 1979, 1983) and Shibata (1981).

In Fig. 5.2, the silica-core fiber has the lowest Raman gain coefficient; the dispersion compensation fiber the highest. This is as expected since the dispersion compensation fiber has both a higher germanium concentration and a smaller effective core area than does the silica-core fiber.

In addition to the waveguide and material dependencies, the Raman scattering process also depends strongly on the state of polarization of the interacting

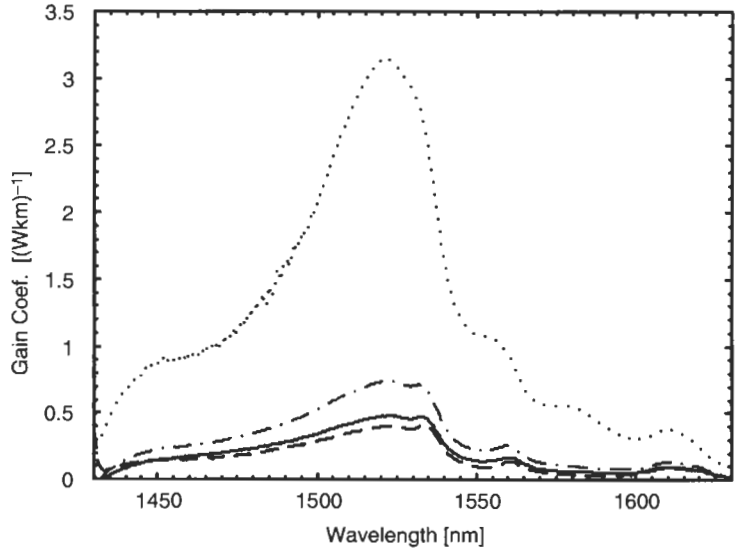


Fig. 5.2 Raman gain spectra for four fibers: a silica-core fiber (dashed), a Lucent TrueWave fiber (dashed/dotted), a standard single-mode fiber (solid), and a dispersion compensating fiber (dotted). The spectra are recorded with a nonpolarized pump source at 1427 nm (Rottwitt, 2000).

waves. It has been shown that the Raman gain coefficient for co-linearly polarized waves is more than an order of magnitude larger than the gain coefficient for orthogonal linear polarizations (Tobin, 1968). In typical optical fibers that do not preserve the states of polarization, the Raman gain coefficient is approximately half of its maximum since the states of polarization of the pump and signal change rapidly and independently along the length of fiber (Stolen, 1979).

When utilizing a nonpolarized pump, the Raman gain coefficient in typical transmission fibers, with effective areas $\sim 55\,\mu m^2$, has a value close to $0.7 \times 10^{-3}\,(\mathrm{Wm})^{-1}$. Finally, the Raman gain coefficient increases linearly with the pump frequency, $\nu_p = 1/\lambda_p$ (Stolen, 1984).

2. Theory of Raman Amplifiers

In optical communications systems, Raman amplification is typically used to amplify signals, and therefore the signals play the role of the Stokes wave. Henceforth we will therefore refer to the Stokes wave as the signal.

A theoretical treatment of Raman amplification is reasonably straightforward when the interacting waves are limited to a monochromatic pump and signal propagating through a lossless fiber. However, accurate modeling of Raman amplification in current communication systems generally requires the model to include bidirectional propagation of pumps, signals, and amplified spontaneous emission (ASE) over a multitude of wavelengths, accurate linear and nonlinear fiber parameters such as loss, gain coefficient over hundreds of nanometers, Rayleigh back-scattering of the signals as well as temperature dependencies.

In this section, we introduce theoretical models of Raman amplification. The model presented in Section 2.1 is detailed and requires numerical simulation. A simplified, approximate model is described in Section 2.2.

2.1 RATE EQUATIONS

One may imagine n signal photons at a frequency ν_s propagating through an infinitesimally short piece of fiber Δz simultaneously with m pump photons at a frequency ν_p. During Stokes scattering, an incident pump photon scatters off from a molecule creating a photon of lower energy, i.e., lower frequency, and a phonon. Conversely, during anti-Stokes scattering, a photon absorbs energy from the material creating a photon of higher energy. Both of these processes may be stimulated or spontaneous. We will denote the probability of stimulated and spontaneous emission of a Stokes-shifted photon as γ_e^{st} and γ_e^{sp}, respectively, and the probability for stimulated and spontaneous absorption through the anti-Stokes process as γ_a^{st} and γ_a^{sp}, respectively. If we denote the probability for counting n signal photons at position $z + \Delta z$ by $\wp_{z+\Delta z}^{n}$, we find

that (Townes, 1957)

$$\begin{aligned}\wp_{z+\Delta z}^{n} = {} & \wp_z^{n-1}\gamma_e^{st}(n-1)\Delta z + \wp_z^{n-1}\gamma_e^{sp}\Delta z \\ & + \wp_z^{n+1}\gamma_a^{st}(n+1)\Delta z + \wp_z^{n+1}\gamma_a^{sp}\Delta z \\ & + \wp_z^{n}\left\{1-(\gamma_e^{st}+\gamma_a^{st})n\Delta z - (\gamma_e^{sp}+\gamma_a^{sp})\Delta z\right\}\end{aligned} \tag{6}$$

The first two terms on the right-hand side of Eq. 6 represent the probability of the emission of a Stokes-shifted photon via stimulated or spontaneous emission. The next two terms account for the probability of absorption of a photon via stimulated and spontaneous anti-Stokes scattering. Finally, the term in brackets represents the probability that neither Stokes nor anti-Stokes scattering occurs in the fiber piece Δz.

The probability of stimulated absorption or emission depends on the number of pump photons. In a complete analysis, one needs to solve two coupled equations, one for the signal, and one for the pump. We will return to these coupled equations later in this section.

Summing over all photon numbers n ($n = 0, \infty$), and evaluating the change in average photon number $\langle n\rangle|_{z+\Delta z} - \langle n\rangle|_z$ in distance Δz, one arrives at the propagation equation for the average signal photon number when Δz approaches an infinitesimal length of fiber:

$$\frac{d\langle n\rangle}{dz} = (\gamma_e^{st} - \gamma_a^{st})\langle n\rangle + \gamma_e^{sp} - \gamma_a^{sp} \tag{7}$$

Assuming that the initial photon number at $z = 0$ is n_0, we find that the average photon number at position z, $\langle n(z)\rangle$, is given by

$$\langle n(z)\rangle = n_0 G(z) + n_{sp}(z) \tag{8}$$

where the signal gain, $G(z)$, and the spontaneous emission, $n_{sp}(z)$, are given by

$$G(z) = \exp\left[\int_{\tilde{z}=0}^{z} (\gamma_e^{st} - \gamma_a^{st})\, d\tilde{z}\right] \tag{9a}$$

$$n_{sp}(z) = G(z)\int_{\tilde{z}=0}^{z} \frac{\gamma_e^{sp} - \gamma_a^{sp}}{G(\tilde{z})}\, d\tilde{z} \tag{9b}$$

The probabilities of emission and absorption are related to the Raman cross-section of the fiber $\tilde{\sigma}$ and the thermal equilibrium excitation $\eta_T = \left(\exp\left(h(\nu_p - \nu_s)/kT\right) - 1\right)^{-1}$ through $\gamma_e^{st} = \tilde{\sigma}(\eta_T + 1)n_p$, $\gamma_e^{sp} \propto \tilde{\sigma}(\eta_T + 1)n_p$, $\gamma_a^{st} = \tilde{\sigma}\eta_T n_p$ and $\gamma_a^{sp} = \tilde{\sigma}\eta_T n_s$ (Yariv, 1989). Note that the gain coefficient in Eq. 7, $\gamma_e^{st} - \gamma_a^{st}$, is temperature-independent, due to the subtraction of the probabilities of stimulated absorption (i.e., anti-Stokes scattering) from stimulated

emission. However, note that the spontaneous emission γ_e^{sp} is temperature-dependent. This issue will be discussed in further detail in Section 4.4. For a more detailed description of the Raman amplification process see Mills, 1998.

While the above analysis hopefully provides an insight into the fundamental properties of Raman scattering, it is often more useful to derive equations based on more easily measured quantities such as pump and signal powers. Within a relatively small bandwidth B_0, Eq. 7 can be rewritten as

$$\frac{dP_s}{dz} = g_r P_p P_s + h\nu_s g_r P_p B_0 \tag{10}$$

where g_r is the Raman gain coefficient described in Section 1.2, and P_p and P_s are the powers of the pump and signal, respectively. In Eq. 10, it is assumed that the pump is nonpolarized whereas the signal may have a well-defined state of polarization. The Raman gain coefficient is an average taken over all states of polarization of the pump. Under these assumptions, an explicit treatment of the evolution of the states of polarization of the pump and signals is not required. In this situation, the gain coefficient plotted in Fig. 5.2 applies. If both the pump and signal have well-defined polarization states, integration of Eq. 10 becomes much more complicated, since the state of polarization for pump and signal typically is not preserved along the fiber. This topic is outside of the scope of this chapter.

In Eq. 10 it is assumed that the pump is at a single frequency. It should be noted that sometimes a factor of 2 is included on the spontaneous term in Eq. 7 to account for the fact that the signal typically occupies one state of polarization, whereas the spontaneous emissions are equally generated in both polarizations guided by the optical fiber.

The rate equation for the pump is

$$\frac{dP_p}{dz} = -\frac{\nu_p}{\nu_s} g_r P_s P_p - h\nu_p P_p \int_{\tilde{\nu}=\nu_p}^{\infty} g_r \, d\tilde{\nu} \tag{11}$$

The ratio (ν_p/ν_s) on the right-hand side of Eq. 11 accounts for the energy that is lost to the fiber in the form of phonons when pump photons scatter to signal photons. The second term on the right-hand side accounts for the spontaneous emission of Stokes-shifted photons at all frequencies with nonzero gain coefficients.

In the above treatment, we have presented two different approaches to predict the performance of a Raman amplifier: in Eq. 7 by counting photons, and in Eq. 10 by evaluating the evolution of optical powers. The advantage of counting photons is that the photon number is conserved when intrinsic fiber loss is neglected. Thus, as we will show in Section 3, an analytical solution of the coupled equations between pump and signal photons is obtainable. This is not possible in the approach evaluating the power evolution because of the

term (ν_p/ν_s) in Eq. 11. However, evaluating optical powers has the advantage of predicting directly measurable quantities.

When evaluating the Raman amplification process on a realistic optical fiber, Eq. 7 and Eq. 10 need to be expanded to include several other effects. When no amplification is provided, optical powers are reduced during propagation due to the intrinsic attenuation of the fiber. This attenuation may be treated by including the term $(-\alpha_s P_s)$ on the right-hand side of Eq. 10 and $(-\alpha_p P_p)$ on the right-hand side of Eq. 11, where α_s and α_p are the loss coefficients of the fiber at the signal and pump wavelengths, respectively. If the signal (and/or the pump) is propagating in the negative z direction, the right-hand side of Eq. 10 (and/or Eq. 11) needs to be multiplied by (-1).

In a typical Raman-amplified communication system, there are many signal wavelengths, and there may also be several pump wavelengths. In principle, any of these wavelengths may propagate in either direction in the fiber. Locally, the behavior of the Raman process is independent of the direction of propagation of the interacting waves. Thus, a continuous-wave treatment of Raman scattering with multiple wavelengths propagating in arbitrary directions needs only treat all possible combinations of Stokes and anti-Stokes scattering among the waves. However, it should be noted that a full temporal model of Raman interactions between waves propagating in arbitrary directions must include effects such as walkoff. These effects make copropagating and counterpropagating Raman interactions very different. This aspect will be discussed in Section 4 when crosstalk among pumps and signals is addressed.

Decomposing the waves into monochromatic spectral components, the power P_s at any spectral component ν_s, evolves as

$$\begin{aligned}\frac{dP_s}{dz} &= g_r^{s,s-1} P_s P_{s-1} + h\nu_s g_r^{s,s-1} P_{s-1} \\ &\quad - \frac{\nu_s}{\nu_{s+1}} g_r^{s,s-1} P_s P_{s+1} - h\nu_{s+1} \frac{\nu_s}{\nu_{s+1}} g_r^{s,s-1} P_{s+1} - \alpha_s P_s + B\alpha_s^R P_s^-\end{aligned} \tag{12}$$

The first two terms on the right-hand side of Eq. 12 treat the stimulated and spontaneous Stokes scattering of power into spectral component ν_s from waves of at shorter wavelengths. If there is power at multiple wavelengths shorter than λ_s, the factor $g_r^{s,s-1} P_{s-1}$ is replaced by a sum over all wavelengths shorter than λ_s of powers multiplied with the appropriate gain coefficients.

The next two terms treat depletion of the power at the spectral component ν_s. These terms arise from the fact that light at frequency ν_s will not only be amplified by power at shorter wavelengths but will also amplify waves at longer wavelengths. If there is power at more than one wavelength greater than λ_s, the factor $g_r^{s,s-1} \frac{r_s}{r_{s+1}} P_{s+1}$ needs to be replaced by a sum of the powers of all waves with wavelengths greater than λ_s multiplied by the appropriate gain coefficients and frequency ratios.

It should also be noted that in Eq. 12, P_s is the signal power propagating in the forward direction, whereas P_{s+1} and P_{s-1} are the total powers at longer and shorter wavelengths, regardless of their direction of propagation.

The fifth and sixth terms account for intrinsic loss and Rayleigh backscattering, respectively. Rayleigh scattering is the linear scattering of light from density fluctuations whose dimensions are smaller than the wavelength of light. Rayleigh scattering is the dominant source of loss in optical fibers at typical communications wavelengths. Rayleigh backscattering is that fraction of the Rayleigh scattered light that is recaptured by the fiber and is traveling in the direction opposite to that of the incident light. Due to the long lengths of fiber used in Raman amplifiers, Rayleigh scattering may play a critical role in determining the noise properties of the amplifier. This issue is addressed below in Section 4.3.

In general, one needs to solve Eq. 12 for the powers propagating in both directions. This is required not only to account for propagation of the pumps, signals, and Rayleigh-reflected signals, but also the Raman-scattered spontaneous emission.

Equation 12 is appropriate for treating the propagation of a polarized signal. However, the spontaneous emission is generated in arbitrary states of polarization along the length of fiber. This may be included in Eq. 12 as a factor of 2 in the second and fourth terms. Alternatively, an additional differential equation similar to Eq. 12 may be used that treats the orthogonal polarization.

2.2 *UNDEPLETED PUMP REGIME*

A particularly simple and useful scenario considers one monochromatic pump propagating with a weak signal through a fiber with intrinsic losses but with insignificant Rayleigh backscattering. Under these circumstances, the evolution of the pump will be independent of the signal and ASE powers and will be determined solely by the fiber intrinsic loss. If

$$g_r \frac{\nu_p}{\nu_s} P_s(z) + 2h\nu_p P_p \int_{\tilde{\nu}=\nu_p}^{\infty} g_r \, d\tilde{\nu} \ll \alpha_p \tag{13}$$

Eqs. 10 and 11 decouple and may easily be solved for the signal and pump powers. This so-called undepleted pump approximation is very powerful and can often be used to estimate the levels of gain and spontaneous emission. In a typical dispersion-shifted fiber with $\alpha_p \sim 0.057\,\mathrm{km}^{-1}$ (=0.25 dB/km), $g_r \sim 0.7 \times 10^{-3}$ (Wm^{-1}), the undepleted pump approximation is valid for $P_s(z) \ll 80\,\mathrm{mW}$.

In the remainder of this section we will use this simple approximation. For many practical situations it provide useful insight into Raman amplifiers.[1]

In the undepleted pump regime, Eq. 10 may be easily solved for the signal gain at the output of a fiber of length L,

$$G(L) = \exp\left(\int_{\tilde{z}=0}^{L} g_r P_p(\tilde{z})\, d\tilde{z} - \alpha_s L\right) \tag{14}$$

Eq. 14 predicts the net gain (or loss) experienced by a signal traversing a length of fiber with a pump in the undepleted pump regime. The Raman gain is often quoted in terms of the on-off gain: the ratio of the signal output power with the pump on to that with the pump off. The on-off gain is easily calculated in the undepleted pump regime by omitting the factor $\exp(-\alpha_s L)$ in Eq. 14.

The noise power P_{ASE} at the output end, $z = L$, is given by

$$P_{ASE} = G(L) \int_{\tilde{z}=0}^{L} \frac{(1+\eta_T) h \nu_s B_0 g_r P_p(\tilde{z})}{G(\tilde{z})}\, d\tilde{z} \tag{15}$$

Again, this is the spontaneous emission power in a single polarization state.

Since we assume that the pump wave is undepleted by the Raman interaction, the pump power decays exponentially with the intrinsic fiber loss. Thus $P_p^+(z) = P_p^0 \exp(-\alpha_p z)$ for a copumped Raman amplifier, whereas $P_p^-(z) = P_p^L \exp(-\alpha_p (L - z))$ for a counterpumped Raman amplifier.

In the undepleted pump regime, the on-off signal gains in the co- and counterpumped Raman amplifier are identical, and are given by

$$G = \exp\left(g_r L_{eff} P_p^{Launch}\right) \tag{16}$$

where P_p^{Launch} is the launched pump power, and where L_{eff} is the effective fiber length for the Raman interaction, in the following referred to as the Raman effective length, given by

$$L_{eff} = \frac{1 - \exp(-\alpha_p L)}{\alpha_p} \tag{17}$$

In a lossless optical fiber, the same Raman interaction would be achieved if the length of fiber were the Raman effective length, L_{eff}. The pump power only provides a significant contribution to the gain over the effective length of fiber. In a typical transmission fiber, the loss at the pump wavelength is ~0.25 dB/km, and therefore the Raman effective length is ~20 km.

Equation 16 shows that the on-off gain in dB increases linearly with the pump power in watts. For a typical dispersion-shifted transmission fiber (with $A_{eff} \sim 55\,\mu m^2$), an unpolarized pump, and an optical fiber much longer than the effective length, the on-off Raman gain is approximately 55 dB/W.

3. Noise in Raman Amplifiers

The noise properties of Raman amplifiers are at least as important as the gain characteristics for the application of Raman amplifiers to lightwave communication systems. The noise in an optical communication system may be interpreted as any phenomenon that degrades the bit-error-rate performance of a digital system or the signal-to-noise ratio of an analog communication system.

Common sources of noise in lightwave communication systems include the beating of the signal with ASE and nonlinear impairments due to four-wave mixing and self- and cross-phase modulation. While these sources of noise are also relevant to Raman-amplified systems, there are additional potential sources of noise that require careful attention. First, due to the long lengths of fiber used in Raman amplifiers, noise from Rayleigh scattering of light is an issue. This includes both the single Rayleigh backscattering of ASE as well as the double Rayleigh backscattering of the signals. Additionally, unlike EDFAs, the response time of Raman scattering is extremely fast. This particularly impacts systems utilizing copropagating Raman amplification where amplitude modulation from one signal may be coupled to another via the pump wave, or where fluctuations of the pump power itself may be coupled directly to a signal.

In this section we will focus on the accumulation of ASE. After developing the concept of noise figure in Section 3.1, we discuss the improvements in noise figure permitted by distributed Raman amplification in Section 3.2. In Section 3.3, we briefly discuss the impact of nonlinearities in systems utilizing distributed amplification.

The analysis presented in this section is based solely on a continuous-wave analysis. We do not attempt to correlate noise figure with bit-error rate, nor do we attempt to treat other temporal effects such as pump-signal walk-off. For these effects we refer the reader to references such as Marcuse, 1990, and Mahlein, 1984.

3.1 BASIC DEFINITIONS

Noise figure is a well-accepted metric for characterizing the noise performance of discrete amplifiers. In the following sections, we introduce the familiar concepts of signal-to-noise ratio and noise figure for optical amplifiers. However, the use of distributed amplifiers adds another layer of complexity to the characterization of noise performance. For instance, it is common for commercially available discrete EDFAs to have noise figures less than 6 dB, falling within a factor of two of the well-established 3 dB lower limit for high-gain amplifiers (Desurvire, 1999). However, one must remember that this noise figure limit applies only to amplifiers with large net gains. Ideally, one should be able to attain noise figures less than 3 dB when the net gain of a discrete amplifier

is small. However, when one considers distributed Raman amplifiers, where the amplification occurs over a long length of fiber with significant intrinsic loss, the net gain of the amplifier may be small, and yet the noise figure of the distributed amplifier may be significantly higher than 3 dB. Under these circumstances, one should not directly compare the noise figure of the distributed amplifier to a discrete EDFA. In order to facilitate a direct comparison to the performance of discrete amplifiers, we introduce the concept of effective noise figure for distributed amplifiers. In this section, we introduce these concepts and discuss some of the more frequently used formulas related to characterization of noise performance.

3.1.1 Signal-to-Noise Ratio

The concept of signal-to-noise ratio was first applied to lightwave communication systems to evaluate the performance of analog systems, and is defined as the squared average of the photocurrent induced by the signal in a receiver relative to the variance of the total current (Smith, 1987). This signal-to-noise ratio is called the electrical signal-to-noise ratio $(SNR_e)^2$ and is calculated as

$$SNR_e \equiv \frac{I_{sig}^2}{\sigma_d^2} \tag{18}$$

where I_{sig} is the mean time average of the current induced by the signal, and σ_d^2 is the variance of the current generated in the detector (Olsson, 1989; Desurvire, 1994; Haus, 1998). When ASE is the only source of noise generated by an amplifier, the variance of the photo current due to noise measured at the output of the amplifier may be sorted into three categories (Olsson, 1986): shot noise, beat noise, and thermal noise. Shot noise dominates at low signal values and is created by the quantization of light and current into photons and electrons. The dependence of the variance of shot noise on bandwidth and current is known as Schottky's formula (Boyd, 1983). Beat noise is noise created by the interference of signal with ASE or between components of the ASE itself. Thermal noise is the spontaneous noise generated by the detection circuits. In the example of an avalanche photo detector (Olsson, 1989; Desurvire, 1994) show that this is expressed as

$$\begin{aligned} \sigma_{shot}^2 &= \eta e 2B_e(GI_s + MI_N) \\ \sigma_{beat}^2 &= 2\eta^2 GI_s I_N \frac{2B_e}{B_0} + M\eta^2 I_N^2 \frac{2B_e}{B_0^2}\left(B_0 - \frac{B_e}{2}\right) \end{aligned} \tag{19}$$

where η is the quantum efficiency of the detector, M is the number of ASE modes (Desurvire, 1994), G is the signal gain, e is the electronic charge, I_s is the input signal current, I_N is the noise current within the bandwidth B_0, and B_e is the electrical bandwidth of the receiver. Eq. 19 does not account for

excess noise involved with the variance of the signal light when launched. In an ideal detector without circuit noise, where $\eta = 1$ and $B_e = B_0/2$, the total variance of the current reduces to

$$\sigma_d^2 = eB_0(GI_s + MI_N) + 2GI_sI_N + \frac{3}{4}MI_N^2 \tag{20}$$

When comparing transmission elements such as inline amplifiers, optical filters or transmission segments, it makes sense to assume that the detector is ideal. Under this assumption the electrical signal-to-noise ratio is conveniently expressed as

$$SNR_e = \frac{\langle n_{sig} \rangle^2}{V_n} \tag{21}$$

where $\langle n_{sig} \rangle$ is the mean number of photons being detected and V_n the variance of the number of photons being detected. In a practical system, the added noise has to be much smaller than the signal for the system to work, and therefore $\langle n_{sig} \rangle$ in Eq. 21 approximates the mean number of pure signal photons (Desurvire, 1994; Olsson, 1989; Haus, 1988).

For completeness, it should be noted that under the assumption that the beating between signal and ASE dominates the variance of the detected current, the electrical signal-to-noise ratio approximates

$$SNR_e \approx \frac{P_{sig}}{2\tilde{P}} \tag{22}$$

where P_{sig} is the signal power, and $\tilde{P}$ the noise power in the signal mode. Equation 22 illustrate that the electrical signal-to-noise ratio is 3 dB higher than the ratio of the optical signal power relative to the noise power.

3.1.2 Noise Figure

Noise figure represents the degradation in signal-to-noise ratio from the input to the output of a device and is defined as

$$F \equiv \frac{(SNR_e)_{in}}{(SNR_e)_{out}} \tag{23}$$

where F is the noise figure, $(SNR_e)_{in}$ and $(SNR_e)_{out}$ are the input and output electrical signal-to-noise ratios measured with an ideal detector, and assuming the input signal is shot-noise limited. (Desurvire, 1994).

For a high-gain discrete amplifier, the noise figure has a lower limit of 3 dB. However, in the low gain limit the noise figure approaches 0 dB. For a component with loss, the noise figure has a lower limit that equals the loss in dB. We will demonstrate this in Section 3.2.

3.1.3 Effective Noise Figure

For distributed amplification, where the signal is amplified in a transmission span that when unpumped may be quite lossy, the degradation of the signal-to-noise ratio is typically much larger than in a discrete amplifier. For comparative purposes, we therefore introduce the concept of an effective discrete amplifier with an effective noise figure and effective gain. The noise performance of the distributed amplifier is equivalent to the performance of an identical passive fiber span followed by the effective discrete amplifier. The effective noise figure is defined as

$$F_{eff} \equiv \frac{F}{T_{sp}} \tag{24}$$

where T_{sp} is the transmission through the passive fiber span to the point where the effective discrete amplifier is defined. For example, in a 100 km-long fiber with an intrinsic loss of 0.2 dB/km, the effective noise figure in decibels is $F_{eff}(\text{dB}) = F(\text{dB}) - 20\,\text{dB}$. The effective gain is simply the on-off Raman gain. We elaborate on the application of this construct below.

3.2 SPONTANEOUS EMISSION

A Raman amplifier is characterized by the gain and ASE generated during the amplification process. Equations 9a to 14 provide tools to calculate these. In the following section, we illustrate the manner in which distributed Raman amplifiers can be used to reduce the noise accumulation within communication systems by applying the above definitions to the following five cases: I) a passive fiber, II) a discrete amplifier, III) a passive fiber followed by a discrete amplifier, IV) a Raman amplifier, and V) a Raman amplifier followed by a discrete amplifier. We will assume that the discrete amplifiers are erbium doped fiber amplifiers (EDFAs).

3.2.1 A Passive Fiber

The mean number of signal photons emitted at the output end of a passive fiber is

$$\langle n(z)\rangle = n_0 \exp(-\alpha_s z) \tag{25}$$

where n_0 is the mean number of photons at $z = 0$ and α_s the background loss. The mean number of photons is related to signal power by $P_{sig} = \int h\nu n(\nu, t)\, d\nu$. If the signal is monochromatic with a frequency of ν_s, the signal power is $P_{sig} = h\nu_s \langle n\rangle$.

The variance V_n may be calculated from the mean number of photons and the second momentum of n, as shown by (Townes, 1957). Applying this, the variance is given by

$$V_n(z) = n_0 \exp(-\alpha_s z)\{1 - \exp(-\alpha_s z)\} + V_0 \exp(-2\alpha_s z) \tag{26}$$

where V_0 is the variance of the photon number at $z = 0$. If the signal source emits photons with Poissonian statistics, i.e. $V_0 = n_0$, the optical signal-to-noise ratio equals

$$SNR_e = n_0 \exp(-\alpha_s z) \tag{27}$$

The noise figure F after a fiber segment with length L is therefore $F = \exp(\alpha_s L)$.

The point is that while a passive fiber does not contribute to the noise power by generating ASE, it does degrade the signal-to-noise ratio through shot noise.

3.2.2 Discrete Amplification, an EDFA

The mean photon number and the variance of the photon number are well described in many publications, see, for example, Desurvire, 1994. The mean and variance are

$$\langle n \rangle = G_e n_0 + M\tilde{n} \tag{28a}$$

and

$$V_n = G_e^2(V_0 - n_0) + G_e n_0(2\tilde{n} + 1) + M\tilde{n}(\tilde{n} + 1) \tag{28b}$$

where $\tilde{n}$ and G_e are the noise photon number and the gain of the amplifier, respectively, and M is the number of signal modes (Desurvire, 1994). Assuming the signal is monochromatic with frequency ν_s, the noise power in the signal mode is related through $\tilde{P} = h\nu_s\tilde{n}$. Assuming $V_0 = n_0$, the optical signal-to-noise ratio equals

$$SNR_e = \frac{G_e n_0}{2\tilde{n} + 1 + (M\tilde{n}/G_e n_0)(1 + \tilde{n})} \approx \frac{G_e n_0}{2\tilde{n} + 1} \approx \frac{G_e n_0}{2\tilde{n}} \tag{29}$$

The first approximation is true when the noise power is much less than the signal power, i.e., $G_e n_0/M\tilde{n} \gg 1$, whereas the second approximation is valid only when the added spontaneous emission is significant, i.e., $\tilde{n} \gg 1/2$. The result of the second approximation corresponds to Eq. 22. When the input signal is shot-noise limited and therefore has Poissonian statistics ($V_0 = n_0$), the noise figure equals

$$F = \frac{(1 + 2\tilde{n}) + (M\tilde{n}/G_e n_0)(1 + \tilde{n})}{G_e} \approx \frac{1 + 2\tilde{n}}{G_e} \tag{30}$$

This approximation is valid if the output signal power exceeds the noise power, i.e., $G_e n_0/M\tilde{n} \gg 1$. This is true in almost all cases. Expressing the added noise

in terms of power, the approximated noise figure in Eq. 30 can be rewritten as

$$F = \frac{2\tilde{P}}{G_e h\nu B_0} + \frac{1}{G_e} \tag{31}$$

where $\tilde{P}$ is the spontaneously emitted power within the bandwidth B_0 and in the signal polarization.

3.2.3 A Passive Fiber Followed by a Discrete Amplifier

It was shown above that the mean and variance in photon number of a signal that has traversed a fiber of length L are both $(n_0 \exp(-\alpha_s L))$. By inserting these results into Eqs. 28a and 28b for the statistics of the input light into a discrete amplifier, we find the mean and variance at the output of the amplifier are given by

$$\langle n \rangle = G_e n_0 \exp(-\alpha_s L) + M\tilde{n} \tag{32a}$$

and

$$V_n = G_e n_0 \exp(-\alpha_s L)\{2\tilde{n} + 1\} + M\tilde{n}(\tilde{n} + 1) \tag{32b}$$

where G_e is the gain of the discrete amplifier. Assuming that the signal output power is much larger that the output noise power, $M\tilde{n}/\{n_0 G_e \exp(-\alpha_s L)\} \ll 1$, the noise figure for this configuration equals

$$F = \frac{2\tilde{n} + 1}{G_e \exp(-\alpha_s L)} \tag{33}$$

Since $G_e \exp(-\alpha_s L)$ is the "gain" of the transmission segment, Eq. 33 is the analog to Eq. 30 for a discrete amplifier. Denoting the noise figure of a passive fiber F_f and the gain or loss as $G_f = \exp(-\alpha_s L)$, the total noise figure may be written as

$$F = F_f + \frac{F_e - 1}{G_f} \tag{34}$$

where F_e is the noise figure of the EDFA. This is recognized as the Friss Formula (Carlson, 1986). In fact, this relationship is generally true for any two concatenated components, whether they are passive or active, as we will see below.

By comparing Eqs. 30 and 33, we see that the addition of a passive attenuator, such as a passive optical fiber, to the input to an amplifier increases the noise figure of the total system dB-for-dB with the loss of the attenuator. It is also worth noting that a similar analysis would show that if the passive attenuator were placed after the amplifier, the noise figure of the combination would be dramatically improved compared to the opposite ordering. This is simply because an input attenuator attenuates only input signal photons, whereas an output attenuator attenuates both signal photons and noise photons. Therefore, strictly from noise considerations, it is always better to have gain before loss.

3.2.4 A Raman Amplifier

Next we consider the case of a distributed Raman amplifier. The mean number of photons and their variance at the output of the amplifier is governed by equations similar to Eqs. 28a and 28b:

$$\langle n \rangle = G_R n_0 + M \tilde{n}_R \tag{35a}$$

and

$$V_n = G_R n_0 (2\tilde{n}_R + 1) + M \tilde{n}_R (\tilde{n}_R + 1) \tag{35b}$$

where the number of spontaneously emitted photons $\tilde{n}_R$ and the gain G_R must be determined by numerical integration as described in Section 2.

For the purpose of comparison, it is useful to represent the distributed counterpumped Raman amplifier as a passive fiber followed by an effective discrete postamplifier, as mentioned in Section 3.1. The noise added by the effective discrete amplifier must match the noise generated by the distributed Raman amplifier, and the gain of the effective discrete amplifier must equal the on-off Raman gain. Under these conditions, the effective gain G_{eff} and noise figure F_{eff} of the discrete amplifier are

$$G_{eff} = G_R \exp(\alpha_s L) \tag{36a}$$

and

$$F_{eff} = F_R \exp(-\alpha_s L) \tag{36b}$$

where G_R and F_R are the gain and noise figure of the original distributed Raman amplifier.

In order to demonstrate the typical performance of a distributed Raman amplifier, we compare in Fig. 5.3 the gain and the normalized electrical signal-to-noise ratios for the following three scenarios: a) a discrete amplifier followed by 100 km of passive fiber; b) 100 km of passive fiber followed by a discrete amplifier; and c) a distributed, counterpumped Raman amplifier in 100 km of fiber. In all three cases, the level of amplification offsets the passive loss in order to make the system transparent. For the Raman amplifier, we assume that the signal and pump wavelengths are 1555 nm and 1455 nm, respectively, and that the fiber has a Raman gain coefficient of $0.7\,(\mathrm{Wkm})^{-1}$ for the pump and signal. Furthermore, we assume the loss at the signal and pump wavelengths are 0.2 dB/km and 0.24 dB/km, respectively.

From the signal power evolutions depicted in Fig. 5.3a, one would expect the system with the discrete amplifier first (system a) to have the best output SNR_e, the system with the discrete amplifier last (system b) to have the worst SNR_e, and the SNR_e generated by the distributed Raman amplified system (system c) to fall somewhere between these two. There are two reasons why one would expect this. First, as stated at the end of case III, strictly from noise considerations, it is always better to have gain before loss. Secondly,

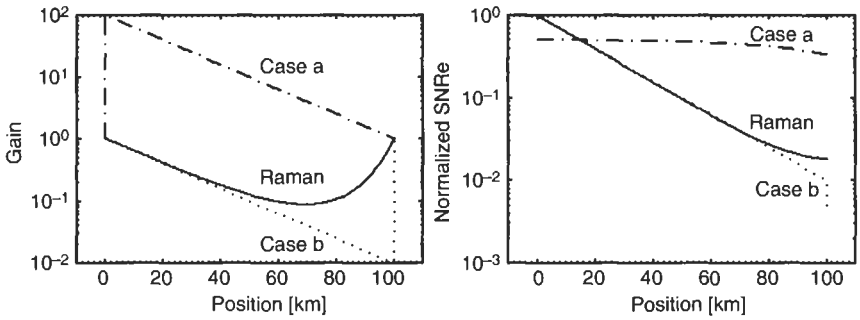


Fig. 5.3 Gain and electrical signal-to-noise ratio (SNR_e) versus position for: a) a discrete amplifier followed by 100 km of passive fiber; b) 100 km of passive fiber followed by a discrete amplifier; and c) a distributed, counterpumped Raman amplifier in 100 km of fiber. The noise figure of the discrete amplifier assumed to be 3 dB independently of the signal power level into the amplifier. The SNR_e's are normalized to the SNR at the signal input end.

when comparing systems with low-noise amplifiers and the same net gain, the system that allows the signal power to drop to the lowest level will produce the worst noise performance. The lost SNR_e of a signal can never be recovered. Once the signal power has dropped to a low level, the SNR_e is already degraded by shot noise and can never be recovered.

From one point of view, the improvement in SNR_e produced by systems with distributed Raman amplifiers over systems with discrete amplifiers is generated simply by the relocation of a fraction of the gain from a position after the span into the last 20 km of the span. This relocation of the gain improves the noise performance by reversing the order of the gain and loss in the last 20 km of fiber.

It is also worth emphasizing the fact that the SNR_e's degrade with increasing distance in all of the plots shown in Fig. 5.3b. Raman amplification does not improve SNR_e's. It simply prevents SNR_e from degrading as quickly as it would have without Raman amplification.

In addition, Fig. 5.3b also illustrates the meaning of the effective noise figure. An effective discrete amplifier located at the output end of the passive fiber would have had to have a noise figure less than 1 (i.e., negative in decibels) in order to match the performance of the distributed Raman amplifier; therefore the effective noise figure may be negative.

After comparing the output SNR_e's in Fig. 5.3, one might wonder why one would not simply use more powerful amplifiers at the beginning of each span and forgo the use of distributed Raman amplifiers completely. In many cases, it is not possible to increase the launch power into the fiber span due to transmission impairments from optical nonlinearities. Under these circumstances, it is advantageous to use distributed Raman amplification. We will return to the subject of signal power evolution and path-average powers in Section 3.3.

3.2.5 A Raman Amplifier Followed by an EDFA

In typical communication systems, a distributed Raman amplifier is followed by an EDFA. We now turn our attention to this case. Building upon the results of the previous case, we find that the mean and variance of the photon number following the EDFA is given by

$$\langle n \rangle = Gn_0 + M\tilde{n} \tag{37a}$$

and

$$V_n = Gn_0(2\tilde{n} + 1) + M\tilde{n}(\tilde{n} + 1) \tag{37b}$$

where G is the total gain G_eG_R, and $\tilde{n}$ is the final output noise $G_e\tilde{n}_R + \tilde{n}_e$. Eqs. 37a and 37b are analogous to those derived for an EDFA. The signal-to-noise ratio in this case may be derived from Eq. 29 by replacing G_e and $\tilde{n}$ with the total gain G_eG_R and the final output noise $G_e\tilde{n}_R + \tilde{n}_e$. We find that the total noise figure is

$$F = \frac{1 + 2\tilde{n}}{G} = F_R + \frac{F_e - 1}{G_R} \tag{38}$$

where G_R and F_R are the net gain (not the on-off gain) and noise figure of the Raman amplifier, and F_e is the noise figure of the erbium amplifier.

Figure 5.4 illustrates the noise power in a 0.1 nm bandwidth in one of the two polarization modes supported by an optical fiber at the output of a transmission segment described by Section V above. The signal is at 1560 nm and

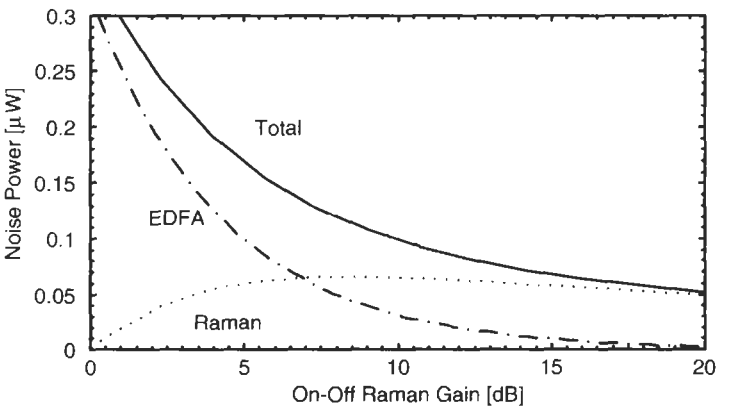


Fig. 5.4 The noise power at the output of a transparent 100 km transmission segment versus the on-off Raman gain from case V of the main text. The noise power is calculated in a 0.1 nm bandwidth and in a single polarization state using simplistic formulas for gain and noise figure (i.e., no pump depletion and no Rayleigh scattering) (Rottwitt, 2000). The noise figure of the EDFA is assumed to be 6 dB for all data points. The pump wavelength is 1455 nm, whereas the signal wavelength is 1560 nm.

with the pump wavelength at 1455 nm the signal is located at the Raman gain peak, which in this case equals 0.7 $(\text{Wkm})^{-1}$. The loss at the pump wavelength is 0.24 dB/km resulting in a Raman effective length of 18 km. The 100 km-long transmission segment contains a Raman amplifier with gain G_R and an EDFA with gain G_e. The sum of the on-off Raman gain and the EDFA gain is kept at 20 dB to exactly compensate for the fiber background loss. The noise power is calculated as a function of the on-off Raman gain. That is, for each data point, the gain provided by the EDFA is adjusted to assure that the sum of the on-off Raman gain and the EDFA gain equals 20 dB. The total output noise power, the solid curve in Fig. 5.4, contains a contribution from the Raman amplifier, the dotted curve in Fig. 5.4, a contribution from the EDFA, the dotted curve in Fig. 5.4. The noise figure of the EDFA is assumed to be 6 dB for all the data points independently of the input signal power.

Figure 5.4 illustrates that for a system with low Raman gain and high EDFA gain, the output noise is dominated by EDFA noise, and the output noise power is relatively high compared to the case when more Raman gain is provided. When the on-off Raman gain exceeds 7 dB in this example, the noise from the Raman amplifier dominates, and the noise from the EDFA is "hidden." This is explained by the high Raman gain, which helps reduce the impact of the noise from the EDFA, as can be seen from Eq. 38.

3.3 SIGNAL EFFECTIVE LENGTH/PATH AVERAGE POWER

In typical high-capacity digital transmission systems, the signal power launched into a fiber span is adjusted to minimize the bit-error rate (BER) of the received signal. Often the BER will worsen not only at lower launched ignal powers due to SNR_e limitations, but also at higher launched signal powers due to optical nonlinear impairments. These nonlinear impairments include pulse distortion arising from self- and cross-phase modulation as well as four-wave mixing. In general, the launched power is adjusted until the complicated interplay between the various sources of noise and distortion results in a minimum BER.

The strength of many nonlinear impairments (e.g., self- and cross-phase modulation) is proportional to the product of the launched signal power and the effective length of the optical signal (Chraplyvy, 1990). We will refer to this length as the signal effective length and denote this as L^s_{eff} to distinguish it from the Raman effective length defined in Eq. 17. In general the signal effective length is defined as the path average power relative to the launched signal power, $P_s(z = 0)$ times the physical length, L:

$$L^s_{eff} = \int_{z=0}^{L} \frac{P_s(z)}{P_s(z=0)}\, dz = L\frac{\langle P(z)\rangle}{P_s(z=0)} \tag{39}$$

In a passive fiber, the signal power decays exponentially, $P_s(z) = P_s^0 \exp(-\alpha_s z)$, and therefore the signal effective length equals $L_{eff}^s = (1-\exp(-\alpha_s L))/\alpha_s$. Note that in this case, the signal effective length is independent of the launched signal power. In a distributed Raman amplifier, one must directly calculate the effective length using Eq. 39, and the signal effective length may depend upon the launched signal power.

The signal effective length of a counterpumped Raman amplifier in the undepleted pump regime with a physical length of 100 km is plotted in Fig. 5.5 as a function of on-off Raman gain. With a background loss of 20 dB, the effective signal length is 21.5 km in the passive fiber. Note that for an on-off gain of 26 dB the effective signal length is increased by a factor of 2. Using these high gains and the same launched signal power, the nonlinear impairments are increased significantly compared to systems utilizing lower on-off Raman gains. However, the improved noise performance of systems utilizing distributed Raman amplification also permits the use of lower launch signal powers. This tradeoff between improved noise figure and nonlinear impairments must be considered carefully when optimizing a system for use with distributed Raman amplification. It is also worth mentioning that the above results are valid only in the undepleted-pump regime. Pump depletion will be treated in Section 4.

We now return to compare the three cases previously examined in Fig. 5.3: a) a discrete amplifier followed by 100 km of passive fiber; b) 100 km of passive fiber followed by a discrete amplifier; and c) a distributed counterpumped Raman amplifier in 100 km of fiber. We compare the evolution of

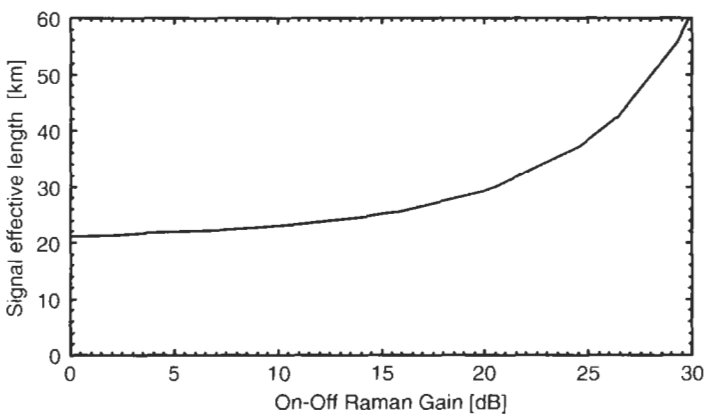


Fig. 5.5 The signal effective length of a Raman amplifier versus on-off Raman gain. The curve is calculated for a signal wavelength of 1560 nm. The Raman gain coefficient is $0.7\,(\mathrm{Wkm})^{-1}$ corresponding to a pump wavelength of 1455 nm. The loss at the pump wavelength is 0.24 dB/km.

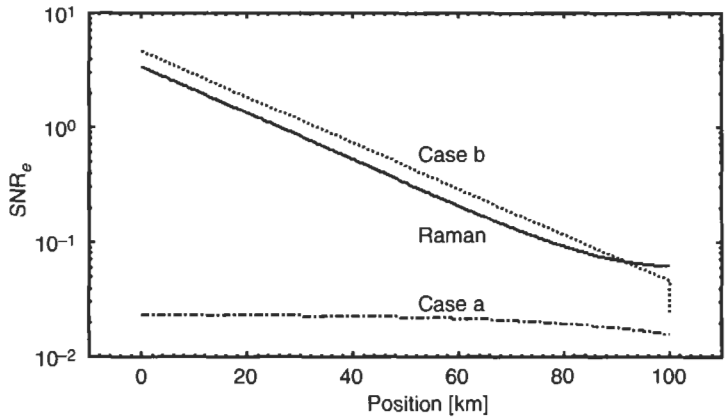


Fig. 5.6 Electrical signal-to-noise ratio as a function of position for the following three cases: (a) a discrete amplifier followed by 100 km of passive fiber; (b) 100 km of passive fiber followed by a discrete amplifier; and (c) a distributed, counterpumped Raman amplifier in 100 km of fiber. The launched signal powers have been adjusted such that all three cases have the same path-average signal power.

the electrical signal-to-noise ratios after we have adjusted the launched signal powers such that all three cases have the same path-averaged signal power (Rottwitt, 1993). The results, depicted in Fig. 5.6, clearly illustrate the benefit of distributed Raman amplification when a system is limited by optical nonlinearities.

Figure 5.6 demonstrates that the distributed Raman amplifier performs better than either of the two cases with discrete amplifiers when compared on equal path-averaged power and hence presumably equal nonlinear impairments.

4. Constraints of Real-World Raman Amplifier

The previous sections have largely treated Raman amplifiers in the undepleted-pump regime and where both the pump and signals are monochromatic. In order to model communications systems with multiple signal and pump wavelengths, a large set of coupled differential equations must be solved. This task becomes more complicated if pump depletion, crosstalk, Rayleigh scattering, and temperature effects are included. In this section, we provide an introductory description of these phenomena. Section 4.1 focuses on pump depletion, Section 4.2 on crosstalk effects, Section 4.3 deals with the impact from Rayleigh scattering, and finally Section 4.4 describes the temperature dependence.

4.1 PUMP DEPLETION

In many realistic Raman amplifiers, the rate at which pump power is lost exceeds the exponential decay rate originating from the intrinsic fiber background loss. This phenomenon occurs when a significant fraction of the pump power is transferred to the signal via Raman amplification. For example, this may occur when amplifying a large number of signals, or a signal of very high power, or when the amplifier is providing a very high gain. Under such conditions the simple model of gain and noise performance described in Section 2.2 is not appropriate.

In the photon model of the Raman amplifier described in Section 2, considering n_p pump photons at frequency ν_p and n_s signal photons at frequency ν_s, each photon may be accounted for in the case when the intrinsic fiber loss is neglected, that is, every pump photon that scatters is transferred to a signal photon. When the loss at the pump and signal frequency are assumed to be the same, each photon may also be accounted for when the amplifier is forward pumped, since any photon added to the signal is a Raman scattered photon.

Applying this to Eq. 7 for signal photons and its counterpart for pump photons, one arrives at the output signal photon number

$$n_s(z) = n_s^0 \frac{1+r}{r + G^{-(1+r)}} \exp(-\alpha z) \tag{40}$$

where r is the ratio of signal photons to pump photons at launch $r = n_s^0/n_p^0$ and G the on-off Raman gain at position z. See Auyeung, 1978 and Agrawal, 1995.

When using real values for the intrinsic fiber loss, the effects of depletion must be evaluated numerically. Figure 5.7 illustrates the effects of pump depletion in a counterpumped fiber Raman amplifier. Plotted in Fig. 5.7a is the on-off Raman gain versus signal input power, and the corresponding effective noise figure is shown in Fig. 5.7b. The figures show three curves, each calculated for fixed launched pump powers of 150 mW, 300 mW and 600 mW.

Figure 5.7 illustrates that the on-off Raman gain decreases as the launched signal becomes sufficiently powerful. This can be explained by considering the rate of loss of the pump. At any position along the fiber, the intrinsic fiber loss and the product of the signal power and the Raman gain coefficient determine the rate of loss of the pump. For a sufficiently powerful signal, the rate of pump loss will be dominated by the term including the signal power times the Raman gain coefficient, i.e., the Raman scattering. As the pump is more quickly attenuated, the effective length of the Raman interaction is reduced, reducing the on-off Raman gain.

Figure 5.7 also illustrates that the noise figure of the Raman amplifier will increase as the pump is depleted. As the length of the Raman interaction is reduced, the signal power is allowed to drop to a lower minimum value within the span, resulting in a higher span noise figure.

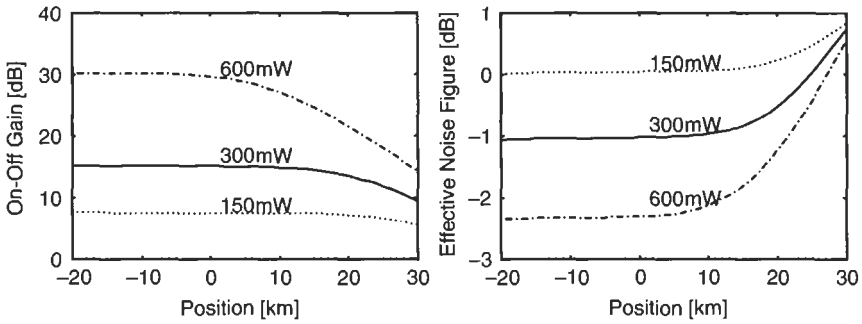


Fig. 5.7 On-off Raman gain and corresponding effective noise figure versus input signal power for three pump powers. Each data point is calculated for a 100 km long counterpumped Raman amplifier. The signal wavelength is 1555 nm and the pump wavelength is 1455 nm. The loss coefficient at the pump wavelength is 0.25 dB/km, creating an effective Raman length of 17 km, and the Raman gain coefficient is $0.7\,(\mathrm{Wkm})^{-1}$.

4.2 CROSSTALK IN RAMAN AMPLIFIERS

In this section, we will refer to any coupling of amplitude modulation between two wavelengths in a Raman-amplified, wavelength-division-multiplexed system as crosstalk. There are three basic types of Raman-related crosstalk in lightwave communication systems: interchannel crosstalk, pump-signal crosstalk, and signal-pump-signal crosstalk.

4.2.1 Interchannel Crosstalk

This type of crosstalk occurs due to Raman scattering between signals alone, without any involvement with a Raman pump. When two or more signals propagate through an optical fiber, they will interact through Raman scattering. The channels with longer wavelengths will experience gain at the expense of the channels with shorter wavelengths. For systems with many closely spaced channels, this effect will tend to tilt the output power spectrum. In order to illustrate this effect, we plot in Fig. 5.8 the differential gain experienced by 20 channels propagating through a 100 km length of transmission fiber under typical conditions. In this simulation, we treated the channels as continuous wave signals. We will return to the issue of temporal dependence below.

In this example, the channels with the longest wavelengths experience approximately 0.3 dB of gain at the expense of the shortest wavelength channels, which experience 0.3 dB of loss. The interchannel Raman gain depends strongly upon channel power, wavelength separation and count.

In order for the signals to interact through Raman scattering, they must overlap in space and time. Thus, in a digital communication system where each channel is carrying bits (i.e., marks and spaces), the interchannel Raman gain

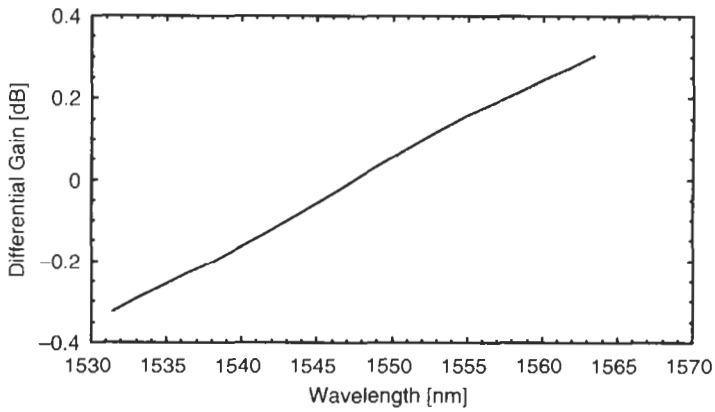


Fig. 5.8 Interchannel Raman gain among 20 channels with average power of 1.5 mW per channel. The channels are launched into a 100 km length of transmission fiber with peak Raman gain coefficient $\sim 0.7\,(\mathrm{Wkm})^{-1}$ and 0.2 dB/km loss. No Raman pump is applied.

clearly depends on the bit pattern of each channel. The crosstalk is further complicated since channels are experiencing different group velocities due to group-velocity dispersion of the transmission fiber. The group-velocity mismatch leads to a "walkoff" between the channels. The walkoff is beneficial in the sense that it averages the interchannel Raman gain over many bits. In systems with many channels, the result of the walkoff and the large channel count is that the interchannel Raman gain becomes almost deterministic (Forghieri, 1995; Tariq, 1995). For this reason the differential gain illustrated in Fig. 5.8 is calculated using continuous-wave signals with 1.5 mW of power per channel at launch. Due to the walkoff and large channel count, this representation is suitable for an information-carrying multichannel signal with 1.5 mW of average signal power per channel.

In Fig. 5.8 the Raman gain peak is assumed to equal $0.7\,(\mathrm{Wkm})^{-1}$. This value is measured on a Lucent TrueWave type fiber using an unpolarized pump source. In the case of interchannel crosstalk, the transmitted signals that participate in the crosstalk typically have well defined polarization. Hence, in long fibers that do not maintain polarization, the true Raman coupling is a complicated average over the entire transmission since the state of polarization for each signal channel changes independently during propagation.

Interchannel Raman gain is also present when a Raman pump is utilized. However, the only relevant effect due to the presence of a Raman pump is that the signals do not decay exponentially, and therefore the effective signal lengths of individual channels is modified as described in Section 3.3. For further details see Forghieri, 1995.

4.2.2 Pump-Signal Crosstalk

One particularly important source of noise in Raman amplifiers is the direct coupling of amplitude noise from the pump to the signal through the Raman gain. The root cause of this source of noise is the extremely fast response time of the Raman process. Unlike EDFAs, Raman amplifiers do not have a long upper-state lifetime to buffer the gain from fluctuations in pump power. Pump-signal crosstalk can be dramatically reduced by the use of a counterpropagating pump geometry. The use of a counterpropagating pump uses the transit time through the effective length of the amplifier to average fluctuations in the pump power. The effectiveness of this approach is depicted in Fig. 5.9. Here, a 1240 nm pump generated by a fiber laser is used to amplify a 1310 nm signal. Note that the amplitude noise of the pump source is effectively transferred to the signal in the copropagating pump geometry but is dramatically reduced in the counterpropagating geometry. This noise source is one of the main reasons that Raman amplifiers are almost exclusively counterpumped.

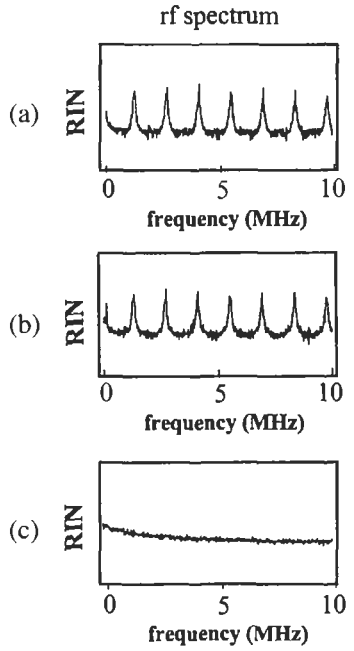


Fig. 5.9 Electrical power spectrum of (a) a 1240 nm pump; (b) a 1310 nm signal after amplification by the 1240 nm pump in a copropagating pump geometry; and (c) in a counterpropagating pump geometry.

4.2.3 Signal-Pump-Signal Crosstalk

We now turn to the subject of signal-pump-signal crosstalk, also known as pump-mediated crosstalk, as described in Mahlein, 1984; D. Cotter, 1984; Jiang, 1989; and Forghieri, 1994. When a signal becomes sufficiently powerful to deplete a Raman pump, amplitude modulation of the signal may be impressed upon the pump via the depletion process. If the Raman pump is simultaneously amplifying multiple signals, the fluctuations impressed upon the pump may be transferred to another signal via the Raman gain. This form of crosstalk is known as signal-pump-signal crosstalk. The efficiency of the coupling depends on the walkoff between the pump and the signals as well as the amount of Raman gain and the degree of depletion. The effect is greatly enhanced if the signals and pump are copropagating. However, as for the interchannel Raman gain, walkoff and large channel count diminish the effect by averaging. This is discussed in further detail in Forghieri, 1995.

When the Raman pump and signals are counterpropagating, the noise penalty due to signal-pump-signal crosstalk is greatly reduced. As opposed to the copropagating geometry where the pump and signals propagate with almost the same group velocity, a counterpropagating geometry has the pump and signals propagating relative to one another at twice the speed of light, effectively diminishing crosstalk at all but the lowest frequencies. Forghieri, 1994, illustrates the large benefit in a counterpropagating pump geometry with respect to signal-pump-signal crosstalk. In Forghieri, 1994, one signal is launched into a Raman amplifier as a sinusoidal modulated signal, whereas another signal is launch as a continuous wave (CW). Figure 5.10 illustrates the induced modulation on the CW tone. The effect of the averaging when pump and signal propagate in opposite directions is clearly shown. In the counterpumped Raman amplifier, the induced modulation strongly diminishes for modulation frequencies above 1 kHz. In the copumped Raman amplifier, the modulation frequency has to exceed 10 MHz before the induced modulation vanishes.

4.3 RAYLEIGH SCATTERING

As opposed to Raman scattering, Rayleigh scattering is an elastic process, and therefore the scattered light has the same frequency as the incoming light. Rayleigh scattering varies as λ^{-4} and is largely responsible for the intrinsic loss at wavelengths less than 1600 nm. At wavelengths greater than 1600 nm, the fiber loss is dominated by material and bending losses. A minimum is located near 1.55 μm where the intrinsic loss is close to 0.2 dB/km.

A portion of the Rayleigh scattered light is recaptured in the fiber. Half of this light propagates in the forward direction; half in the backward direction. The backward propagating light recaptured by a fiber due to Rayleigh scattering of power P_s in length dz is $B\alpha_s^R P_s dz$, where α_s^R is the Rayleigh scattering loss and B is the recapture fraction. The recapture fraction may be calculated based

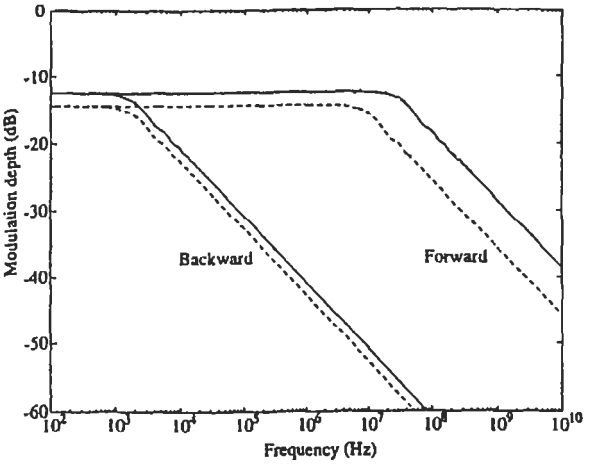


Fig. 5.10 Signal crosstalk in forward- and backward-pumped Raman amplifiers using dispersion shifted (solid curves) and conventional (dashed curves) fiber. The dispersion-shifted fiber has effective area of 55 μm^2 and zero dispersion at 1560 nm, whereas the conventional fiber has effective area of 85 μm^2 and zero dispersion at 1300 nm. The length of the Raman amplifier is 30 km and the Raman gain is chosen exactly to counterbalance the intrinsic fiber loss. The average input signal power is 5 mW. (Reprinted with permission from Forghieri, 1994.)

on fiber design (Hartog, 1984). The product $B\alpha_s^R$ is the Rayleigh-backscatter coefficient. Experimentally, the Rayleigh backscatter coefficient may be easily determined by measuring the light reflected from a very long length of fiber of known loss, $P_R(z = 0)$, using the following relationship:

$$\frac{P_R(z=0)}{P^0} = \frac{B\alpha_s^R}{2\alpha}\left(1 - \exp(-2\alpha_s L)\right) \rightarrow \frac{B\alpha_s^R}{2\alpha}, \quad \text{for } L \rightarrow \infty \tag{41}$$

where P^0 is the power of the launched light.

In standard dispersion-shifted fibers with $\alpha_s = 0.2$ dB/km, the Rayleigh reflected power is typically 30 dB lower than the launched light for an infinitely long fiber. Using Eq. 41, we find a Rayleigh backscatter coefficient of 10^{-4} per km. The value is only approximate and a value specific for a given fiber design needs to be determined for accurate modeling.

Rayleigh scattering may have two significant effects on the noise performance of a Raman amplifier. First, ASE that was initially counterpropagating relative to a signal may backscatter, adding to the noise in the system. Second, a signal may undergo two, or any even number, of Rayleigh backscattering events and contribute to the system noise in the form of multiple-path interference (MPI). The latter effect is commonly referred to as double Rayleigh scattering. A particularly troublesome aspect of double Rayleigh scattering is that the MPI contribution to the noise has the same optical frequencies as

the signal, making it impossible to measure with the simplest spectral measurements. In order to quantify the amount of double Rayleigh scattering, one needs to use alternative methods such as electrical (Movassaghi, 1998) or time-extinction measurements (Lewis, 2000). These effects were first characterized in distributed Raman amplifiers by Hansen (1998a).

Using the simple Raman model described in Section 2 and assuming that the powers of the Rayleigh-reflected ASE and double-reflected signal remain relatively low, we may derive simple expressions for these powers at the output of the fiber. For the remainder of this section, we will assume that the pump is counterpropagating relative to the signal and that the pump is injected into the fiber at $z = L$. Denoting the ASE powers that are co- and counterpropagating relative to the signal as P_{ASE^+} and P_{ASE^-}, respectively, we find that the Rayleigh-reflected portion of P_{ASE^-} at the end of the fiber into which the pump is injected is given by

$$P^R_{ASE^-}(L) = G(L) \int_{z=0}^{L} \frac{\left(B\alpha_s^R\right) P_{ASE^-}(z)}{G(z)}\, dz \tag{42}$$

where $G(x)$ is the net gain, not the on-off gain. In Fig. 5.11, we compare the reflected ASE power $P^R_{ASE^-}$ to the unreflected power P_{ASE^+} at the end of the fiber into which the pump is injected. In this example, a 100 km length of transmission fiber is pumped at 1455 nm, and a Raman gain coefficient of $0.7\,(\mathrm{Wkm})^{-1}$ and a Rayleigh backscatter coefficient $10^{-4}\,\mathrm{km}^{-1}$ are used.

Figure 5.11 illustrates two regimes of operation. For on-off Raman gains less than 20 dB, the Rayleigh-reflected ASE is significant less than the forward

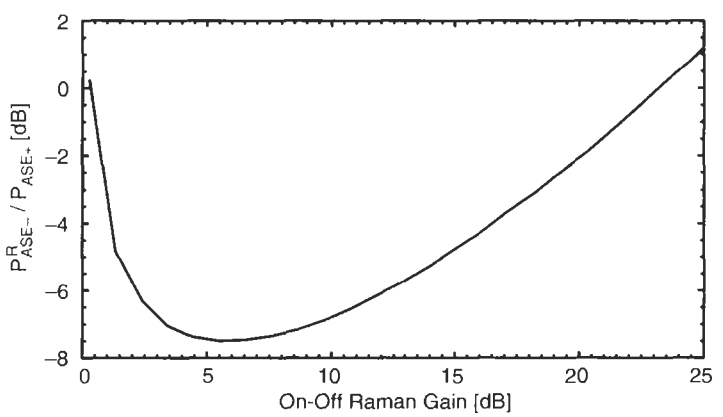


Fig. 5.11 Ratio of reflected ASE $P^R_{ASE^-}$ to forward-propagating ASE P_{ASE^+} at the fiber end into which a Raman pump is injected versus on-off Raman gain. A 100 km length of fiber with peak Raman gain coefficient of $0.7\,(\mathrm{Wkm})^{-1}$ and loss coefficient for the signal and pump of $\alpha_s = 0.2\,\mathrm{km}^{-1}$ and $\alpha_p = 0.25\,\mathrm{km}^{-1}$ were used.

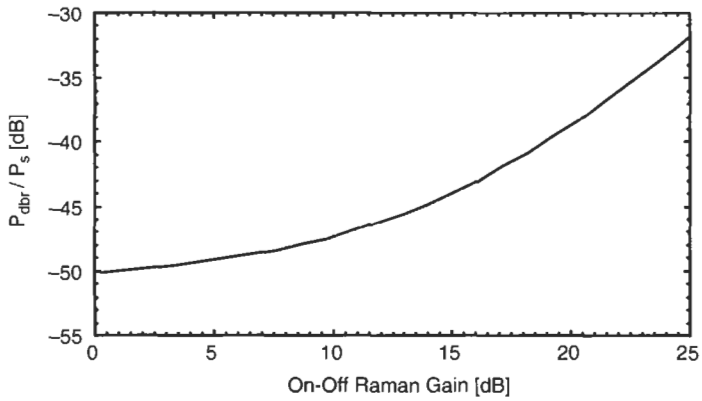


Fig. 5.12 Ratio of double reflected power relative to signal power versus on-off Raman gain in a 100 km long backward-pumped Raman amplifier. Same parameters as those used in Fig. 5.11.

propagating ASE, whereas for on-off Raman gains greater than 23 dB, the Rayleigh-reflected ASE exceeds the forward-propagating ASE.

The double Rayleigh reflected signal power at $z = L$ is

$$P_{dbr}(L) = G(L)P_s^0(B\alpha_s^R)^2 \int_{\tilde{z}=0}^{L} \frac{1}{G^2(\tilde{z})} \int_{\tilde{z}}^{L} G^2(x)dx\, d\tilde{z} \tag{43}$$

where again $G(x)$ is the net gain, not the on-off Raman gain. Plotted in Fig. 5.12 is the ratio power of the double-reflected signal P_{dbr} relative to the unscattered signal at the end of the transmission fiber, $P_s = P_s^0 G(L)$.

The double-reflected signal power from a discrete amplifier with gain G surrounded by two discrete reflectors is proportional to G^2 (Wan, 1995). In this situation, the ratio of the double-reflected power to the signal power at the output of the amplifier is proportional to G and therefore will increase dB-for-dB with the net gain of the amplifier. Note that in Fig. 5.12 the slope of this ratio with respect to on-off Raman gain is 1.5 dB for on-off Raman gains close to 20 dB.

To fully evaluate the impact of double Rayleigh scattering on lightwave communication systems, one needs to translate the above results into BER penalties. This requires detailed knowledge of the transmitters and receivers utilized in a system and is outside the scope of this chapter. We refer the reader to Takahashi et al. (1996) and Rasmussen et al. (1999) for further reading.

4.4 TEMPERATURE DEPENDENCE OF RAMAN AMPLIFICATION

The temperature dependence of Raman scattering must be included in any modeling of Raman amplifiers if an accurate prediction of noise performance

is required at temperatures above zero Kelvin. As will be shown below, the model developed in Section 2 is perfectly capable of accurately treating the temperature dependence of Raman amplifiers.

According to the model, the signal gain is independent of temperature, whereas the spontaneous emission is temperature dependent. This agrees with measurements reported by Lewis et al. (1999). In Fig. 5.13 is plotted the measured (left) and predicted (right) effective noise figures of a 15 km-long dispersion-shifted fiber versus on-off Raman gain. The measurements were recorded at room temperature and with the on-off Raman gain varied from 2 dB to 26 dB in 2 dB increments. By cooling the same fiber to 77 K using liquid nitrogen and repeating the experiment, the results plotted in Fig. 5.14 were obtained.

Both Figs 5.13 and 5.14 demonstrate that the theoretical model accurately predicts the experimental results. The small deviations, especially at high values of the Raman gain, are due to spurious end-reflections.

The temperature dependence is strongest on wavelengths closest to the pump wavelength. Tracing a loop on one of the curves, starting from wavelengths close to the pump, corresponding to zero on-off gain in Figs 5.13 and 5.14, the effective noise figure increases drastically up to a global peak, after which the effective noise figure drops, and the gain peaks at a wavelength that corresponds to a frequency close to 13 THz below the pump. Then, for

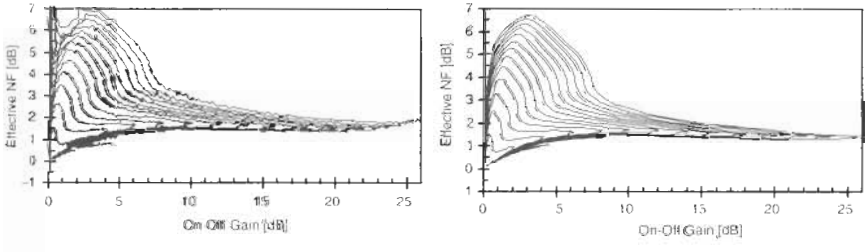


Fig. 5.13 Effective noise figure versus on-off Raman gain, at room temperature (295 K). Left: measured; right: predicted. The pump wavelength was 1455 nm.

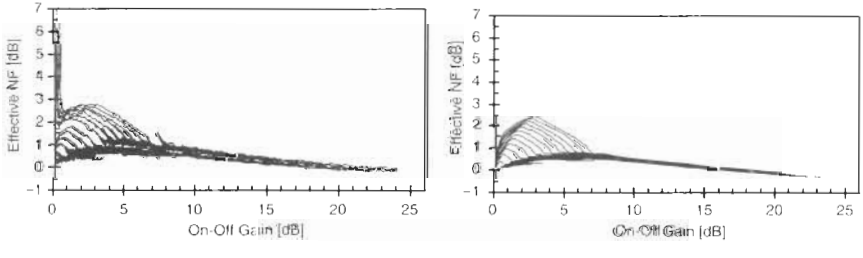


Fig. 5.14 Effective noise figure versus on-off Raman gain at 77 K. Left: measured; right: predicted.

even longer wavelengths, the gain and the noise figure roll off to zero. There are two reasons why the curves make a loop. One is the wavelength dependence of the intrinsic fiber loss, and the other, and most dominant, effect is the wavelength dependence of the temperature dependence of the spontaneous emission.

For completeness it should be noted that for an amplifier installed in the field, providing significant gain in commonly used signal wavelengths bands, the noise performance changes by only a few tenths of a dB when the temperature varies from −25°C to 75°C (Rottwitt, 2000).

5. Pump Sources

One can easily solve Eq. 16 for the pump power required for the generation of a given amount of Raman gain in the undepleted pump regime. For instance, in typical transmission fibers with a Raman effective area of 55 μm^2 and where $g \sim 0.7(\text{Wkm})^{-1}$ for a depolarized pump, one requires ~300 mW of pump power to generate 15 dB of on-off Raman gain. The required pump power increases to ~500 mW in standard single-mode fiber and increases further still if the amplifier is operated in depletion. Clearly, Raman amplification in standard communication fibers requires substantial pump power.

Given the strong polarization dependence of the Raman process, it is beneficial to use a polarization-multiplexed or depolarized pump source in order to eliminate any polarization-dependent gain with pump polarization diversity.

Finally, given that the gain bandwidth of a Raman amplifier can be broadened and flattened through the use of multiple pump wavelengths, it is advantageous for the Raman pump source to contain a number of pump wavelengths. In this section, we briefly describe two pump sources that have been frequently used to pump Raman amplifiers: multiplexed semiconductor laser and cascaded Raman fiber lasers.

5.1 MULTIPLEXED SEMICONDUCTOR LASERS

Depicted in Fig. 5.15 is a schematic illustration of a set of polarization- and wavelength-multiplexed semiconductor lasers. Typically the lasing wavelengths of the lasers are determined by a short-period fiber Bragg grating that is spliced to the output fiber pigtails of the semiconductor pumps. Bragg gratings guarantee the wavelength stability of the pump source and therefore the stability of the Raman gain shape. This type of pump source certainly meets the requirements for sufficient output power, polarization diversity, and multiple pump wavelengths.

The largest pump source yet constructed with this technology utilized twenty-four pump diodes at twelve wavelengths, producing 2 W of output power and generating a very flat, 100 nm-wide Raman gain spectrum (Emori, 1999).

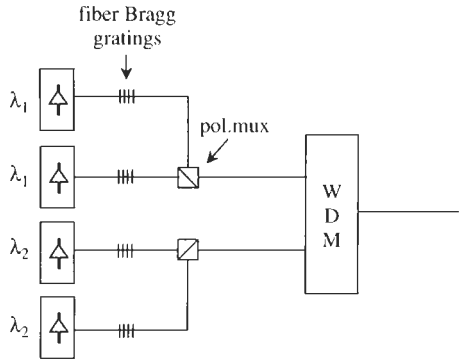


Fig. 5.15 Schematic illustration of a Raman pump source constructed with wavelength-stabilized semiconductor pump lasers that are polarization- and wavelength-division multiplexed.

5.2 CASCADED RAMAN FIBER LASERS

An alternative pump source, known as a cascaded Raman fiber laser, relies on Raman amplification itself. The device is depicted schematically in Fig. 5.16. Intense pump light is injected into the Raman laser at a wavelength easily generated by a ytterbium-doped, cladding pumped fiber laser. The initial pump wavelength emitted from the ytterbium laser is typically near 1100 nm. The presence of the intense pump light in the length of small core germanosilicate fiber generates Raman gain at 1155 nm. A low-loss resonator is created at that wavelength with fiber Bragg gratings. When light begins lasing in the 1155 nm resonator, it serves to pump another fiber laser at 1218 nm. This light pumps another laser at 1288 nm and so forth, until light of the desired wavelength is created. In the example shown in Fig. 5.16, the desired light at 1455 nm is coupled out of the device with an output grating whose reflectivity is substantially less than 100%. Since the wavelengths of the initial pump light and the intermediate lasers may be substantially tuned without significant degradation in the conversion efficiency of the device, any wavelength between 1100 nm and 1650 nm may be generated. The device has demonstrated extremely high efficiencies, converted light from 1100 nm to 1455 nm with 55% slope efficiencies[3] (i.e., 70% slope efficiency in converting input photons to output photons), and the entire device can be spooled on a relatively compact spool of fiber. Recently, this technology has been extended to simultaneously produce three power-stabilized output wavelengths using tunable fiber Bragg gratings (Mermelstein, 2001).

Similar devices to those shown in Fig. 5.16 have been created using fused fiber couplers and rings to create the lasing cavities (Chernikov, 1995, 1998). Also Dianov (1997) has shown that by using phosphosilicate fiber, Raman

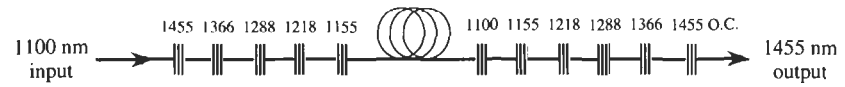


Fig. 5.16 Schematic illustration of a cascaded Raman fiber laser. Input pump light at 1100 nm is converted to output light at 1455 nm through a series of nested Raman fiber lasers. Low-loss and highly wavelength-selective fiber Bragg gratings are used to define laser cavities.

frequency shifts of ~40 THz may be efficiently achieved, eliminating all but two of the Raman fiber lasers.

6. Published Results

In this section, we review selected results from the literature. Certainly we cannot describe all of the results found in the literature. A more complete selection of some of the many excellent papers on Raman amplification can be found in the reference list at the end of this chapter.

6.1 EARLY WORK

After the initial measurement of the Raman gain spectrum in germanosilicate fiber by R. H. Stolen in 1973, interest in Raman amplification in optical fibers remained relatively low for the next decade. During this period, neither low-loss transmission fibers nor high-power pump lasers were widely available. However, in the mid-1980s, with the widespread deployment of optical fibers, Raman amplification became a popular research topic, as engineers searched for a means to extend the distances between electronic repeaters.

In the early 1980s, Aoki (1983, 1985) and coworkers demonstrated the potential of high-gain, low-noise Raman amplification. In these experiments, the Raman pump source was a YAG laser operating at 1.32 μm. Experiments were conducted with the laser operating either CW, Q-switched, or mode-locked. An InGaAsP laser diode operating at 1.4 μm was developed to produce a signal at an appropriate wavelength for the 1.32 μm Raman pump. The transmission fibers had 1.1 and 2.5 dB/km losses at the pump and signal wavelengths, respectively. In one set of experiments, on-off gains as large as 21 dB were generated with net gains of 10 dB (Aoki, 1983). In another set of experiments, the first evaluation of the bit-error rate of signals amplified with a Raman amplifier was conducted. An on-off Raman gain of 14 dB of a 100 Mb/s signal was obtained with a 1 dB power penalty.

In the mid- to late 1980s, work on Raman-amplified systems continued with improved pump and signal sources. Hegarty (1985) and colleagues demonstrated the first Raman amplification of a signal in the low-loss region of optical fibers near 1.5 μm. A color-center laser was used to generate pump light at 1470 nm, and a net increase in receiver sensitivity of ~3 dB/100 mW of pump

power was demonstrated by bit-error rate measurements of a 1 Gb/s signal. Edagawa (1987) and coworkers demonstrated the first use of semiconductor lasers as pump sources for Raman amplifiers. The outputs of a pair of semiconductor pump lasers were polarization multiplexed to produce a total pump power of 60 mW at 1470 nm. This light was used to amplify three wavelength-division-multiplexed signals. In 1988, Mollenauer and Smith used a color-center laser producing light at 1497 nm as a Raman pump source to compensate for the loss in a recirculating loop. Soliton pulses were injected into the loop and were transmitted over 4000 km (Mollenauer, 1988). Finally, in 1988, Aoki and coworkers studied distributed preamplifiers and discrete booster amplifiers. The impact of noise from pump-signal crosstalk was discussed, and the concept of an effective discrete Raman amplifier was introduced (Aoki, 1988).

In 1988, groups at the University of Southampton and Bell Laboratories simultaneously reported the efficient operation of low-noise erbium-doped fiber amplifiers (Mears, 1987; Desurvire, 1987). This discovery dramatically decreased interest in Raman amplification in optical fibers as a viable commercial technology for several years.

6.2 DISCRETE AMPLIFIERS

In the early 1990s, the primary focus on Raman scattering was as a source of nonlinear impairment to optical transmission systems. However, by the mid-1990s, the development of compact high-power pump sources and the desire to develop optical amplifiers with the flexibility to be tuned to new wavelength regimes generated new interest in Raman amplification.

Substantial interest in discrete Raman amplifiers at 1300 nm was revived by the results of Grubb and coworkers in 1994 and subsequently Chernikov and coworkers in 1995. These efforts relied on cascaded Raman fiber lasers to produce the requisite 1240 nm pump light and focused on the improved power conversion efficiency of the amplifiers. Grubb directly passed the signal through a cascaded Raman fiber laser similar to the device depicted in Fig. 5.16 but terminating at the pump wavelength of 1240 nm. Unfortunately, these devices, although quite efficient, suffered from extremely poor noise performance due to pump-signal crosstalk and double Rayleigh scattering. Fibers highly doped with germanium were developed to further improve the power conversion efficiency of the discrete amplifiers (Dianov, 1995).

Unlike typical distributed Raman amplifiers, discrete Raman amplifiers are useless unless they generate a substantial amount of net Raman gain. However, when generating a substantial amount of gain, discrete Raman amplifiers become susceptible to penalties from double Rayleigh scattering. Recognition of the role of double Rayleigh scattering as a dominant noise source in discrete Raman amplifiers led to the development of the first high-gain, low-noise discrete Raman amplifiers in 1996. Stentz (1996) reported the construction of a

1.3 μm analog-grade Raman amplifier with an external noise figure of 4.3 dB. This amplifier utilized a novel ring design with intracavity pumping. The pump light counterpropagated through a two-stage amplifier with an interstage isolator. The strictly counterpropagating pump geometry eliminated pump-signal crosstalk, and the multistage design with an interstage isolator dramatically reduced the noise from double Rayleigh scattering. A similar design was constructed as a preamplifier that produced a peak gain of 41 dB with a noise figure of 4.2 dB as well as a booster amplifier with an output power of 18 dBm. Nielsen (1998) and coworkers used these amplifiers to demonstrate the unrepeatered transmission of eight channels at 10 Gb/s over 141 km of fiber at 1300 nm. Discrete Raman amplifiers were also constructed at new wavelengths. A 1400 nm Raman ring amplifier was demonstrated (Srivastava, 1998), as well as a multistage amplifier at 1520 nm pumped by semiconductor pumps (Kani, 1998). Although these amplifiers were successful, the design of discrete Raman amplifiers that produce high gains at high signal powers when pumped by reasonable pump powers remains a challenge due to the limitations of double Rayleigh scattering.

Hansen (1998b) and coworkers demonstrated that a modest amount of Raman pump power was required to offset the loss of a dispersion-compensating module. The small effective-core areas and high germanium concentration of dispersion-compensating fibers make these fibers very efficient Raman gain media. Simply pumping the fibers to a transmission of unity provides substantial system benefits. Emori (1998) and coworkers extended this concept to an arrangement with four pump wavelengths and eight pump lasers. The loss of the dispersion-compensating fiber was compensated to within 0.5 dB over a 50 nm bandwidth.

Another exciting area of research has focused on the use of hybrid Raman-erbium amplifiers with extended gain bandwidths. Typically a Raman pump with a wavelength near 1480 nm was utilized to generate Raman gain at wavelengths longer than the conventional C-band of an erbium-doped fiber amplifier. These hybrid experiments were conducted with both distributed and discrete Raman amplifiers. Bandwidths as large as 108 nm were generated (Musada, 1998).

Recent work on discrete Raman amplifiers has focused on the S-band, the wavelength band immediate below the conventional C-band of EDFAs, where the loss of transmission fibers is still reasonably low. Works by Bromage (2001) and coworkers and Puc (2001) and coworkers have demonstrated multispan transmission of dense wavelength-division-multiplexed signals in this wavelength range.

6.3 RENEWED INTEREST IN DISTRIBUTED AMPLIFIERS

In the mid-1990s, interest in extending the reach between electronic regenerators in repeaterless communication systems revived interest in distributed

Raman amplification. In the same timeframe, cascaded Raman fiber lasers and semiconductor pump lasers with higher output powers became available as Raman pump sources. A commonly employed technique was to remotely pump a small piece of erbium-doped fiber. The action of the remote pumping provided additional gain not only from the erbium-doped fiber but also from distributed Raman amplification. For instance, Hansen (1995) and coworkers demonstrated transmission of 2.5 Gb/s over 529 km without electrical power in the transmission line through the use of remotely pumped erbium-doped fiber and distributed Raman amplification.

Over the last few years, interest in Raman amplification has steadily increased. Experiments have demonstrated the utility of distributed Raman amplification for increasing fiber span losses, extending system reach, decreasing optical nonlinearities, and increasing system capacity. Hansen (1997) and coworkers demonstrated a power budget increase of up to 7.4 dB through the use of distributed Raman amplification. This improvement allowed system capacities to be increased four-fold through either time- or wavelength-division multiplexing. Nissov (1997) and coworkers demonstrated transmission over 7200 km of fiber using nearly all Raman gain. A recirculating loop with 11 Raman-pumped spans and one erbium-doped fiber amplifier was utilized. Ma (1998) and coworkers demonstrated transmission over 5280 km with spans of 240 km length through the use of locally and remotely pumped erbium-doped fiber amplifiers and distributed Raman amplification. Finally, Hansen (1999) and coworkers and Takachio (1999) and coworkers simultaneously demonstrated the use of distributed Raman amplification to reduce system impairments from optical nonlinearities. The system power budget improvement provided by Raman amplification was utilized to reduce the launched signal power and therefore reduce system impairments from nonlinearities such as self- and cross-phase modulation and four-wave mixing. Ironically, the use of a particularly optical nonlinearity can be used to decrease the impairments caused by other optical nonlinearities.

Very recently, interest in transmission at 40 Gb/s line rates has further increased interest in distributed Raman amplification. Without Raman amplification, it is very difficult to achieve the optical signal-to-noise ratios required for 40 Gb/s transmission through common span losses. In 1999, Nielsen and coworkers demonstrated the transmission of 40 channels of 40 Gb/s over four spans of 100 km of fiber. This was the first experiment to demonstrate 40 Gb/s transmission over multiple 80–100 km long spans without optical time-domain multiplexing. These results were made possible by the extensive use of distributed Raman amplification and were extended to 80 channels at 40 Gb/s in 2000 (Nielsen, 2000).

It is now common for ultra-high-capacity transmission-system experiments to utilize distributed Raman amplification, and the commercial applications of these results are following quickly behind.

7. Conclusions

Over the past thirty years, Raman amplification in optical fibers has made the transition from a laboratory curiosity to a researcher's tool for hero experiments, and finally to an essential feature of every ultra-high-capacity fiberoptic communication system. Undoubtedly with the development of improved pump sources and optimized optical fibers, Raman amplification will continue to find new applications as it expands the ultimate capabilities of fiberoptic communication systems.

Acknowledgments

Karsten Rottwitt is supported financially by the Danish Technical Research Council and OFS Fitel Denmark.

Notes

1. Analytical results are obtainable assuming the same intrinsic loss for pump and signal (Desurvire, 1986).
2. There exists also an optical signal-to-noise ratio (Desurvire, 1994). However, to avoid confusion between this and the ratio of signal power to ASE power in some bandwidth we will restrict this chapter to the electrical signal-to-noise ratio.
3. Personal communication with C. Headley, Lucent Technologies.

References

Agrawal, G. P., *Nonlinear Fiber Optics*, Academic Press, San Diego, 1995.

Aoki, Y., "Properties of fiber Raman amplifiers and their applicability to digital optical communication systems," J. of Lightwave Technol., v. 6, p. 1225, 1988.

Aoki, Y., Kishida, S., Honmou, H., Washio, K., and Sugimoto, M., "Efficient backward and forward pumping cw Raman amplification for InGaAsP laser light in silica fibres," Electron. Lett., v. 19, p. 620, 1983.

Aoki, Y., Kishida, S., Washio, K., and Minemura, K., "Bit error rate evaluation of optical signals amplified via stimulated Raman process in an optical fibre," Electron. Lett., v. 21, p. 191, 1985.

Auyeung, J., and Yariv, A., "Spontaneous and stimulated Raman scattering in long low loss fibers," IEEE J. of Quantum Electron., v. QE-14, p. 347, 1978.

Becker, P. C., Olsson, N. A., and Simpson, J. R., *Erbium-doped Fiber Amplifiers, Fundamentals and Technology*, Academic Press, 1999.

Boyd, R. W., *Radiometry and the Detection of Optical Radiation*, John Wiley & Sons, New York, 1983.

Bromage, J., Bouteiller, J.-C., Thiele, H. J., Brar, K., Park, J. H., Headley, C., Nelson, L. E., Qian, Y., DeMarco, J., Stulz, S., Leng, L., Zhu, B., and Eggleton, B. J., "S-band

all-Raman amplifiers for 40 × 10 Gb/s transmission over 6 × 100 km of non-zero dispersion fiber," OFC, pd4, 2001.

Carlson, A. B., *Communication Systems: An Introduction to Signals and Noise in Electrical Communication*, McGraw-Hill, 1986.

Chernikov, S. V., Platonov, N. S., Gapontsev, D. V., Chang, D., Guy, M. J., and Taylor, J. R., "Raman fibre laser operating at 1.24 μm," Electron. Lett., v. 34, p. 680, 1998.

Chernikov, S. V., Zhu, Y., Kashyap, R., and Taylor, J. R., "High-gain, monolithic, cascaded fibre Raman amplifier operating at 1.3 μm," Electron. Lett., v. 31, p. 472, 1995.

Chinn, S. R., "Temporal observation and diagnostic use of double Rayleigh scattering in distributed Raman amplifiers," IEEE Photon Technol. Lett., v. 11, p. 1632, 1999.

Chraplyvy, A. R., "Limitations on lightwave communications imposed by optical-fiber nonlinearities," J. of Lightwave Technol., v. 8, p. 1548, 1990.

Chraplyvy, A. R., "Optical power limits in multichannel wavelength-division-multiplexed systems due to stimulated Raman scattering," Electron. Lett., v. 20, p. 58, 1984.

Chraplyvy, A. R., and Henry, S. P., "Performance degradation due to stimulated Raman scattering in wavelength-division-multiplexed optical-fibre systems," Electron. Lett., v. 16, p. 641, 1983.

Cotter, D., and Hill, A. M., "Stimulated Raman crosstalk in optical transmission: Effects of group velocity dispersion," Electron. Lett., v. 20, p. 185, 1984.

Dakss, M. L., and Melman, P., "Amplified spontaneous Raman scattering and gain in fiber Raman amplifiers," J. of Lightwave Technol., v. LT-3, p. 806, 1985.

Davey, S. T., Williams, D. L., Ainslie, B. J., Rothwell, W. J. M., and Wakefield, B., "Optical gain spectrum in GeO2-SiO2 Raman fibre amplifiers," IEE proc., v. 136, p. 301, 1989.

Desurvire, E., "Spectral noise figure of Er3+-doped fiber amplifiers," IEEE Photonics Technol. Lett., v. 2, p. 208, 1990.

Desurvire, E., *Erbium-Doped Fiber Amplifiers: Principles and Applications*, John Wiley & Sons, New York, 1994.

Desurvire, E., DiGonnet, M., and Shaw, H. J., "Theory and Implementation of a Raman Active Fiber Delay Line," J. of Lightwave Technol., v. LT-4, p. 426, 1986.

Desurvire, E., Simpson, J. R., and Becker, P. C., Opt. Lett., v. 12, p. 888 1987.

Dianov, E. M., Abramov, A. A., Bubnov, M. M., Shipulin, A. V., Prokhorov, A. M., Semjonov, S. L., Schebunjaev, A. G., Devjatykh, G. G., Guryanov, A. N., and Khopin, V. F., "Demonstration of 1.3 μm Raman fiber amplifier gain of 25 dB at a pumping power of 300 mW," Opt. Fiber Technol., v. 1, p. 236, 1995.

Dianov, E. M., Grekov, M. V., Bufetov, I. A., Vasiliev, S. A., Medvedkov, O. I., Plotnichenko, V. G., Koltashev, V. V., Belov, A. V., Bubnov, M. M., Semjonov, S. L., and Prokhorov, A. M., "CW high power 1.24 μm and 1.48 μm Raman lasers based on low loss phosphosilicate fibre," Electron. Lett., v. 33, p. 1542, 1997.

Edagawa, N., Mochizuki, K., and Iwamoto, Y., "Simultaneous amplification of wavelength-division-multiplexed signals by a highly efficient fibre Raman amplifier pumped by high-power semiconductor lasers," Electron. Lett., v. 23, p. 196, 1987.

Emori, Y., Akasaka, Y., and Namiki, S., "Broadband lossless DCF using Raman amplification pumped by multichannel WDM laser diodes," Electron. Lett., v. 34, p. 2145, 1998.

Emori, Y., and Namiki, S., "100 nm bandwidth flat gain Raman amplifiers pumped and gain-equalized by 12-wavelength-channel WDM high power laser diodes," OFC, pd19, San Diego, 1999.

Forghieri, F., Tkach, R. W., and Chraplyvy, A. R., "Effect of modulation statistics in Raman crosstalk in WDM systems," IEEE Photon. Technol. Lett., v. 7, p. 101, 1995.

Forghieri, F., Tkach, R. W., and Chraplyvy, A. R., "Bandwidth of cross talk in Raman amplifiers," Opt. Fiber Commun. Conference, pap. FC6, San Jose, CA, Feb. 20–25, 1994.

Galeener, F. L., "Band limits and the vibrational spectra of tetrahedral glasses," Phys. Rev. B., v. 19, p. 4292, 1979.

Galeener, F. L., Leadbetter, A. J., and Stringfellow, M. W., "Comparison of the neutron, Raman, and infrared vibrational spectra of vitreous SiO2, GeO2 and BeF2," Phys. Rev. B., v. 27, p. 1052, 1983.

Galeener, F. L., Mikkelsen, J. C. Jr., Geils, R. H., and Mosby, W. J., "The relative Raman cross sections of vitreous SiO2, GeO2, B2O3, and P2O5," Appl. Phys. Lett., v. 32, p. 34, 1978.

Grubb, S. G., Erdogan, T., Mizrahi, V., Strasser, T., Cheung, W. Y., and Reed, W. A., "1.3 μm cascaded Raman amplifier in germanosilicate fiber," Proc. OSA Topical Meeting Optical Amplifiers and Their Applications, Breckenridge, CO, pap. PD3, 1994.

Grubb, S. G., Strasser, T. A., Cheung, W. Y., Reed, W. A., Mizrahi, V., Erdogan, T., Lemaire, P. J., Vengsarkar, A. M., and DiGiovanni, D. J., "High-power 1.48 μm cascaded Raman laser in germanosilicate fibers," Opt. Amp. and Appl., p. 197, Davos, Switzerland, 1995.

Hansen, P. B. et al., "529-km unrepeatered transmission at 2.488 Gb/s using dispersion compensation, forward error correction, and remotely pumped post- and pre-amplifiers," Electron. Lett., v. 31, p. 1460, 1995.

Hansen, P. B., Eskildsen, L., Grubb, S. G., Stentz, A. J., Strasser, T. A., Judkins, J., DeMarco, J. J., Pedrazzani, R., and DiGiovanni, D. J., "Capacity upgrades of transmission systems by Raman amplification," IEEE Photon Technol. Lett., v. 9, p. 262, 1997.

Hansen, P. B., Eskildsen, L., Stentz, A. J., Strasser, T. A., Judkins, J., DeMarco, J. J., Pedrazzani, R., and DiGiovanni, D. J., "Rayleigh scattering limitations in distributed Raman pre-amplifiers," IEEE Photon. Technol. Lett., v. 10, p. 159, 1998.

Hansen, P. B., Jacobovitz-Veselka, G., Gruner-Nielsen, L., and Stentz, A. J., "Raman amplification for loss compensation in dispersion compensating fibre modules," Electron. Lett., v. 34, p. 1136, 1998.

Hansen, P. B., Stentz, A., Nielsen, T. N., Espindola, R., Nelson, L. E., and Abramov, A., "Dense wavelength-division multiplexed transmission in 'zero-dispersion' DSF by means of hybrid Raman/erbium-doped fiber amplifiers," OFC, pd8, San Diego, 1999.

Hartog, A. H., and Gold, M. P., "On the theory of backscattering in single-mode optical fibers," J. of Lightwave Technol., v. LT-2, p. 76, 1984.

Haus, H. A., "The Noise Figure of Optical Amplifiers," IEEE Photon Technol. Lett., v. 10, no. 11, pp. 1602–1604, 1998.

Hegarty, J., Olsson, N. A., and Goldner, L., "CW pumped Raman preamplifier in a 45 km-long fibre transmission system operating at 1.5 μm and 1 Gbit/s," Electron. Lett., v. 21, p. 290, 1985.

Jiang, W., and Ye, P., "Crosstalk in fiber Raman amplification for WDM systems," J. of Lightwave Technol., v. 7, p. 1407, 1989.

Kani, J., Jinno, M., and Oguchi, K., "Fibre Raman amplifier for 1520 nm band WDM transmission," Electron. Lett., v. 34, p. 1745, 1998.

Lewis, S. A. E., Chernikov, S. V., and Taylor, J. R., "Temperature-dependent gain and noise in fiber Raman amplifiers," Opt. Lett., v. 24, p. 1823, 1999.

Lewis, S. A. E., Chernikov, S. V., and Taylor, J. R., "Characterization of double Rayleigh scatter noise in Raman amplifiers." IEEE Photon Technol. Lett., v. 12, p. 528, 2000.

Lin, C., and Stolen, R. H., "Backward Raman amplification and pulse steepening in silica fibers," Appl. Phys. Lett., v. 29, p. 428, 1976.

Ma, M. X., Kidorf, H. D., Rottwitt, K., Kerfoot, F., III, and Davidson, C., "240-km repeater spacing in a 5280-km WDM system experiment using 8 × 2.5 Gb/s NRZ transmission," IEEE Photon. Technol. Lett., v. 10, p. 893, 1998.

Mahlein, H. F., "Crosstalk due to stimulated Raman scattering in single mode fibres for optical communication in wavelength division multiplex systems," Optical and Quantum Electronics, v. 16, p. 409, 1984.

Marcuse, D., "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," J. of Lightwave Technol., v. 8, p. 1816, 1990.

Mears, R. J., Reekie, L., Jauncie, I. M., and Payne, D. N., Elect. Lett., v. 23, p. 1026, 1987.

Mermelstein, M. D., Headley, C., Bouteiller, J.-C., Steinvurzel, P., Horn, C., Feder, K., and Eggleton, B. J., "A High-Efficiency Power-Stable Three-Wavelength Configurable Raman Fiber Laser," OFC, pd3, 2001.

Mills, D. L., *Nonlinear Optics: Basic Concepts*, Springer Verlag, New York, 1998.

Mollenauer, L. F., Stolen, R. H., and Islam, M. N., "Experimental demonstration of soliton propagation in long fibers: loss compensated by Raman gain," Opt. Lett., v. 10, p. 229, 1985.

Movassaghi, M., Jackson, M. K., Smith, V. M., and Hallam, W. J., "Noise Figure of Erbium-Doped Fiber Amplifiers in Saturated Operation," J. of Lightwave Technol., v. 16, no. 5, May, 1998.

Musada, H., Kawai, S., and Aida, K., "108-nm gain-band optical amplification using distributed Raman amplification and erbium-doped fiber amplifier," OECC, pap. PD2-1, Makuhari Messe, July, 1998.

Musada, H., Kawai, S., and Aida, K., "76-nm 3-dB gain-band hybrid fiber amplifier without gain-equalizer," ECOC, pap. PD7-2, Madrid, Spain, 1998.

Nakazawa, M., Nakashima, T., and Seikai, S., "Raman amplification in 1.4–1.5 μm spectral region in polarization-preserving optical fibers," J. Opt. Soc. Am. B., v. 2, p. 515, 1985.

Nielsen, T. N., Hansen, P. B., Stentz, A. J., Aquaro, V. M., Pedrazzani, J. R., Abramov, A. A., and Espindola, R. P., "8 × 10 Gb/s 1.3 μm unrepeatered transmission over a distance of 141 km with Raman post- and pre-amplifiers," IEEE Photon. Technol. Lett., v. 10, p. 1492, 1998.

Nielsen, T. N., Stentz, A. J., Hansen, P. B., Chen, Z. J., Vengsarkar, D. S., Strasser, T. A., Rottwitt, K., Park, J. H., Stulz, S., Cabot, S., Feder, K. S., Westbrook, P. S., and Kosinski, S. G., "1.6 Tb/s (40 × 40 Gb/s) transmission over 4 × 100 km nonzero-dispersion fiber using hybrid Raman/erbium-doped inline amplifiers," ECOC, pd2-2, 1999.

Nielsen, T. N., Stentz, A. J., Rottwitt, K., Vengsarkar, D. S., Chen, Z. J., Hansen, P. B., Park, J. H., Feder, K. S., Cabot, S., Peckham, D. W., Hsu, L., Kan, C. K., Judy, A. F., Park, S. Y., Nelson, L. E., and Gruner-Nielsen, L., "3.28-Tb/s transmission over 3 × 100 km of nonzero-dispersion fiber using dual C- and L-band distributed Raman amplification," IEEE Photon Technol. Lett., v. 12, p. 1079, 2000.

Nissov, M., Davidson, C. R., Rottwitt, K., Menges, R., Corbett, P. C., Innis, D., and Bergano, N. S., "100 Gb/s (10 × 10 Gb/s) WDM transmission over 7200 km using Distributed Raman Amplification," ECOC, pdp., 1997.

Olsson, N. A., "Lightwave systems with optical amplifiers," J. of Lightwave Technol., v. 7, p. 1071, 1989.

Olsson, N. A., and Hegarty, J., "Noise Properties of a Raman Amplifier," J. of Lightwave Technol., v. LT-4, 1986.

Po, H., Cao, J. D., Laliberte, B. M., Minns, R. A., Robinson, R. F., Rockney, B. H., Tricca, R. R., and Zhang, Y. H., "High power neodymium-doped single transverse mode fibre laser," Electron. Lett., v. 29, p. 1500, 1993.

Puc, A. B., Chbat, M. W., Henrie, J. D., Weaver, N. A., Kim, H., Kaminski, A., Rahman, A., and Fevrier, H., "Long-haul WDM NRZ transmission at 10.7 Gb/s in S-band using cascade of lumped Raman amplifiers," OFC, pd39, 2001.

Rasmussen, C. J., Fenghai, L., Pedersen, R. J. S., and Jorgensen, B. F., "Theoretical and experimental studies of the influence of the number of crosstalk signals on the penalty caused by incoherent optical crosstalk," OFC, pap. TuR5-1, 1999.

Rottwitt, K., and Kidorf, H. D., "A 92 nm Bandwidth Raman Amplifier," OFC '98, pap. pd 6, 1998.

Rottwitt, K., Bromage, J., Du, M., and Stentz, A., "Design of distributed Raman amplifiers," ECOC, pap. 4.4.1, 2000.

Rottwitt, K., Povlsen, J. H., Bjarklev, A., Lumholt, O., Pedersen, B., and Rasmussen, T., "Noise in Distributed Erbium-Doped Fibers," IEEE Photonics Technol. Lett., v. 5, no. 2, 1993.

Shibata, N., Horigudhi, M., and Edahiro, T., "Raman spectra of binary high-silica glasses and fibers containing GeO_2, P_2O_5 and B_2O_3," J. of Non-Crystalline Solids, v. 45, pp. 115–126, 1981.

Shimoda, K., Takahasi, H., and Townes, C. H., "Fluctuations in amplification of quanta with application to laser amplifiers," J. of the Physical Soc. of Japan, v. 12, p. 686, 1957.

Smith, R. G., "Optical power handling capacity of low loss optical fibers as determined by stimulated Raman and Brillouin scattering," Appl. Optics, v. 11, p. 2489, 1972.

Smith, R. G., and Personick, S. D., "Receiver design for optical fiber communication systems," ch. 4 in *Semiconductor Devices for Optical Communication*, edited by H. Kressel, second updated edition, Springer Verlag, 1987.

Srivastava, A. K., Tzeng, D. L., Stentz, A. J., Johnson, J. E., Pearsall, M. L., Mizuhara, O., Strasser, T. A., Dreyer, K. F., Sulhoff, J. W., Zhang, L., Yeates, P. D., Pedrazzani, J. R., Sergent, A. M., Tench, R. E., Freund, J. M., Nguyen, T. V., Manar, H., Sun, Y., Wolf, C., Choy, M. M., Kummer, R. B., Kalish, D., and Chraplyvy, A. R., "High speed WDM transmission in AllWave fiber in both the 1.4 μm and 1.55 μm bands," OAA '98, PD2, Vail, Colorado, July, 1998.

Stentz, A. J., Grubb, S. G., Headley, C. E. III, Simpson, J. R., Strasser, T., and Park, N., "Raman Amplifier with Improved System Performance," Optical Fiber Communication, TuD3, San Jose, CA, 1996.

Stentz, A. J., Nielsen, T., Grubb, S. G., Strasser, T. A., and Pedrazzani, J. R., "Raman Ring Amplifier at 1.3 μm with Analog-Grade Noise Performance and an Output Power of 23 dBm," Optical Fiber Communication, PD16, San Jose, CA, 1996.

Stolen, R. H., "Polarization effects in fiber Raman and Brillouin lasers," IEEE J. of Quantum Electron., v. QE-15, p. 1157, 1979.

Stolen, R. H., and Ippen, E. P., "Raman gain in glass optical waveguides," Appl. Phys. Lett., v. 22, p. 276, 1973.

Stolen, R. H., Ippen, E. P., and Tynes, A. R., "Raman oscillation in glass optical waveguide," Appl. Phys. Lett., v. 20, p. 62, 1972.

Stolen, R. H., Lee, C., and Jain, R. K., "Development of the stimulated Raman spectrum in single-mode silica fibers," JOSA B, v. 1, p. 52, 1984.

Takachio, N., Suzuki, H., Masuda, H., and Koga, M., "32 × 10 Gb/s distributed Raman amplification transmission with 50 GHz channel spacing in zero-dispersion region over 640 km of 1.55 μm dispersion-shifted fiber," OFC, pd9, San Diego, 1999.

Takahashi, H., Oda, K., and Toba, H., "Impact of crosstalk in an arrayed-waveguide multiplexer on NxN optical interconnection," J. of Lightwave Technol., v. 14, p. 1092, 1996.

Tariq, S., and Palais, J. C., "A computer model of non-dispersion-limited stimulated Raman scattering in optical fiber multiple-channel communications," J. of Lightwave Technol., v. 11, p. 1914, 1993.

Tobin, C. T., and Baak, T., "Raman spectra in some low-expansion glasses," J. of the Optical Soc. of America, v. 58, p. 1459, 1968.

Urquhart, W. P., and Laybourn, P. J., "Effective core area of stimulated Raman scattering in single-mode optical fibres," IEE Proc., v. 132, p. 201, 1985.

Wan, P., and Conradi, J., "Double Rayleigh backscattering in long-haul transmission systems employing distributed and lumped fibre amplifiers," Electron. Lett., v. 31, p. 383, 1995.

Willard, H. H., Merritt, L. L., Dean, J. A., and Settle, F. A., *Instrumental Methods of Analysis*, Wadsworth Publishing Company, California, 1988.

Yariv, A., *Quantum Electronics*, John Wiley & Sons, New York, 1989.

Chapter 6 | Electrooptic Modulators

Amaresh Mahapatra

Linden Photonics, Inc., Acton, Massachusetts

Edmond J. Murphy

JDS Uniphase, Windsor, Connecticut

1. Introduction

The advent of wavelength-division multiplexing in deployed systems over the last five years has resulted in a new emphasis on optical modulation techniques. On the one hand, the use of several wavelengths over a single fiber tends to mitigate the need for higher and higher modulation speeds. On the other hand, the explosive growth in demand for bandwidth continuously pushes on all methods for increasing the data-carrying capacity of a fiber, including higher modulation rates. By some measures, optical modulation speeds have doubled every 10 months over the last few years, thus exceeding the well known Moore's law for microprocessors, where chip complexity doubles every 18 months.

As of this writing, the deployment of optics in the long haul is accepted practice, while deployment in the metro and access markets has only just commenced. In either case, deployment has happened in a transient state; that is, the optimum technologies for a specific information link are not necessarily known or understood, but the urgency of the market requires the installation of the best available option as soon as possible. Hence the abundance of new, venture-funded system vendors. As a result, research and development into all aspects of component technology continues at a blistering pace, since there is as yet no industry standard. Hence the abundance of venture-funded component manufacturers.

Specifically in the electrooptic modulator arena there is work in three different integrated optic modulator technologies: lithium niobate modulators, polymeric electrooptic modulators, and semiconductor electrooptic modulators.

In this chapter we survey the design, fabrication, and performance of lithium niobate and polymeric modulators. Semiconductor electrooptic modulators of the interferometric kind have so far not found commercial acceptance. The single major technical issue is the high fiber pigtailed insertion loss of these modulators, typically 8 to 10 dB. However, recently there has been renewed interest in these because of the possibility of integration of

OPTICAL FIBER TELECOMMUNICATIONS,
VOLUME IVA

ISBN: 0-12-395172-0

lasers and interferometric modulators and semiconductor optical amplifiers (SOAs) on a single substrate where the SOA can compensate for the loss of the modulator. For information on semiconductor electrooptic modulators the references will serve as a good starting point (Walker, 1991; Walker, 1995; Griffin, et al., 2001).

Section 2 presents a comparison between electrooptic and electroabsorption modulator technologies. Section 3 details fabrication technology for lithium niobate modulators from lithography to fiber attachment and packaging. Section 4 details design issues of lithium niobate modulators, such as velocity matching and bandwidth optimization. Section 5 discusses the performance of lithium niobate modulators in optical systems. Section 6 covers the design and fabrication aspects of polymer electrooptic modulators.

2. Comparison of Electrooptic and Electroabsorption Modulators

All three types of electrooptic modulators compete with electroabsorption modulators (EAM), which have now been demonstrated in the laboratory at speeds as high as 40 GHz (Ouagazzaden, 2001).

An EAM works on the principle of the Franz-Keldysh effect, which is the observed lengthening in wavelength of the optical absorption edge of a semiconductor with the application of an electric field (Noda, 1986; Reinhart, 1976; Stillman, 1976; Dutta, 1984; Noda, 1985). In multiquantum well structures this is often referred to as the quantum confined Stark effect. An electroabsorption modulated laser (EML) combines in one chip a continuously operating distributed-feedback (DFB) laser and an electroabsorption modulator that acts like a very fast shutter, blocking the DFB laser's output or letting it pass. A lithium niobate modulator is external to the laser. Such modulators are referred to as external modulators. They are bulkier and require the light from the laser to be coupled via fiber into the input of the external modulator.

An EML is an attractive device for systems because of small size, low drive voltage and low chirp compared to direct modulation. An EML is specially attractive, because the high coupling efficiency between the modulator and the laser in the integrated structure leads to high output power and small size. Researchers have also succeeded in tuning the laser element by standard tunable laser diode technology through several hundred GHz, thus reducing the need for maintaining an inventory of EMLs for every wavelength of the ITU grid. However, the chirp of EMLs, while small, is not zero. Time-resolved chirp has been measured for EMLs grown by selective area metalloorganic vapor phase epitaxy (Runge, 1988; Johnson, et al., 1994) and shown to be an aggregate of about 0.11 Å over the rising and falling edges of a pulse. The chirp results from a small change in the reflectivity of the modulator segment

as electrons are injected into it to produce absorption. This chirp limits the transmission distance of EML-based 10 Gb/s links to a maximum of 100 km.

Externally modulated 10 Gb/s links using lithium niobate have been demonstrated to distances of several hundred km. The extinction of EMLs is about 10 dB while that of lithium niobate modulators can be as high as 25 dB. The switching voltage, on the other hand, for EMLs is as low at 2.5 V, while lithium niobate requires as much as 7–8 V at the high end of the modulation band.

3. Fabrication of Lithium Niobate Optical Modulators

3.1 ELECTROOPTIC EFFECT

The linear electrooptic effect, also known as the Pockels effect, is the change in refractive index of a material that is caused by and is proportional to an applied electric field. This effect exists only in crystals that do not possess inversion symmetry. The constant of proportionality between the refractive index and the electric field is called the electrooptic constant, usually designated in the literature by r. The optical propagation characteristics in crystals are fully described by means of the index ellipsoid. The nine elements of the refractive index tensor are reduced to six using the well-known reduced notation and are designated by n_i ($i = 1$ through 6). Since the electric field is a vector with three components, in the most general case r has eighteen independent components that are designated by two subscripts, r_{ij}. These components are further reduced in number by the point symmetry of the specific crystal.

3.2 OPTICAL AND ELECTROOPTICAL PROPERTIES OF LITHIUM NIOBATE

Lithium niobate has been the material of choice for the fabrication of electrooptic modulators due to its combination of high electrooptic coefficients and high optical transparency in the near infrared wavelengths. It exhibits several interesting material properties, such as the piezoelectric, pyroelectric, and acoustooptic effects, all of which have been exploited for diverse device concepts. The property that is of special interest in optical modulators is the electrooptic effect.

Lithium niobate has 3 m point symmetry. Applying the 3 m symmetry to the electrooptic tensor reduces the number of independent electrooptic constants to four, $r_{22}, r_{33}, r_{13}, r_{14}$. The constant r_{14} is responsible for polarization rotation in lithium niobate polarization controllers but is not relevant to modulators. The values of the other three, in units of 10^{-12} m/V, are: $r_{22} = 3.4$, $r_{13} = 8.6$, $r_{33} = 30.8$. Clearly r_{33}, which measures the change in the extraordinary index n_e as a result of a z-directed electric field, is significantly larger than the others. Therefore most lithium niobate modulators for communications are designed to exploit this electrooptic constant. From a design point of view, this implies

that electrodes must be positioned such that the electric field generated within the optical guide is along the z crystal axis. This fact will be important in understanding several of the designs discussed in the rest of this chapter.

The lithium niobate crystals arrive from the manufacturer with an optical grade polish on the face to be processed and an inspection grade polish on the backside. Orientation is specified by suitable fiduciary marks on the substrate. Currently the largest lithium niobate crystals available commercially are 6″. It should be noted that, in cleaning and processing, lithium niobate is intolerant of rapid heating and cooling and may crack if subjected to these conditions. In addition, z-cut lithium niobate exhibits a strong pyroelectric effect. A pyroelectric solid exhibits a change in spontaneous electric polarization vector, ΔP, as a function of temperature. It is a linear effect and can be written as $\Delta P = p\Delta T$ where p is the pyroelectric tensor. In lithium niobate this effect is due to the movement of the lithium and niobium ions relative to the oxygen layers. Since these ions move only in a direction parallel to the c-axis, the only change in volume polarization is along the z-direction and is characterized by the pyroelectric constant $p_3(-4 \times 10^{-5}$ Coulomb/°C/m^2) (Savage, 1966). This implies that as the crystal is heated, negative charge collects on the $+z$ face while positive charge collects on the $-z$ face. The pyroelectric effect must be considered when designing devices and processing steps.

3.3 WAVEGUIDE FABRICATION

Two methods of fabricating waveguides in lithium niobate are popular: titanium indiffusion and annealed proton exchange.

Considerable data has been reported in the literature pertaining to titanium diffusion in lithium niobate (Schmidt and Kaminow, 1974; Minaka, et al., 1979; Pearsall, et al., 1976; Sugii, et al., 1978; Burns, et al., 1979; Stultz, 1979). Titanium indiffusion is a high-temperature process that increases both the ordinary (Δn_o) and extraordinary (Δn_e) indices and therefore provides channel guides that support both TE and TM modes. There is a dependence of diffusion constant on stoichiometry (Holmes and Smyth, 1984). The dispersion of Δn (Fouchet, 1987) impacts the design of channel guides for different wavelengths or the use of the same device over a range of wavelengths. The standard procedure for making titanium indiffused waveguides in lithium niobate is as follows. The waveguide pattern is delineated on the lithium niobate surface using standard lithography. Titanium is then deposited on the substrate using thermal evaporation, e-beam deposition, or sputter deposition followed by liftoff to define a titanium strip. The deposition technique impacts the film density and hence the thickness used for fabricating the guides. The smoothness of the titanium lines determines the propagation loss of the channel guides. With good lithography practice, propagation loss can be reduced to 0.1 dB/cm at 1300 and 1550 nm. Thermal evaporation and e-beam are the

preferred methods of deposition. Typical titanium layer thickness ranges from of 500 Å to 1200 Å and strip width ranges from 4 μm to 10 μm.

The next step is to diffuse the titanium into the lithium niobate; the diffusion temperature typically ranges from 950°C to 1100°C, while the diffusion time ranges from 4 to 12 hours. One significant problem is the outdiffusion of lithium from lithium niobate at temperatures higher than 700°C. This is not desirable, since lithium depletion increases n_e and results in a weak slab guide on which the titanium channel guide sits. Lithium outdiffusion is avoided by use of flowing wet oxygen in the diffusion tube during the titanium diffusion process. Using appropriate values in the ranges indicated above it is possible to fabricate single-mode channel guides for all wavelengths from 700 nm to 1600 nm. Almost all lithium niobate modulators fabricated commercially are designed for operation at a nominal wavelength of 1550 nm.

Annealed proton exchange is a low-temperature method of fabricating waveguides in lithium niobate (Jackel, et al., 1982; Suchoski, et al., 1988; Yi-Yan, 1983; Howerton, et al., 1991). The fabrication requires that lithium atoms in the substrate be replaced by hydrogen atoms. To effect the exchange, a metal mask defining a suitable channel guide is described on the crystal surface using aluminum, chrome, or titanium metallization and standard photolithography. The substrate is then dipped in benzoic acid at a temperature between 150°C and 250°C for times ranging from minutes to hours (several acids other than benzoic have been used in the literature). This produces an exchange between hydrogen from the acid melt and lithium from the substrate resulting in an increase in refractive index of about 0.1 in the exchanged region. It is difficult to make a single-mode guide with such a large index change. Furthermore, proton-exchanged guides have very high propagation losses of about 6 dB/cm. Therefore, the guides are subsequently annealed at temperatures ranging from 250°C to 450°C to reduce the hydrogen concentration and reduce the index difference to about 0.01. Appropriately annealed proton-exchanged guides have propagation losses of about 0.1 dB/cm, similar to titanium diffused guides (Suchoski, et al., 1988).

3.4 ELECTRODE FABRICATION

The above fabrication conditions may be used both for x-cut, y-propagating and z-cut, y-propagating waveguides, which are the two common orientations used in commercial devices. The next step is the fabrication of electrodes. The electrode alignment with respect to the two waveguide configurations is shown in Fig. 3.1. For both orientations, the goal is to exploit the largest electrooptic constant of lithium niobate, r_{33}. This requires that for x-cut, y-propagating guides, the electrodes be deposited on either side of the guide so that the predominant electric field is along the crystal z-axis. However, to achieve the same result in z-cut, y-propagating devices, one electrode needs to be deposited on top of the guide. In this geometry, it is essential to deposit a thin dielectric

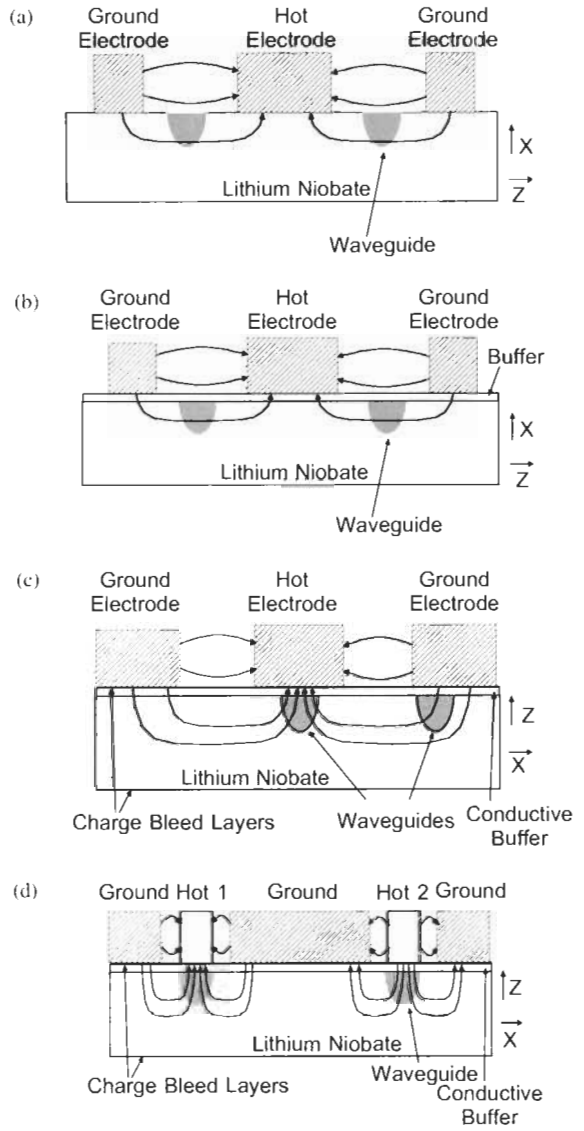


Fig. 3.1 Most common electrode configurations for (a) nonbuffered x-cut, (b) buffered x-cut, (c) buffered single-drive z-cut, and (d) buffered dual drive z-cut.

buffer layer between the guide and the electrodes to avoid excessive attenuation of the optical field. Note that metal electrodes laid directly on an optical guide without a buffer can result in 1 dB/mm propagation loss, while a good buffer layer can lead to guides with 0.1 dB/cm propagation loss. However, buffer layers are used in x-cut, y-propagating devices on occasion, not to eliminate

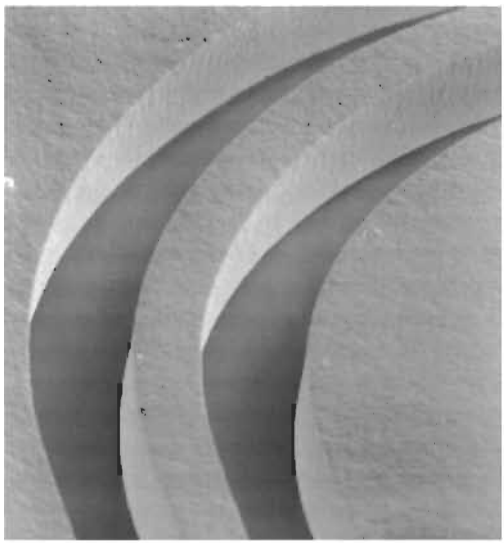


Fig. 3.2 SEM picture of 18 μm-thick gold-plated CPW electrodes at a region where the electrodes are making a bend.

metal loading effects, but to achieve velocity matching between RF and optical waves. Hence the four configurations shown in Fig. 3.1.

Electrodes are typically fabricated with chrome/gold or titanium/gold where the chrome and titanium enhance adhesion. Standard lithography is used to define the gold, while the metallization is done by vacuum deposition or electroplating. Traveling-wave modulators typically use electrodes that are in excess of 15 μm thick, so that electroplating is used exclusively for electrode fabrication. Plating processes yielding high-purity metal, small grain size, and smooth electrode side walls and surfaces along with reproducible dimensions are key to predictable high-frequency performance. Electroplating quality is optimized by controlling composition, pH, temperature of the bath, and also the rate of deposition. Figure 3.2 shows a scanning electron microscope (SEM) picture of a typical gold-plated coplanar waveguide electrode fabricated for a 10 Gb/s modulator.

3.5 DICING AND POLISHING

In contrast to semiconductor materials such as GaAs or InP, lithium niobate substrates do not have convenient cleavage planes. The only cleavage plane contains the x-axis, while the z optic axis is at an angle of 32.75° from the cleavage plane (Kaminow, 1978). Therefore, it becomes essential to polish the requisite plane of the substrate mechanically so that light can be coupled into the waveguide by end-fire coupling. This is done as follows. Substrates containing an array of finished modulators are cut from the lithium niobate

wafer using conventional water-cooled diamond saws. The substrate end faces are cut at an angle to the waveguides in order to eliminate optical reflections and are then polished to an optical finish. Most polishing processes require the substrate to be sandwiched between two other backing plates with a very thin glue line between the three parts. Since the lithium niobate is in the middle of the sandwich, the critical edge where the guide cross-sections reside are not rounded in the polishing process. A good optical finish and a sharp edge are required at both the input and output optical facets of the device to ensure good fiber-to-waveguide coupling.

3.6 PIGTAILING, PACKAGING, AND TESTING

In a real world application, the integrated optical chip must be pigtailed and packaged in order that optical and electrical signals can be efficiently and effectively coupled into and out of the device. Three principal subassemblies are used in the manufacture and packaging of lithium niobate substrates. They are the lithium niobate chip, fiber carrier assemblies for input and output coupling, and a housing that provides for electrical and RF connection to the chip.

Lithium niobate modulators can be packaged either in hermetic or nonhermetic housings, depending on the application and operating environment (Moyer, et al., 1998). For devices designed to operate within telecommunication central offices, nonhermetic packaging has proven both sufficient and cost effective for meeting reliability and qualification requirements (Maack, 1999).

Due to the polarization dependence of the electrooptic effect, the polarization state of the input light supplied to the modulator must be carefully maintained to achieve optimum performance. Hence, most lithium niobate modulators use polarization-maintaining fiber to couple light into the device while the output fiber is typically a standard single-mode fiber. During the preparation of the input and output fiber subassemblies, a small capillary or etched silicon block is attached to the end of each fiber, which serves to provide a large rigid body to hold during the fiber alignment and gluing process. In the case of the input polarization-maintaining fiber, this block also provides a method of transferring the optical axis of the fiber to reference planes on the block, which can be easily viewed in an imaging system or identified by machine vision systems.

Electrical interconnections are attached and soldered to the modulator housing, thereby creating the third subassembly. The pigtailed lithium niobate chip is attached to the package using a compliant adhesive that mechanically decouples the optical assembly from the package and absorbs thermally induced strains (O'Donnell, 1995). Last, electrical connection between the package and the lithium niobate chip is accomplished using either ribbon or

wire bonding. Length and quality of bond wire are critical determinants of high-frequency performance of the modulator.

Once all three subassemblies are completed, the modulator must be sealed and tested as a complete functional unit. Key parameters that are typically measured during final test include optical loss, switching voltage, optical on/off ratio (also called extinction ratio) bias stability, and microwave properties such as S11 and S21. Often, some of these parameters are measured over the operating temperature range of the device.

3.7 REDUCTION OF DRIVE VOLTAGE V_π

Several numerical calculations have been made of electrical fields in the gap and in the vicinity of traveling wave electrodes used in lithium niobate optical modulators. Let us consider the two geometries popularly used in commercial devices shown in Figure 3.1. In all cases the electrodes are positioned such that the guides see an electric field directed along the z-axis of the crystal, since this allows use of the largest EO coefficient, r_{33}.

For x-cut lithium niobate, the electrodes are positioned on either side of the guides so that the tangential electric field parallel to the substrate surface is along the crystal z-axis. The guides are generally placed in the center of the gap (Figure 3.1a, b). However, it has been shown (Becker and Kincaid, 1993) that the best overlap of electric and optical field occurs with the guide located off center and closer to the hot electrode of the coplanar three-electrode structure. This is because of the edge effect, which enhances the tangential electric field in the vicinity of the electrode edge because of a concentration of field lines in this region. This enhancement occurs both at the edge of the ground electrode and the hot electrode but is more pronounced at the hot electrode, as seen in Fig. 3.3. It has been shown (Becker and Kincaid, 1993) that by locating the guides exactly adjacent to the hot electrode, V_π can be reduced by 25% compared to the guides being located at the center of the electrodes.

For z-cut lithium niobate, push–pull modulation is achieved by locating the center electrode over one guide and one ground electrode over the other guide as shown in Fig. 3.1c, d. In this geometry, it is the field perpendicular to the substrate surface that is along the crystal z-axis and is responsible for the modulation. Numerical calculations have shown that there is enhancement of the perpendicular electric field in this geometry due to the edge effect (Marcuse, 1981; Ramer, 1982). Indeed, the enhancement is larger than the edge effect enhancement of the tangential field in the x-cut geometry by a factor of 2. However, the guide under the ground electrode does not see as large an electric field as the guide under the center electrode. Even so the net result is that this geometry typically yields a V_π that is 20% to 40% smaller than for x-cut lithium niobate with electrodes of equal length and impedance.

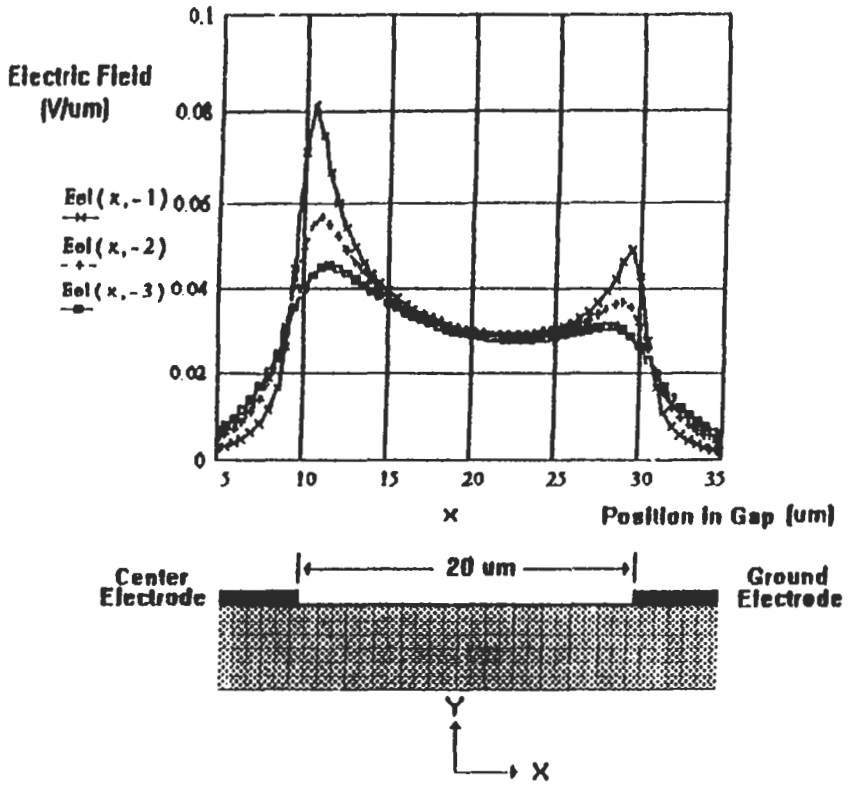


Fig. 3.3 Edge enhancement of tangential electric field in the vicinity of the electrode edge. (Reprinted with permission from Becker and Kincaid, 1993, Improved electrooptic efficiency in guided wave modulators. J. Lightwave Technol. V 11, No. 12, p. 2076, copyright © 1993 IEEE.)

3.8 ETCHING OF LITHIUM NIOBATE TO ENHANCE V_π

Experimental results (Noguchi, et al., 1998) show that in a conventional *z*-cut modulator, velocity match is achieved with a silicon dioxide buffer layer of 0.9 μm and a coplanar waveguide electrode with a gap, width, and thickness of 15 μm, 8 μm and 20 μm, respectively. The product of drive voltage V_π and electrode length L is 13 V · cm (Dolfi and Ranganath, 1992; Onaka, et al., 1996). However, the characteristic impedance becomes 35 Ω, which degrades the modulation response rapidly in the low frequency range. To keep the characteristic impedance near 50 Ω, buffer layer thickness must be larger than 2 μm though $V_\pi L$ increases to about 20 V · cm.

Several attempts have been made to enhance V_π by etching the lithium niobate on either side of the optical guide. This essentially results in a partial ridge waveguide as shown in Fig. 3.4 (Noguchi, et al., 1998). A titanium

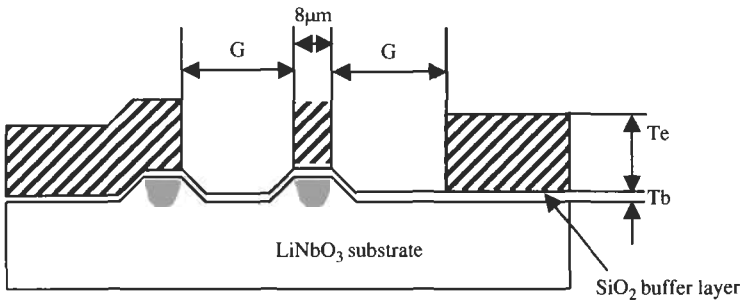


Fig. 3.4 View of the cross-section of a ridged Mach-Zehnder optical modulator.

diffused waveguide is formed in a z-cut lithium niobate substrate. In the interaction region, both sides of the waveguides are etched to form ridges. The substrate is coated with a buffer layer of silicon dioxide. A coplanar waveguide gold electrode is formed on the buffer layer. In the ridge structure, high-dielectric lithium niobate is removed and replaced with low-dielectric materials such as air and silicon dioxide. Thus, the effective dielectric constant of the substrate becomes lower. In a typical design, the substrate was etched to form a ridge with width and depth of 9 μm and 3.6 μm respectively. A buffer layer of 0.9 μm thickness was deposited prior to formation of coplanar waveguide electrodes with a gap, width, and thickness of 25 μm, 8 μm and 29 μm, respectively. The impedance of this modulator was measured at 47 Ω, and $V_\pi L$ was 10.5 V · cm. The modulation bandwidth S_{21} and reflected rf power S_{11} of this device are shown in Fig. 3.5. This corresponds to a 30% improvement in $V_\pi L$. Similarly, others have reported a 1 V improvement in the V_π of a 40 GHz modulator through use of a ridge waveguide where the ridge width and depth were 11 μm and 4.1 μm respectively (Burns, et al., 1999).

Several techniques have been developed to etch z-cut lithium niobate. Argon ion beam milling has been used (Burns, et al., 1999). However, ion milling suffers from the problem of redeposition, where the material physically removed by the high energy bombarding ions often redeposits on some other part of the substrate. Reactive ion beam etching of lithium niobate and titanium-indiffused lithium niobate using C_2F_6 has been demonstrated with AZ1350B photoresist as a masking material (Zhang, et al., 1984). At a beam current density of 0.4 mA/cm^2 the etch rates were as follows:

Lithium niobate	300 Å/sec
Ti-indiffused lithium niobate	400 Å/sec
Niobium pentoxide	700 Å/sec
AZ1350B photoresist	98 Å/sec

Niobium pentoxide has the highest etch rate, probably because niobium easily forms the volatile compound niobium pentafluoride when etched in

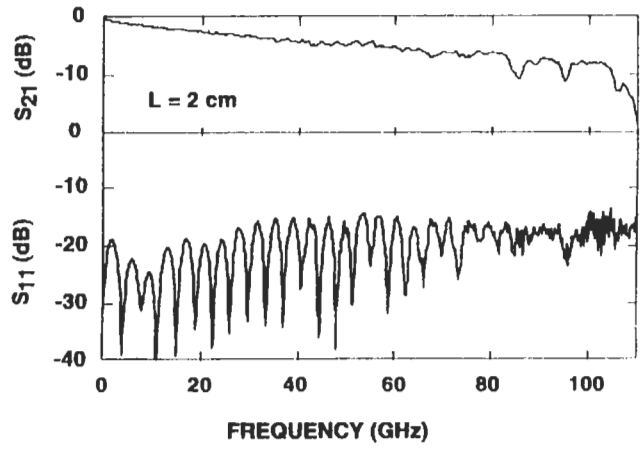


Fig. 3.5 Modulation bandwith, S_{21}, and reflected RF power, S_{11}, for a z-cut lithium niobate modulator with buffer layer and CPW. (Reprinted with permision from Noguchi, K., et al., 1998. Millimeter-wave Ti:$LiNbO_3$ optical modulators. J. Lightwave Technol., V 11, No. 4, p. 615, copyright © 1998 IEEE.)

a fluorinated hydrocarbon. However, lithium fluoride at NTP has a boiling point of 1676°C and is not easily etched. It has been difficult to find an etching gas that reacts with both niobium and lithium to give volatile species.

Some experiments have been performed to use proton exchange to etch lithium niobate. Proton exchange in lithium niobate using different acids results in an exchanged region consisting of $Li_{1-x}H_xNbO_3$. With strong acids x can be close to one (Rice and Jackel, 1984). The exchanged region can then be etched in nitric acid and in aqua regia. However, because of diffusion constant limitations, it is difficult to exchange to a depth of more than 2 μm. Hence, this method has not been used to make etched-ridge traveling-wave modulators.

4. Design of Lithium Niobate Amplitude Modulators

4.1 DIRECTIONAL COUPLERS AND MACH-ZEHNDER AMPLITUDE MODULATORS

Two basic designs are used for fabrication of amplitude modulators in lithium niobate: directional couplers and Mach-Zehnder interferometers (MZI). In the directional coupler approach, shown schematically in Fig. 4.1, light is split into two channel guides that are sufficiently close so that there is coupling between them through evanescent fields. Effectively, the two guides act as a single guiding structure with even and odd guided modes. The applied electric field changes the relative propagation velocities of the odd and even modes, so when the guides are separated at the output end of the coupler, the sum of

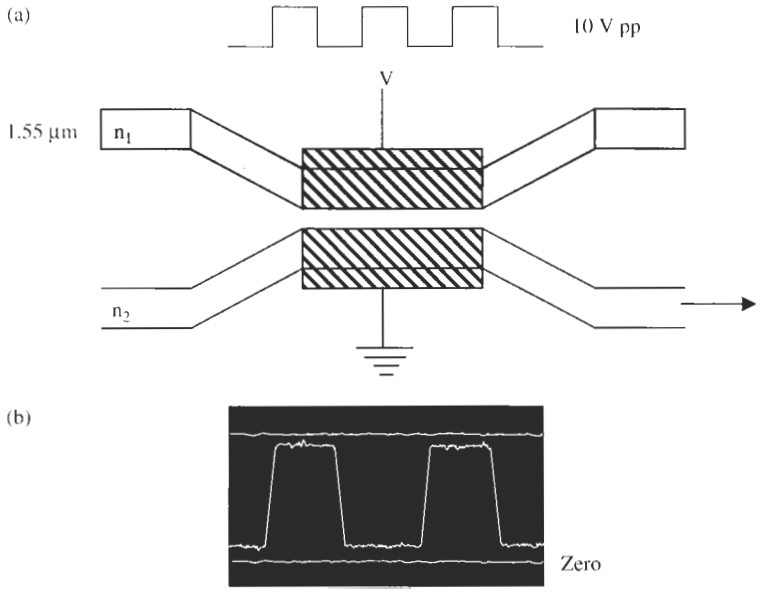


Fig. 4.1 Waveguide directional coupler: (a) Schematic and (b) typical switching waveform on application of voltage to the electrodes.

the optical powers in the two guides is constant, but the optical power in any one is a function of the applied electric field. The two most common examples of the directional coupler type are the reversed delta-β coupler (Schmidt and Alferness, 1979) and the digital optical switch (Silberberg, et al., 1987; Burns, et al., 1976; Burns, et al., 1992; Okayama and Kawahara, 1994; Murphy, et al., 1994). The reversed delta-β coupler is compact and can be operated with modest voltages (10–20 V). The digital switch requires higher drive voltages (40–50 V) but can be constructed as a polarization-independent switch.

The small guide separation in directional coupler switches does not allow the design of traveling-wave electrodes (discussed below) with 50 Ω impedance. This is essential for high-bandwidth operation. Hence directional coupler-type switches are typically used for lower-speed switching applications, where small size and polarization diversity may be required.

A typical MZI (Alferness, 1982) configuration is shown in Fig. 4.2a. The input light is split in a Y branch. When the light from the upper and lower waveguides recombine in the output Y branch, there is either constructive or destructive interference depending on the optical path difference between the two branches. Since lithium niobate is electrooptic, the refractive index in the region of the guides can be modified by applying an electric field to them through the electrodes shown. The two parallel optical waveguide arms form two phase modulators, which operate in a push–pull manner when the electrodes are fabricated as shown. Thus by changing the voltage on the center

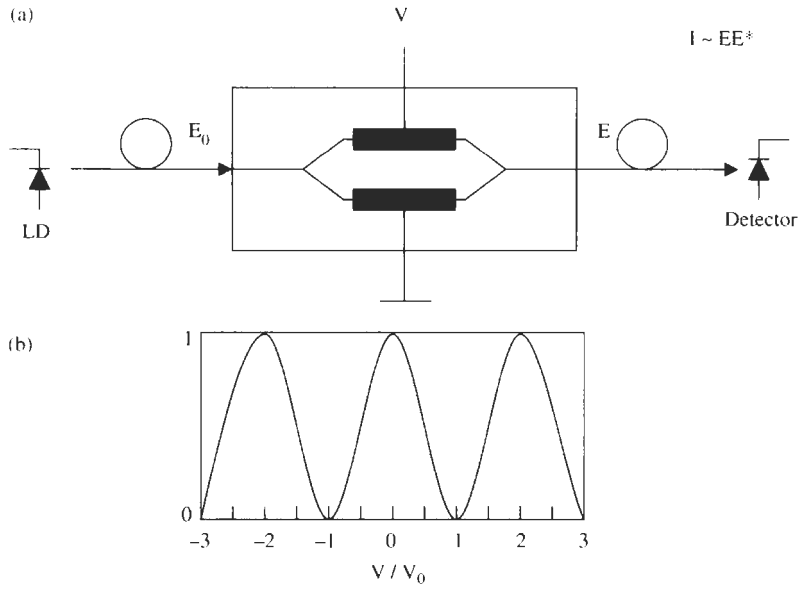


Fig. 4.2 Waveguide Mach-Zehnder interferometer: (a) Schematic and (b) theoretical switching curve.

electrode, the output can be modulated from a maximum to a minimum. The transfer curve of this amplitude modulator as a function of applied voltage is sinusoidal as shown in Fig. 4.2b.

Typically, the electrodes are several centimeters long and several tens of microns wide. Because of these small dimensions, the electrode capacitance is very small, about 1 pF/mm, which allows very large modulation bandwidths. To date, bandwidths as high as 75 GHz (Noguchi, 1994) have been demonstrated by use of novel electrode design.

4.2 LUMPED CAPACITANCE AND TRAVELING-WAVE MODULATORS

At modulation frequencies of less than 1 or 2 GHz the electrode can be modeled as a lumped capacitance with a value determined by the length and width of the electrodes, typically a few picofarads. These low-speed modulators are called lumped electrode designs. However, at higher frequencies the electrode can no longer be simulated by a lumped capacitor but must be treated as a traveling-wave electrode. In this regime, the RF energy travels along the electrode as in a transmission line while the optical energy travels in the waveguide. The dielectric constant of lithium niobate is approximately 40. It is well known that when a transmission line, such as a microstripline or coplanar waveguide, is

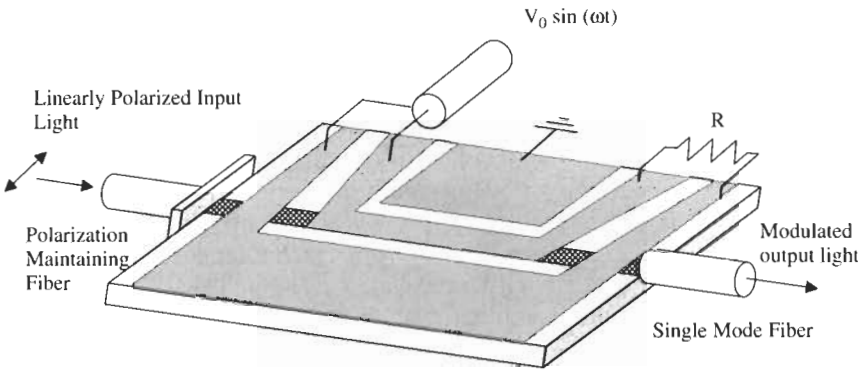


Fig. 4.3 Schematic diagram of a traveling-wave electrooptic modulator.

fabricated on a dielectric substrate (dielectric constant ε) with air overlay, the effective dielectric constant is of the order of $(\varepsilon + 1)/2$, or about 20 for lithium niobate. The RF (refractive index), which is the square root of the effective dielectric constant, is therefore about 6, while the optical refractive index is about 2.2. This translates into an optical velocity that is three times larger than the RF velocity. This limits the length of the electrodes since as the two fields walk off, the phase modulation in the first part of the electrode is canceled by that in the second part. Without any attempts at velocity matching, the product of length and modulation bandwidth for a lithium niobate modulator is about 8 GHz · cm. However, with velocity matching, this product can be increased to about 160 GHz · cm. Designs that implement velocity matching are called traveling-wave modulators and are shown schematically in Fig. 4.3. Note that the electrode is terminated with a chip resistor equal to the characteristic impedance of the transmission line, which is typically designed to be close to 50 Ω. Any mismatch in the termination and the characteristic impedance results in a reflected RF wave propagating counter to the optical field and results in degraded low-frequency response (Gopalakrishnan, 1994).

The MZI works well with high-bandwidth traveling-wave electrodes where the electrode length and gap is of the order of 20 μm and 2–4 cm respectively to keep the drive voltages low (<6 V). In this section we will discuss some of the design issues for MZIs.

4.3 CRYSTAL ORIENTATION AND BUFFER LAYERS

The first variable encountered in designing a lithium niobate modulator is the orientation of the crystal axes to the waveguides and electrodes. The crystal cut affects both modulator efficiency, as measured by the half-wave voltage V_π and modulator chirp, which is described by the chirp parameter α. Figure 3.1 shows the four most common electrode structures used in MZI-type switches. In all these configurations, the following principle is followed. To achieve the

lowest drive voltage it is essential to exploit the largest electrooptic constant, r_{33}. This requires that for any given crystal cut, the electrode position with respect to the guides must be such that the guide sees an electric field along the z-axis of the crystal. This requires that the waveguide be placed between the electrodes for an x-cut substrate and beneath the electrodes for an z-cut crystal. Because the electrodes are placed on top of the waveguide, z-cut devices always require a dielectric buffer layer between the guide and electrode to minimize attenuation of the optical field due to ohmic losses in the metal. Z-cut devices also typically employ partially conductive buffer layers and charge bleed layers to mitigate dc drift and temperature instability problems from the pyroelectric effect in lithium niobate (Moyer, et al., 1998). X-cut devices do not inherently need a buffer layer because the electrodes are not placed directly above the waveguides. However, to achieve multigigahertz operation, broadband x-cut devices do use a buffer layer to facilitate velocity matching of the RF and optical waves and to design electrodes with impedance close to 50 Ω.

4.4 MODULATION EFFICIENCY

The applied electric field and modulation efficiency of various electrode geometries can be modeled using quasi-static techniques such as finite-element or finite-difference methods (Noguchi, et al., 1998). These techniques also provide the microwave properties of the electrode (velocity, impedance, and loss). RF loss can also be adequately determined at high frequencies (≥2 GHz) with Wheeler's inductance rule (Wheeler, 1942). In general, electrodes of a thickness of 15 μm fabricated by electroplating have low RF loss of the order of $0.5\,\text{dB} \cdot \text{cm}^{-1} \cdot \text{GHz}^{-0.5}$ and enhanced velocity matching, due to the presence of electric flux in the air gap between electrodes and in the buffer layer, which has smaller dielectric constant than the lithium niobate (Rangaraj, 1992; Gopalakrishnan, 1994). Buffer layers are required for broadband velocity matching on both x-cut and z-cut devices due to the high dielectric constant of lithium niobate ($\varepsilon_{x,y} = 44, 28$) relative to the optical dielectric constant ($\varepsilon_{x,y} = 4.6, 4.9$).

The use of a shielded velocity matched design (shown in Fig. 4.4) has been demonstrated to improve velocity matching while keeping the drive voltage low (Noguchi, 1991). In these designs the height of the shielding plane above the electrodes is of the order of 10 μm. Modeling has shown (Kawano, 1991) that this height is critically related to the electrode thickness. A 40 GHz bandwidth with a drive voltage of 3.6 V at 1550 nm has been demonstrated (Noguchi, 1993).

X-cut electrode topologies result in chirp-free modulation due to the push–pull symmetry of the applied fields in the electrode gap (Koyama and Iga, 1988). In z-cut devices, the waveguide positioned underneath the hot electrode experiences an RF field flux that is more concentrated, resulting in a factor of two improvement in overlap between RF and optical field relative to x-cut

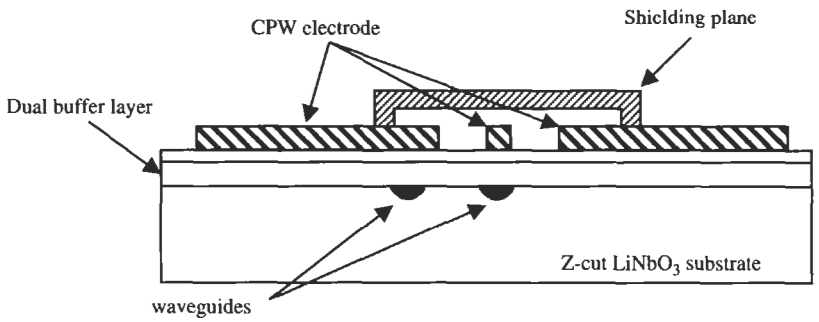


Fig. 4.4 Cross sectional view of Mach-Zehnder modulator with a shielding plane to improve velocity matching.

devices. However, the overlap under the z-cut ground electrode is reduced by a factor of three relative to x-cut devices; therefore, the overall improvement in z-cut V_π relative to the x-cut is about 20% for single drive modulators. The difference in overlap between the two z-cut waveguides results in a chirp parameter of approximately -0.7.

Employing a dual-drive design shown in Fig. 3.1(d), in which the push–pull effect is produced by the driver circuit, the factor of two improvement in overlap under the hot electrode can be utilized in both arms of the MZI at the expense of increased RF drive complexity. Balancing the two drive levels results in zero chirp operation while imbalancing them in a controlled manner provides electronic chirp control.

The electrooptic efficiency versus modulation frequency is compared in Figs. 4.5 and 4.6 for the various electrode geometries and crystal cuts of Fig. 3.1. A narrowband x-cut modulator for RZ pulsing is included as well (Hallemeier, 1999). The response curves represent electrical power at the receiver photodetector per unit of total electrical power into the modulator. For dual-drive modulators this would be the combined RF power into the two complementary drive ports. Conventional electrode and waveguide structures are assumed (Noguchi, 1998; Gopalakrishnan, 1994); that is, CPW electrodes are assumed and no ridges or Δn enhancement is utilized. RF properties and RF-optical overlap have been calculated using finite difference quasi-static field analysis. Optical mode profiles typical of conventional single-mode stripe waveguides are used in the RF-optical field overlap calculation. The electrode properties and dc electrooptic efficiency are inserted into an electrooptic response model that accounts for reflections from the RF termination (Gopalakrishnan, 1994).

Figure 4.5 shows the calculated electrooptic response for several devices designed for broadband digital or RZ pulsing transmission applications at 2.5 or 10 Gb/s using conventional electrode and waveguide structures. The electrode lengths were chosen to be representative of devices used in actual

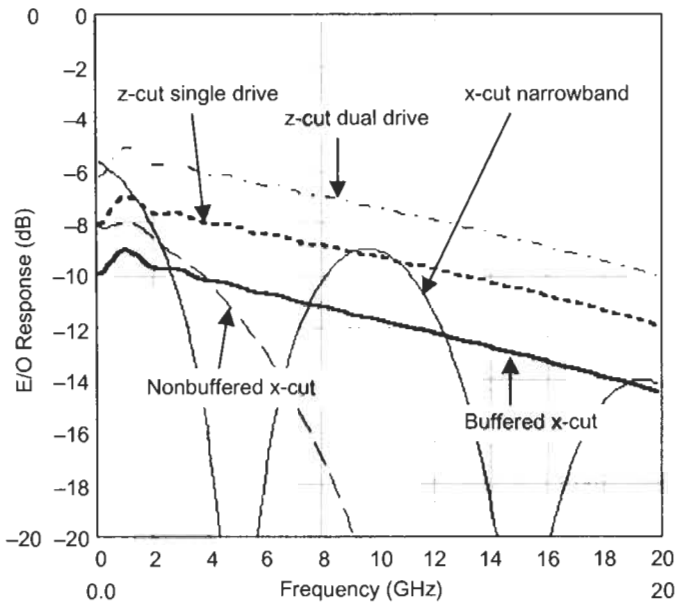


Fig. 4.5 Calculated electrooptic response for several devices designed for broad-band digital or RZ-pulsing applications at 2.5 or 10 Gb/s using conventional electrode and waveguide structures. Representative electrode lengths, shown in Table 1, were chosen for devices typically used in various applications.

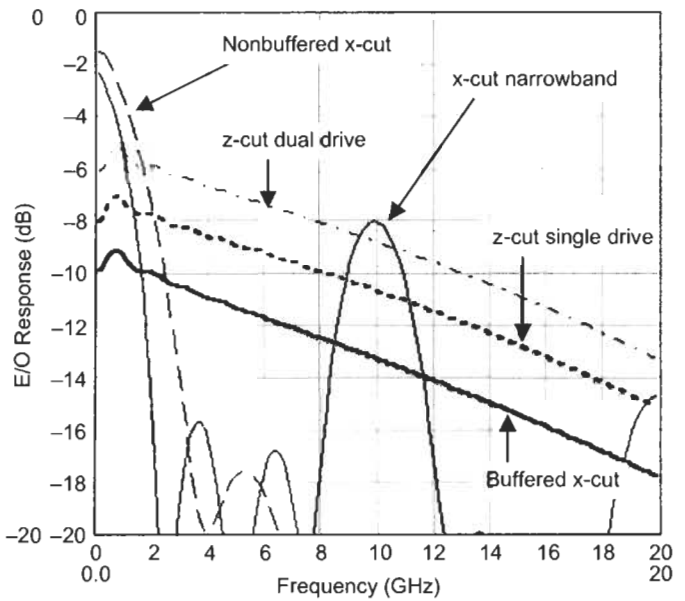


Fig. 4.6 Calculated electrooptic response for the case where all modulators have a 5-cm electrode length with conventional electrodes and waveguide structures.

applications and are shown in Table 2, normalized to the shortest electrode (nonbuffered x-cut). The application for which each device would be best suited is also listed in Table 1. Figure 4.6 shows the electrooptic response for the case where all modulators have a 5 cm electrode length. Note that the plots in Figs. 4.5 and 4.6 are of the same format as an S_{21} network analyzer measurement. RF loss from packaging is neglected in these in order that raw electrooptic performances can be compared.

Table 1 **Modulation Formats**

Modulation Technique	*Optical Spectra*	*Data Format*	*Comments*
AM–NRZ (amplitude modulated – nonreturn-to-zero)	Double sideband, with carrier	NRZ (typical)	Bandwidths typically twice the information bandwidth (or more), significant carrier power
AM–RZ (amplitude modulated – return-to-zero)	Double sideband, with carrier	RZ	Bandwidths typically 4 times the information bandwidth (or more), significant carrier power
SSB (single sideband)	Single sideband	NRZ	Bandwidths 1/2 AM bandwidths, increased dispersion tolerance
DSSC (double-sideband suppressed carrier)	Double sideband suppressed carrier	Duo-binary	Requires special modulation techniques, external modulator typically used
PM (phase modulated)	Phase modulation	PM	Used for linewidth broadening and dispersion compensation

Table 2 **Normalized Electrode Lengths Used in Calculation Shown in Fig. 4.5**

Electrode/Crystal Cut	*Application*	*Electrode Length*
Nonbuffered x-cut	2.5 Gb digital	1.0
Buffered x-cut	10 Gb digital	2.2
x-cut narrowband	10 Gb RZ pulse	1.5
z-cut single drive	10 Gb digital	2.2
z-cut dual drive	10 Gb digital	2.2

The comparison of electrooptic efficiencies in Fig. 4.5 reveals some interesting results. As expected, the z-cut dual drive is the most efficient, given its high overlap efficiency. The 4 dB advantage over the buffered x-cut is due mainly to the factor of two improvement in RF-optical overlap minus the 3 dB penalty for two RF inputs. Single drive z-cut boasts a 2 dB advantage over buffered x-cut and is surpassed slightly at 10 GHz by narrow band x-cut, which is 30% shorter. Nonbuffered x-cut begins near the efficiency of the z-cut single drive but rolls off faster due to RF-optical velocity mismatch. Note that the improved performance of the z-cut dual drive device comes at the expense of greater complexity in delivery of RF drive signals to the modulator, since the two drives must be exactly out of phase over the full bandwidth of the modulator. This is difficult to achieve at 10 Gb/s and even more so at 40 Gb/s.

The electrooptic efficiency for the case where all electrodes are 5 cm is shown in Fig. 4.6. The narrowband x-cut device is the most efficient at 10 GHz, even surpassing the dual-drive structure by 1 dB. The low RF electrode loss and high overlap efficiency of the narrowband electrodes on nonbuffered x-cut substrates account for the improved performance. In the narrowband design, the RF loss can be minimized with little constraint from velocity- or impedance-matching considerations, a degree of freedom that becomes important for long electrode lengths.

4.5 BIAS STABILITY AND TEMPERATURE PERFORMANCE

One other factor that affects modulator design is bias voltage drift. For MZI-type modulators, the optical power versus drive voltage transfer function is sinusoidal, as shown in Fig. 4.2(b), with the ideal bias point near the half-power point. Generally an active feedback circuit is used to maintain quadrature, since quality of transmitted digital data can suffer if the bias point shifts too much over time. DC drift was identified as a potential problem very early on in the history of lithium niobate modulators. Several methods have been adopted to decrease and understand dc drift (Nayyer, 1994; Chuang, 1993; Gee, 1985). The long term dc drift behavior is best described by an RC ladder model (Yamada, 1981; Korotky, 1996). The initial applied field is determined by capacitive voltage division through the ladder. The long-term field is determined by resistive voltage division. The resistors and capacitors in the model are set by process parameters and conditions used to fabricate the waveguide and electrode. Accelerated aging tests reveal that bias voltage for a device using Ti-indiffused waveguides will not change by more than a factor of two from the initial bias voltage over a 20-year lifetime at typical operating temperatures.

Long term dc drift in lithium niobate modulators is also reduced by exclusion of OH^- ions from the substrate (Nagata, 1995). Typically, when a bias voltage is first applied to the modulator, the drift is very fast for the first several hours. Then it reverses direction and changes slowly with the activation

energy of 1.0 eV. The activation energy is measured by performing elevated-temperature-accelerated testing of modulators. Typically a bias voltage change by a factor of 2 over the life of the modulator is considered acceptable. A life of 20 years at room temperature corresponds to a change in bias voltage by a factor of 2 over a few hours at a temperature above 100°C. Activation energy for x-cut modulators has been measured to be about 1.4 eV (Nagata, 2000).

5. System Requirements and Digital Performance

The introduction of wavelength-multiplexed systems to enhance data-carrying capacity of the installed fiber base has significantly increased the deployment of lithium niobate external modulators. Modulators have become a critical component in the majority of today's long-haul terrestrial and submarine optical networks.

Intensity modulation or on-off keying (OOK) is the predominant technique used for telecommunications transmission in optical fiber. An external modulator is switched to define the two states of the system. Data encoding has predominantly utilized a nonreturn-to-zero (NRZ) format, where an arbitrary data stream of ones and zeroes is directly encoded as high and low optical power levels, respectively. Advantages of OOK include minimizing the fiber link degradation due to group velocity dispersion, because NRZ data coding provides a narrow-spectrum signal, and ease of data recovery at the receiver. OOK is adequate for links operating at data rates of 2.5 Gb/s for spans up to 1000 km and 10 Gb/s for spans up to 500 km, utilizing traditional and nonzero-dispersion-shifted single-mode fibers.

A typical fiber link is shown in Fig. 5.1. The goal is to reproduce the data input exactly at the receiver output. A partial list of the several factors that can potentially corrupt the received signal, is as follows:

laser mode-hopping as a result of optical reflections;

RF amplifier distortions as a result of RF reflections from modulator (poor S11);

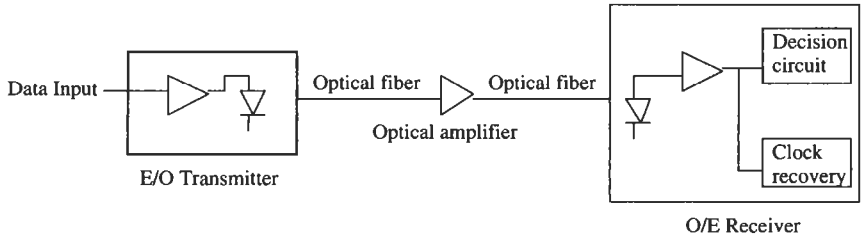


Fig. 5.1 A typical digital fiberoptic communication link.

dispersion in the fiber, resulting in spreading of the pulses;
stimulated Brillouin scattering (SBS) in the fiber;
bandwidth rolloff in the modulator (poor S12).

Consider, for example, the nonflat bandwidth of the modulator. To switch the modulator from on to off, the RF driver must provide a voltage swing equal to the switching voltage V_π of the modulator. However, V_π is a function of frequency and is a factor of 1.414 higher at the 3 dB rolloff point in the bandwidth. Therefore, the RF driver gain cannot be optimized for all frequencies.

Two measures are popularly used to quantify the performance of a fiber link. An eye diagram is the superposition of all one and zero states of a pseudorandom bit sequences (PRBS) within one bit window. Experimentally, this is done by impressing the PRBS on the modulator and using a storage scope to superpose or integrate the output of the receiver over several bits of the data stream. A typical eye diagram is shown in Fig. 5.2a and b. The larger the clean area in the center of the eye, the higher the accuracy of the receiver in differentiating between zeroes and ones. An extinction of 15 dB at the center is excellent, while 12 dB is acceptable in most links. These numbers imply that the modulator extinction measured at a single frequency anywhere in the bandwidth must be better than 20 dB. The expression for link power penalty as a function of extinction ratio is

$$P_{pen(r)} \approx -10\left[\frac{r-1}{r+1}\right]$$

where r is the linear ratio of ON power to OFF power (Ramaswami, 1998). An increase in extinction ratio from 10 dB ($r = 10$) to 20 dB ($r = 100$) results in 0.8 dB improvement in receiver sensitivity. For appropriately designed transmitters utilizing lithium niobate modulators, 20 dB extinction at frequency and >90% eye opening has been demonstrated at 2.5 Gb/s. These transmitters readily maintain >15 dB extinction under all operating conditions. At 10 Gb/s, >15 dB extinction has been achieved at frequency and >13 dB maintained over all conditions.

With the advent of erbium-doped fiber amplifiers (EDFAs), fiber loss no longer presents a fundamental limit to transmission distance and data rate. With the increase in bitrate from 2.5 Gb/s to 10 Gb/s, chromatic dispersion has become the primary impairment for transmission distances in excess of 100 km on standard single-mode (SM) fiber. Various techniques have been suggested to overcome the limitation using modulator-based low-chirp or negative-chirp devices (Lam, 1982) and dispersion compensation techniques (Ouellette, 1991; Suzuki, 1993). To solve some of today's higher bitrate problems, the ability to design a nonzero-chirp lithium niobate modulator has proved important. The design of the modulator's chirp value is used as a degree of freedom, which can be leveraged to extend link distances. Chirp has been shown to enhance

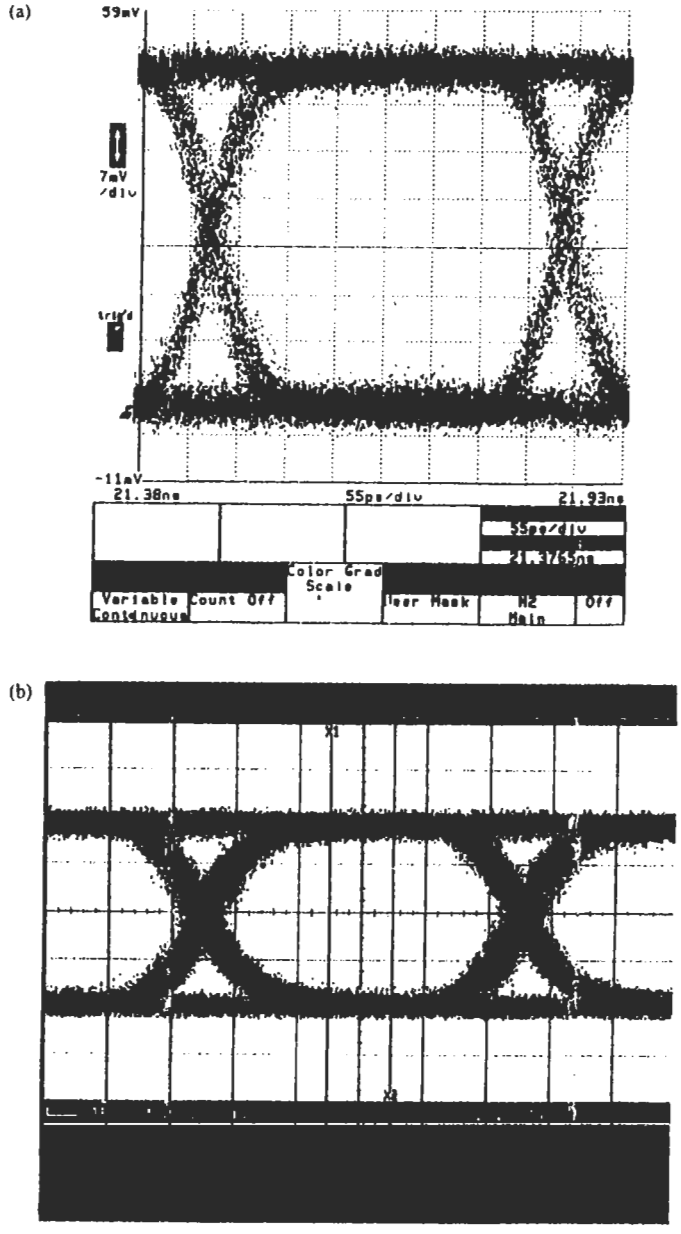


Fig. 5.2 Lithium niobate externally modulated eye diagrams at (a) 2.5 Gb/s and (b) 10 Gb/s. (Reprinted with permissions from Wooten, et al., 2000. A review of lithium niobate modulators for fiberoptic communications. J. Selected Topics Quan. Electron., V 6, No. 1, p. 69, copyright © 2000 IEEE.)

the transmission distances at 10 Gb/s by 25% (from 80 to 100 km) (Agarwal, 1995). Dispersion-shifted fibers (DSF) have been introduced to combat the shortcomings of SM fiber. DSF makes use of waveguide dispersion in order to move the zero-dispersion wavelength from about 1300 nm to 1550 nm, the center of the conventional transmission band (C-band, 1530–1560 nm). However, it is now realized that it is difficult to achieve dense wavelength-division multiplexing (DWDM) in fibers that have dispersion close to zero in the transmission window. This is primarily due to the fiber nonlinearity four-wave mixing, which can be mitigated by designing the fibers to have finite dispersion. This has led to the development of nonzero-dispersion-shifted fiber (NZ-DSF). For example, SM fiber has a dispersion of about 17 ps/nm · km, while typical NZ-DSP fiber has a dispersion of 4 ps/nm · km. As a consequence of the finite dispersion, it has been found that NZ-DSP fiber will require some dispersion compensation for the longest system lengths of about 600 km (Fells, 2000; Poole, 1994). In view of this, the importance of chirp control in every modulator to tunably compensate for dispersion for individual wavelengths will continue to grow.

Optical links also suffer from several nonlinearity problems. As link lengths and transmission bandwidths increase, higher optical powers are launched, which lead to increased manifestation of nonlinear effects. Stimulated Brillouin scattering (SBS) is the most common of these in fibers. The phenomenon was discovered in 1964 (Chioa, 1964); it was found that when an intense laser beam of frequency ω_1 passed through a crystal, a coherent acoustic wave at a frequency ω_s was produced within the crystal, while, simultaneously, an optical beam at a frequency $\omega_1 - \omega_s$ was generated. This generation occurs only above a well-defined input threshold optical power. Because of the wave-vector selection rules due to momentum conservation and the nature of the longitudinal acoustic modes of the cylindrical fiber (Shelby, 1985), the Brillouin wave at frequency $\omega_1 - \omega_s$ is backscattered, while the acoustic wave propagates forward. The threshold for SBS is increased by increasing the linewidth of the optical beam. One way to do this is by incorporation of a lithium niobate phase modulator in the link operating at several gigahertz and driven at around the voltage to produce a π radian phase modulation; this produces optical sidebands separated from the fundamental frequency by multiples of the drive frequency. The SBS threshold can be further increased by novel techniques using multiple drive tones (Hansen, et al., 1995) and depolarized sources (Howerton, 1996).

In recent years the conventional NRZ transmission format has been complemented with other data encoding formats and modulation techniques, as described in Table 1. Lithium niobate affords the advantages of a customizable modulator device technology that can be used to supply enabling building blocks for these alternative modulation formats and techniques. Recent high-bandwidth systems have used the return-to-zero format, especially those requiring long transmission distances. Dispersion-managed soliton and other narrow-pulse transmission techniques can be considered specialized versions

of RZ transmission. In conventional NRZ format, the interaction between self-phase modulation (SPM) and group velocity dispersion (GVD) causes transfer of energy from the center of the pulse toward the pulse edges (Madani, 1999; Hayee, 1999), a situation that is enhanced for a stream of ones. Cross-phase modulation (XPM) worsens for a stream of ones as well, due to the increased cross-wavelength interaction time. Use of the RZ format in a dispersion-managed system greatly reduces SPM and GVD pulse deformation, resulting in greater pulse-to-pulse consistency. The improved performance permits lower optical power levels, due to the reduced duty cycle, which in itself provides additional relief from nonlinear effects.

RZ modulation increases the demands on bandwidth and phase linearity for both the modulator and the drive electronics. Hence, modulator suppliers have designed multistage lithium niobate devices, which ease these requirements. Specifically, the bandwidth requirement for the higher spectral-width RZ format is divided into two optical stages, each with lower-frequency drives. Figure 5.3 shows an RZ transmitter topology implemented using a two-stage lithium niobate modulator. NRZ input signal and RZ transmitter data at 10 Gb/s are shown in Fig. 5.4. RZ transmitters have also been extended to include a third stage, which provides phase modulation, used to synchronously chirp the pulse train and extend even further the transmission link span.

Integration of multiple functions in series on a single chip is required for some of the complex new WDM systems. For example, an integrated phase modulator and amplitude modulator can be used to provide adjustable chirp to the modulated signal launched into a network, allowing for extended reach in networks using dispersive fiber. Another practical example of serially combining functions is the integration of the optical data modulator with an on-chip variable attenuator. Today's WDM systems require channel-to-channel variable attenuation for dynamic power equalization (Ramaswami, 1998) and full channel blocking during live-card installations (hot swaps). Today, several lithium niobate component suppliers are integrating the attenuation and modulation functions into a single package.

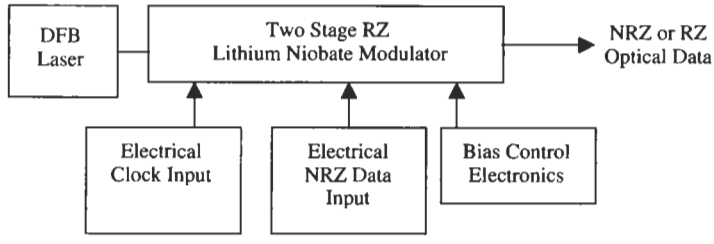


Fig. 5.3 Data modulator topology for NRZ transmission.

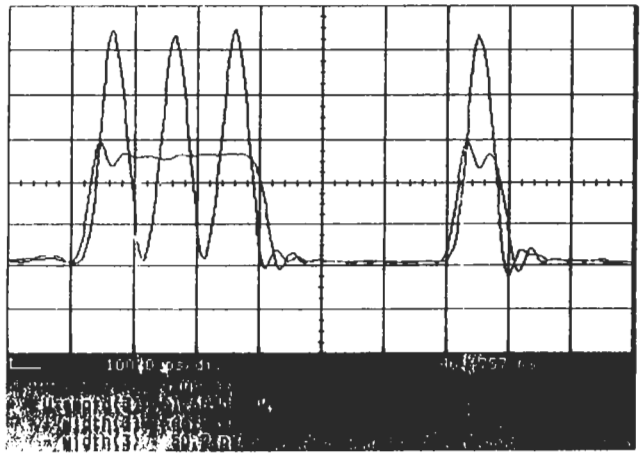


Fig. 5.4 Acquired waveforms of 10 Gb/s NRZ and RZ data. (Reprinted with permission from Wooten, et al., 2000. A review of lithium niobate modulators for fiberoptic communications. IEEE J. Selected Topics Quantum Electronics, V 6, No. 1, p. 69, copyright © 2000 IEEE.)

6. Polymeric Electrooptic Modulators

6.1 INTRODUCTION

Organic nonlinear optical (NLO) polymers are potential candidates for waveguide photonic devices such as Mach-Zehnder (MZ) modulators, directional couplers, and Y-branch structures (Prasad, 1991; Hornak, 1992). A variety of devices have been made in the past few years using NLO polymers (Girton, 1991; Van Tomme, 1991; Teng, 1992; Hikita, 1993; Thackara, 1993; Ermer, 1995; Thackara, 1995; Cross, 1989). Such devices have potential advantages over inorganic devices such as lithium niobate or compound semiconductor devices in the areas of high-speed modulation and effective bandwidth. Such advantages are due to the low dielectric constants of the polymers and the fact that they can be easily processed at low cost. Several methods of fabricating waveguides have been demonstrated, such as photobleaching, UV cross-linking to increase index, and reactive ion etching to form the channel guides. However, several obstacles to the use of polymeric electrooptic modulators in practical photonic applications exist. Single-mode optical waveguides are required for telecommunications systems to allow matching with single mode fibers. The stability of these polymers and their relatively high optical propagation losses have also been obstacles to their wide spread utilization.

In this section we summarize some of the latest results reported in the literature with respect to modulator fabrication and performance.

6.2 DESIGN CONSIDERATIONS

Several factors must be considered in the design of polymeric MZ waveguide modulators for photonic applications. First, single-mode operation is required and can be achieved by controlling the thickness of the guiding layer. The thickness of the guiding layer is inversely proportional to the refractive index difference between the guiding layer and the cladding layers. Second, a low half-wave voltage of the modulator is required for high-speed operation. This can be achieved with a high electrooptic constant. Third, optical loss needs to be considered. The total optical loss of a channel is limited by the inherent absorption loss of the guiding material. Most polymeric host materials have absorption loss of about 0.1 dB/cm for wavelengths shorter than 800 nm. However, in the telecommunication wavelength region of 1550 nm the absorption loss is about 1 dB/cm. This is caused by the overtones of the vibration spectra of the C–H bond in organic compounds. Significant work has been done in reducing this intrinsic loss by replacing the hydrogen with a heavier atomic species so that the vibration spectra move to lower energies. Several groups have demonstrated deuterated, fluorinated, and chlorinated polymers where the absorption loss has been reduced to 0.1 dB/cm in the 1550 nm range. The replacement of hydrogen by a heavier species decreases the refractive index of the polymer. For example, polymethylmethacrylate (PMMA) has an absorption loss of 0.5 dB/cm at 1550 nm, which drops to 0.07 dB/cm for 80% fluorinated PMMA (Eldada, 2000). Concomitantly, the refractive index drops from 1.5 to 1.3 (Lee, et al., 1998).

Electrooptic (EO) polymers require optically nonlinear chromophores that can be incorporated into the polymer host, aligned by a poling field, and finally hardened to maintain the alignment. The poling process aligns the dipole moments of the individual chromophore molecules so that the material exhibits a macroscopic dipole moment, which results in an electrooptic constant and is capable of high modulation bandwidths. Highly nonlinear chromophores have been synthesized and incorporated into a polymer host, but the resultant EO coefficients in devices have not been as large as expected. Recent work (Zheng, et al., 1999) has shown the tendency of chromophores to aggregate into pairs and thus not contribute to an EO effect. This can be avoided by keeping the chromophore concentration at a level of about 10% by weight in the host polymer. At these concentrations stable EO coefficients of 55 pm/V have been demonstrated.

A typical device schematic is shown in Fig. 6.1 (Steier, et al., 1999). The high-speed modulator has a Mach-Zehnder interferometer similar to those used in lithium niobate modulators. However, the high-speed traveling-wave electrodes are microstripline electrodes, unlike the coplanar waveguide electrodes used in lithium niobate (Gupta, et al., 1979). In the microstrip design, the ground plane is below the guiding layer, while the active electrode is placed exactly. The separation of the ground and active electrodes is therefore on

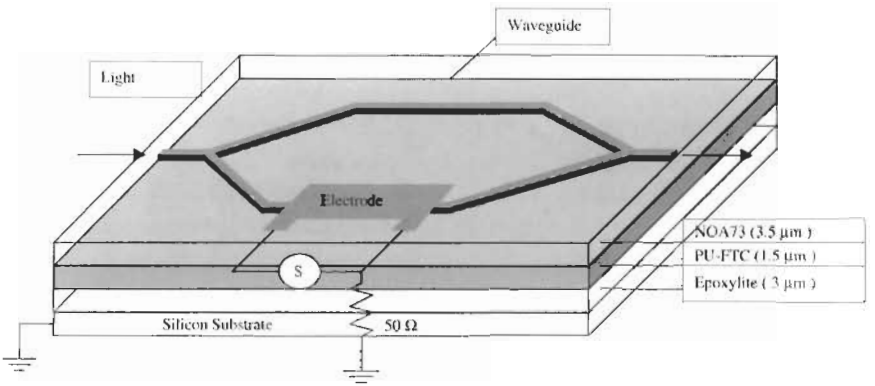


Fig. 6.1 Schematic of a polymeric Mach-Zehnder traveling wave optical modulator. The light is guided through the buried rib waveguides.

the order of the thickness of the spun-on polymer film, which is about 10 to 20 microns. This is not a viable geometry for lithium niobate, since the minimum feasible thickness of the substrate is about 250 μm. In the microstripline geometry, the electrode separation, active electrode width, and polymer dielectric constant are adjusted to achieve 50 Ω characteristic impedance. Velocity matching, as mentioned before, is easier, since the dielectric constant of polymers is about 4.

The polymer and chromophore are used in a guest/host system. A guest/host system is one in which the chromophore is not chemically attached to the polymer backbone. This differs from a side-chained or a copolymer system, in which the dye is attached to the polymer. For a guest/host system, the dye is dissolved in a solvent and mixed with the polymer. Typical steps for device fabrication using EO polymers is shown in Fig. 6.2 (Steier, et al., 1999). In this example, the polymer layers are fabricated on a silicon substrate. The EO polymer in this case uses a chromophore based on a novel tricyanobutadiene acceptor incorporating a furan-derivative ring, referred to as FTC (Wang, et al., 1998; Shi, et al., 1992). The FTC chromophore when doped into a PMMA host has an $r_{33} \cong 55\,\text{pm/V}$ @ 1060 nm. The chromophore has excellent thermal stability (300°C), and the host system (PMMA) has a modest loss of 1 dB/cm at 1300 nm. Alignment of the chromophores to achieve the EO effect is done by electrical poling. An optimum poling profile, determined empirically, is shown in Fig. 6.3. During the precuring step, the cross-linkable PU-FTC polymer is heated to a temperature such as to initiate partial cross-linking. This is to prevent surface damage during poling. There is a tradeoff between the precuring time and the poling efficiency. During the poling step, a dc voltage is applied to the corona needle with the sample temperature T_p. This temperature is such that the dipoles have enough mobility to allow alignment to the external

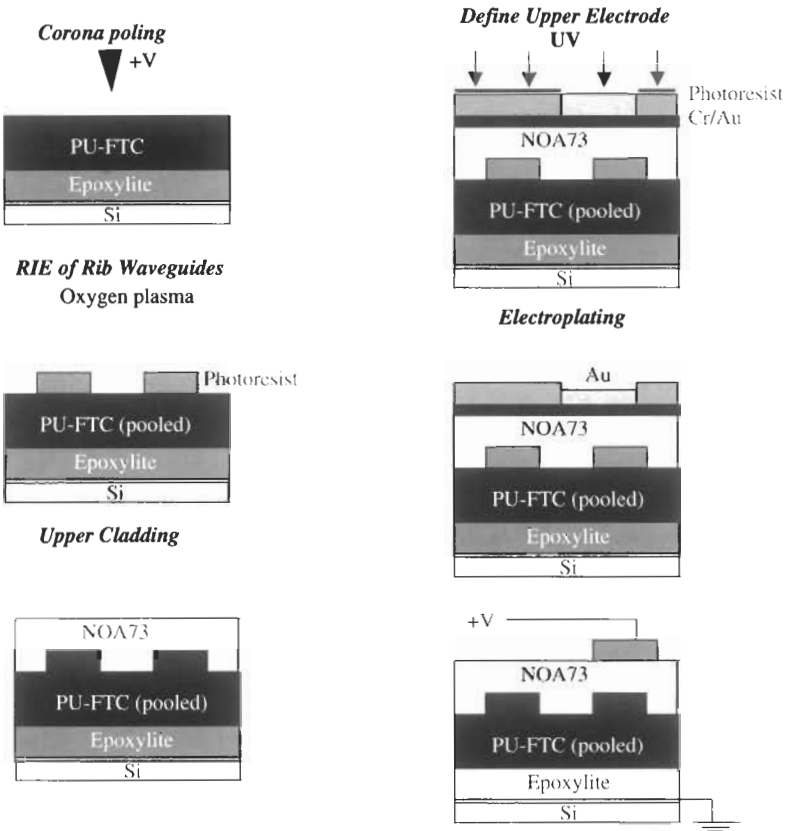


Fig. 6.2 Typical fabrication steps for a polymeric Mach-Zehnder modulator.

field and yet not so high that mobility is impeded by excessive crosslinking. Typical poling temperatures are of the order of 100° to 200°C. In comparison, note that the Curie temperature of lithium niobate is about 1150°C—hence the much higher thermal stability of the electrooptic effect in lithium niobate.

Referring to Fig. 6.2, the ground electrode is deposited directly on the silicon substrate. The lower cladding is a layer of Epoxylite 9653 ($n = 1.54$ @ 1300 nm) on top of the ground electrode. The core FTC layer is spun on the lower cladding. Lateral optical confinement is achieved by a rib structure etched by reactive ion etching (RIE) on the core layer. A commercially available UV-curable polymer, NOA73, is used as the upper cladding because of its good adhesion to electrode metals such as gold and chromium. In a typical device, the waveguide rib width is 6 μm, the core layer thickness is 1.5 μm, and the rib height is 0.3 μm for single-mode operation.

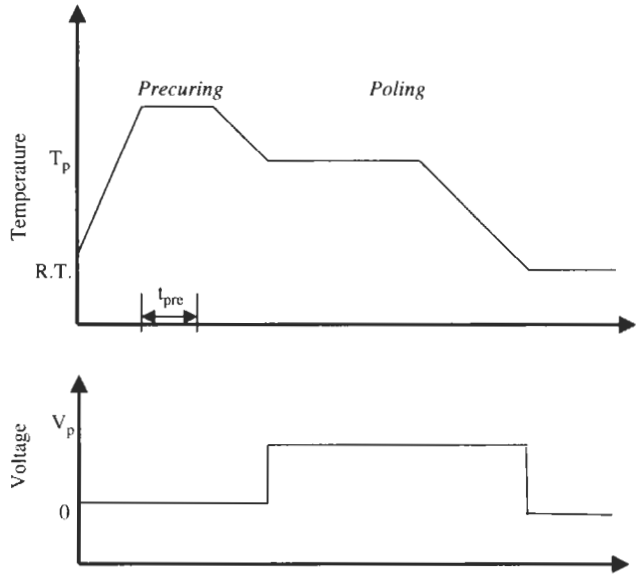


Fig. 6.3 A typical thermal poling cycle for aligning chromophores in a polymeric electrooptic modulator.

The microstrip line is formed between the lower ground electrode and the upper electrode. Its characteristic impedance was designed to be 50 Ω by setting the width of the upper electrode to be 22 μm. To reduce the microwave loss, the thickness of the upper electrode was increased to 4 μm by electroplating. The electrode length was 20 mm.

The measured low frequency V_π was 4.5 V, which corresponds to an EO coefficient of 25 pm/V. The measured bandwidth is shown in Fig. 6.4. This demonstrates some of the advantages of polymer EO modulators: The V_π is low enough, and the bandwidth extends into the millimeter wave range. A polymer EO modulator with an electrode length of 2 cm and bandwidth of 100 GHz has also been demonstrated (Chen, et al., 1997).

Trimming techniques to bias the modulator passively can be used by using the photobleaching property of EO polymers (Chen, et al., 1997). The bleaching process is an irreversible photodecomposition of the chromophores, which reduces refractive index in the bleached region (Girton, et al., 1991).

The main unresolved problem with EO polymers is temperature stability. The aligned chromophores that contribute to the EO effect tend to misalign with time and temperature. The temperature range over which polymers are stable is often characterized by the glass transition temperature T_g. If the thermal coefficient of expansion is plotted as a function of temperature, T_g is the temperature at which the slope of the curve increases discontinuously.

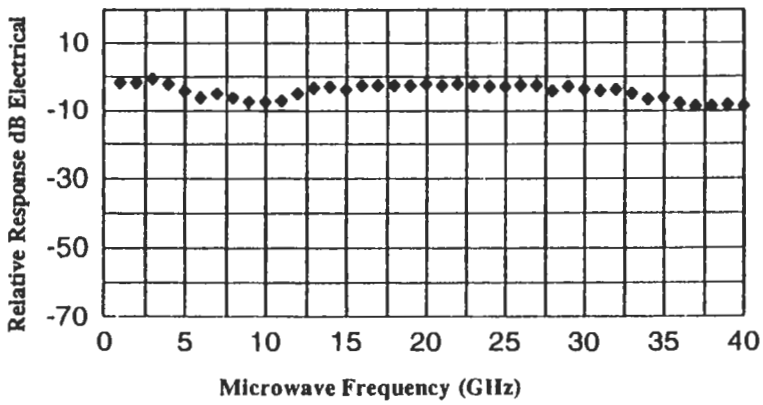


Fig. 6.4 Measured bandwidth of a polymeric electrooptic modulator with microstripline traveling-wave electrodes. (Reprinted from Chemical Physics, Volume 245, W. H. Steier, et al., Polymer electrooptic devices for integrated optics, pp. 487–506, Copyright 1999 with permission from Elsevier Science.)

T_g can also be measured by the in situ second harmonic technique (Oh, et al., 2000). T_g can be increased by using a guest chromophore with high dipole moment in a host that is particularly temperature-stable, such as PMMA.

7. Conclusion

We have reviewed the design, fabrication, and system performance of lithium niobate and polymeric modulators. The lithium niobate technology is mature, and directions for development in the future will relate primarily to manufacturing issues. Over the last several years, these modulators have gone from research laboratories to wide deployment in the telecommunications industry. Pigtailing and packaging of these modulators continues to be a major cost component of their manufacture. Therefore, future work will target manufacturing issues such as automated fiber attachment and increased manufacturing wafer yield. However, work needs to be done on reducing the drive voltage of 40 GHz modulators, since RF power becomes progressively more expensive at higher bandwidths.

Polymeric modulators, on the other hand, are not yet commercially accepted. Several polymeric materials systems with attractive electrooptic constants have been demonstrated. Long-term stability and thermal stability continue to be issues of concern.

References

Agarwal, G. P., 1995. Non-Linear Fiber Optics, 2nd ed. New York: Academic Press.

Alferness, R. C., 1982. Waveguide electrooptic modulators. IEEE Trans. Microwave Theory Tech., V MTT-30, pp. 1121–1137.

Becker, R. A., and Kincaid, B. E., 1993. Improved electrooptic efficiency in guided wave modulators. J. Lightwave Tech., V 11, pp. 2076–2079.

Burns, W. K., 1992. Shaping the digital switch. IEEE Photon. Technol. Lett., V 4, pp. 861–863.

Burns, W. K., Klein, P. H., West, E. J., and Plew, L. E., 1979. Ti diffusion in Ti/$LiNbO_3$ planar and channel optical waveguides. J. Appl. Phys., V 50, pp. 6175–6182.

Burns, W. K., Lee, A. B., and Milton, A. F., 1992. Active branching waveguide modulator. Appl. Phys. Lett., V 29, pp. 790–792.

Chen, A., Chuyanov, V., Marti-Carrera, F. I., Garner, S., Steier, W. H., Mao, S. S. H., Ra, Y., and Dalton, L. R., 1997. Trimming of polymer waveguide Y-junction by rapid photobleaching for tuning the power splitting ratio. IEEE J. Lightwave Technology Lett., V 9, No. 11, pp. 1499–1501.

Chen, D., Fetterman, H. R., Chen, A., Steier, W. H., Dalton, L. R., Wang, W., and Shi, Y., 1997. Demonstration of 110 GHz electrooptic polymer modulators. Appl. Phys. Lett., V 70, pp. 3335–3337.

Chioa, R. Y., Townes, C. H., and Stoicheff, B. P., 1964. Stimulated Brillouin scattering and coherent generation of intense hypersonic waves. Phys. Rev. Lett., V 12, p. 592.

Chuang, W., Chang, W., Lee, W., Leu, J., and Wang, W., 1993. A comparison of the performance of $LiNbO_3$ traveling-wave phase modulators with various dielectric buffer layers. J. Opt. Comm., V 14, No. 4, pp. 142–148.

Cross, G. H., Donaldson, A., Gymer, R. W., Mann, S., Parsons, N. J., Haas, D. R., Man, H. T., and Yoon, H. N., 1989. Polymeric integrated electrooptic modulators. SPIE 1177, p. 307.

Dolfi, D. W., and Ranganath, T. R., 1992. 50 GHZ Velocity matched, broad wavelength $LiNbO_3$ modulator with multimode active section. Electron. Lett., V 28, pp. 1197–1198.

Dutta, N. K., and Olsson, N. A., 1984. Electroabsorption in InGaAsP-InP double heterostructures. Electron. Lett., V 20, pp. 634–635.

Eldada, L., 2000. Advances in polymer integrated optics, IEEE J. Selected Topics in Quantum Electronics, V 6, No. 1, pp. 54–68.

Ermer, S., Anderson, W. W., Van Eck, T. E., and Girton, D. G., 1995. Organic Thin Films for Photonic Applications, Portland, OR, USA, WA2-1, p. 285.

Fells, J. A. J., et al., 2000. Twin fiber grating adjustable dispersion compensator for 40 Gbs. ECOC 2000, PD- 2.4.

Fouchet, S., 1987. Wavelength dispersion of Ti induced refractive index change in $LiNbO_3$ as a function of diffusion parameters. J. Lightwave Technol., V LT-5, No. 5.

Gee, C. M., Thurmond, G. D., Blauvelt, H., and Yen, H. W., 1985. Minimizing dc drift in $LiNbO_3$ waveguide devices. Appl. Phys. Lett., V 47, No. 3, pp. 211–213.

Girton, D. G., Kwiatlowski, S. L., Lipscomb, G. F., and Lytel, R. S., 1991. 20 GHz electrooptic polymer Mach-Zehnder modulator. Appl. Phys. Lett., V 58, No. 16, p. 1730.

Gopalakrishnan, G. K., Burns, W. K., McElhanon, R. W., Bulmer, C. H., and Greenblatt, A. S., 1994. Performance and modeling of broadband $LiNbO_3$ traveling wave optical intensity modulators. J. Lightwave Technol., V 12, No. 10, pp. 1807–1818.

Griffin, R. A., Walker, R. G., Johnstone, R. I., Harris, R., Perney, N. M. B., Whitbread, N. D., Widdowson, T., and Harper, P., 2001. Integrated 10 Gb/s chirped return-to-zero transmitter using GaAs/AlGaAs modulators. Optical Fiber Communication Conference 2001, Anaheim, CA, postdeadline paper PD15.

Gupta, K. C., Garg, R., and Bahl, I. J., 1979. Microstrip Lines and Slotlines. Norwood, MA: Artech House.

Hallemeier, P., Kissa, K., McBrien, G., and Horton, T., 1999. Next generation 10 Gb/s lithium niobate modulator components for RZ based transmission techniques. Proc. National Fiberoptic Engineers Conf., Sept. 1999.

Hansen, P. B., et al., 1995. 2.4488 Gb/s unrepeatered transmission over 529 km using remotely pumped post- and preamplifiers, forward error correction, and dispersion compensation. Technical Digest Conference on Optical Fiber Communication, San Diego, postdeadline paper PD25. Washington, DC: Optical Society of America.

Hayee, M. I., and Wilner, A. E., 1999. NRZ versus RZ in 10–40 Gb/s dispersion-managed WDM transmission systems. J. Lightwave Technol., V 11, pp. 991–993.

Hikita, M., Shuto, Y., Amano, M., Yoshimura, R., Tomaru, S., and Kozawaguchi, H., 1993. Optical intensity modulation in a vertically stacked coupler incorporating electrooptic polymer. Appl. Phys. Lett., V 63, No. 9, p. 1161.

Homes, R. J., and Smyth, D. M., 1984. Titanium indiffusion into $LiNbO_3$ as a function of stoichiometry. J. Appl. Phys., V 55, pp. 3531–3535.

Hornak, L. A. (ed.), 1992. Polymers for lightwave and integrated optics. Dekker, New York.

Howerton, M. M., Burns, W. K., and Gopalakrishnan, G. K., 1996. SBS suppression using a depolarized source for high power fiber applications. J. Lightwave Technol., V 14, No. 3, pp. 417–422.

Howerton, M. M., Burns, W. K., Skeath, P. R., and Greenblatt, A. S., 1991. Dependence of refractive index on hydrogen concentration in proton exchanged $LiNbO_3$. IEEE J. Quantum Electron., V 27, No. 3, pp. 593–601.

Jackel, J. L., Rice, C. E., and Veselka, J., 1982. Proton exchange for high index waveguides in $LiNbO_3$. Appl. Phys. Lett., V 41, pp. 607–608.

Johnson, J. E., et al., 1994. Low-chirp integrated EA modulator/DFB laser grown by selective-area MOVPE. 14th IEEE International Semiconductor Laser Conference, Maui, Hawaii, September, Paper M4.7, pp. 41–42.

Kaminow, I. P., and Stulz, L. W., 1978. Loss in cleaved Ti-diffused $LiNbO_3$ waveguides. Appl. Phys. Lett., V 33, No. 1, pp. 62–64.

Kawano, K., Noguchi, K., Kitoh, T., and Miyazawa, H., 1991. A finite element method (FEM) analysis of a shielded velocity-matched Ti : LiNbO3 optical modulator. IEEE Photonics Technol. Lett., V 3, No. 10, pp. 919–921.

Korotky, S. K., and Veselka, J. J., 1996. An RC analysis of long term Ti : $LiNbO_3$ bias stability. J. Lightwave Tech., V 14, pp. 2687–2697.

Koyama, K., and Iga, K., 1988. Frequency chirping in external modulators. J. Lightwave Tech., V 6, p. 87.

Lam, D. K. W., Garside, B. K., and Hill, K. O., 1982. Dispersion cancellation using optical fiber filters. Opt. Lett., V 7, pp. 291–293.

Lee, H. J., Lee, E. M., Lee, M. H., Oh, M. C., Ahn, J. H., Han, S. G., and Kim, H. G., 1998. Crosslinkable fluorinated polyarylene ethers bearing phenyl ethynyl moiety for low-loss polymer optical waveguide devices. J. Polymer Science, Part A: Polymer Chemistry, V 36, pp. 2881–2887.

Maack, D., 1999. Reliability of lithium niobate Mach-Zehnder modulators for digital optical fiber telecommunication systems. Proc. SPIE Critical Reviews: Reliability of Optical Fibers and Optical Fiber Systems, Boston, MA, pp. 197–230.

Madani, F. M., and Kikuchi, K., 1999. Design of long distance WDM dispersion managed transmission system. J. Lightwave Technol., V 17, pp. 1326–1335.

Marcuse, D., 1981. Optimal electrode design for integrated optics modulators, IEEE J. Quan. Elec., V QE-18, No. 3, pp. 393–398.

Minaka, M., Saito, S., and Shibata, M., 1979. Two dimensional distribution of refractive index changes in Ti-diffused $LiNbO_3$ strip waveguides. J. Appl. Phys., V 59, pp. 3063–3067.

Moyer R. S., Grencavich, R., Judd F. F., Kreshner, R. C., Minford, W. J., and Smith, R. W., 1998. Design and qualification of hermetically packaged lithium niobate optical modulator. IEEE Trans. Comp., Pack., Mant., Technol. B, V 21, pp. 130–135.

Murphy, E. J., 1997. Photonic Switching in Optical Fiber Telecomminications IIIB, Kaminow and Koch (eds.), New York: Academic Press.

Murphy, T. O., Murphy, E. J., and Irvin, R. W., 1994. An 8×8 Ti : $LiNbO_3$ polarization independent photonic switch. Proc. Photonics in Switching/ECOC, Florence, Italy, pp. 174–176.

Nagata, H., 2000. Activation energy of dc-drift of x-cut $LiNbO_3$ optical intensity modulators. IEEE Photonics Technol. Lett., V 12, No. 4, pp. 386–388.

Nagata, H., and Ichikawa, J., 1995. Progress and problems in reliability of Ti : $LiNbO_3$ optical intensity modulators. Optical Engineering, V 34, No. 11, pp. 3284–3293.

Nayyer, J., and Nagata, H., 1994. Suppression of thermal drifts of high speed Ti : $LiNbO_3$ optical modulators. IEEE Photonics Technol. Lett., V 6, No. 8, pp. 952–955.

Noda, Y., Suzuki, M., Kushiro, Y., and Akiba, S., 1985. 1.6 GHz electroabsorption light modulation in InGaAsP/InP double-heterostructures. Electron. Lett., V 21, pp. 1182–1183.

Noda, Y., Suzuki, M., Kushiro, Y., and Akiba, S., 1986. High speed electroabsorption modulator with strip loaded GaInAsP planar waveguide. J. Lightwave Technol., V LT-14, pp. 1445–1453.

Noguchi, K., Kawano, K., Nozawa, T., and Suzuki, T., 1991. A Ti : $LiNbO_3$ optical intensity modulator with more than 20 GHz Bandwidth and 5.2 V driving voltage. IEEE Photonics Technol. Lett., V 3, No. 4, pp. 333–335.

Noguchi, K., Mitomi, O., Kawano, K., and Yanagibashi, M., 1993. Highly efficient 40-GHz bandwidth Ti : $LiNbO_3$ optical modulator employing ridge structure. IEEE Photonics Technol. Lett., V 5, No. 1, pp. 52–54.

Noguchi, K., Mitomi, O., and Miyazawa, H., 1998. Millimeter-wave Ti : $LiNbO_3$ optical modulators. J. Lightwave Tech., V 16, No. 4, pp. 615–619.

Noguchi, K., Miyazawa, H., and Mitomi, O., 1994. 75-GHz Ti : $LiNbO_3$ optical modulator. Optical Fiber Conference 1994. Technical Digest, Paper WB3.

O'Donnell, A., 1995. Packaging and reliability of active optical components. Proc. 7th Eur. Conf. Int. Opt. (ECIO '95), 1995, p. 585.

Oh, M. C., Zhang, H., Szep, A., Chuyanov, V., and Steier, W. H., 2000. Electro-optic polymer modulators for 1.55 μm wavelength using phenyltetraene bridged chromophore in polycarbonate. App. Phys. Lett., V 76, No. 24, pp. 3525–3527.

Okayama, H., and Kawahara, M., 1994. Prototype 32 × 32 optical switch matrix. Electron. Lett., V 30, pp. 1128–1129.

Onaka, H., Miyata, H., Ishikawa, G., Otsuka, K., Ooi, H., Kai, Y., Kinoshita, S., Seino, M., Nishimoto, H., and Chikama, T., 1996. 1.1 Tb/s WDM transmission over a 150 km 1.3 μm zero-dispersion single mode fiber. Proc. OFC '96, San Jose, CA. Postdeadline paper 19.

Ouagazzaden, A., et al., 2001. 40 Gb/s tandem electro-absorption modulator. Optical Fiber Communication Conference 2001, Anaheim, CA, postdeadline paper PD14.

Ouellette, F., 1991. All-fiber filter for efficient dispersion compensation. Opt. Lett., V 16, pp. 303–306.

Pearsall, T. P., Chiang, S., and Schmidt, R. V., 1976. Study of titanium diffusion in lithium niobate low-loss optical waveguides by X-ray photoelectron spectroscopy. J. Appl. Phys., V 47, pp. 4794–4797.

Plastow, R., 1998. Tunable lasers enable new optical networks to meet changing demands. Lightwave, December, pp. 805–806.

Poole, C. D., Wisenfeld, J. M., Giovanni, D. J., and Vengsarkar, A. M., 1994. Optical fiber-based dispersion compensation using higher order modes near cut-off. J. Lightwave Technol., V 12, pp. 1746–1758.

Prasad, P. N., and Williams, D. J., 1991. Introduction to non-linear optical effects in molecules and polymers. New York: John Wiley.

Ramaswami, R., and Sivarajan, K., 1998. Optical Networks: A Practical Perspective. Los Altos, CA: Morgan Kaufman.

Ramer, O. G., 1982. Integrated optic electrooptic modulator electrode analysis, IEEE J. Quan. Elec., V QE-18, No. 3, pp. 386–392.

Rangaraj, M., Hosoi, T., and Kondo, M., 1992. A wide-band Ti : $LiNbO_3$ optical modulator with a conventional coplanar waveguide type electrode. IEEE Photonics Technol. Lett., V 4, No. 9, pp. 1020–1022.

Reinhart, F. K., 1976. Electroabsorption in $Al_yGa_{1-y}As$-$Al_xGa_{1-x}As$ double heterostructure. Appl. Phys. Lett., V 22, pp. 372–374.

Rice, C. E., and Jackel, J. L., 1984. Structural changes with composition and temperature in rhombohedral $Li_{1-x}H_xNbO_3$. Mat. Res. Bull., V 19, pp. 591–597.

Runge, P. K., and Bergano, N. S., 1988. Undersea cable transmission systems in optical fiber telecommunications II, S. E. Miller and I. P. Kaminow (eds.), pp. 879–909. Boston: Academic Press.

Savage, A., 1966. Pyroelectricity and spontaneous polarization in $LiNbO_3$. J. of Appl. Phys., V 37, pp. 3071–3072.

Schmidt, R. V., and Alferness, R. C., 1979. Directional coupler switches, modulators and filters using alternating techniques. IEEE Trans. Circuits Syst., V CAS-26, pp. 1099–1108.

Schmidt, R. V., and Kaminow, I. P., 1974. Metal in-diffused optical waveguides in $LiNbO_3$. Appl. Phys. Lett., V 25, pp. 458–460.

Shelby, R. M., Levenson, D., and Bayer, P. W., 1985. Guided acoustic wave Brillouin scattering. Phys. Rev. B., V 31, No. 8, pp. 5244–5252.

Silberberg, Y., Perlmutter, P., and Baran, J. E., 1987. Digital optical switch. Appl. Phys. Lett., V 51, pp. 1230–1232.

Steier, W. H., et al., 1999. Polymer electro-optic devices for integrated optics. Chemical Physics, V 245, pp. 487–506.

Stillman, G. E., Wolfe, C. M., Bozler, C. O., and Rossi, J. A., 1976. Electroabsorption in GaAs and its applications to waveguide detectors and modulators. Appl. Phys. Lett., V 28, pp. 544–546.

Stultz, L. W., 1979. Titanium in-diffused optical waveguide fabrication. Appl. Opt., V 18, pp. 2041–2044.

Suchoski, P. G., Findakly, T. K., and Leonberger, J., 1988. Stable low-loss proton-exchanged $LiNbO_3$ waveguide devices with no electrooptic degradation. Opt. Lett., V 13, pp. 1050–1052.

Sugii, K., Fukuma, M., and Iwasaki, H., 1978. A study of titanium diffusion into $LiNbO_3$ waveguides by electron probe analysis and X-ray diffraction methods. J. Mat. Sci., V 13, pp. 523–533.

Suzuki, N., and Ozeki, T., 1993. Simultaneous compensation of laser chirp, Kerr effect, and dispersion in 10 Gb/s long-haul transmission systems. J. Lightwave Technol., V 9, pp. 1486–1494.

Teng, C. C. 1992. Appl. Phys. Lett., V 60, No. 13, p. 1538.

Thackara, J. I., Bjorklund, G. C., Fleming, W., Jurich, M., Smith, B. A., and Swalen, J. D., 1993. Polymeric electrooptic phase modulator for broadband data transmission. SPIE V 2025, p. 564.

Thackara, J. I., Chon, J. C., Bjorklund, G. C., Volksen, W., and Burrland, D. M., 1995. Appl. Phys. Lett., V 67, No. 26, p. 3874.

Van Tomme, E., van Daele, P., Baets, R., Mohlmann, G. R., and Diemeer, M. B. J., 1991. Guided wave modulators and switches fabricated in electrooptic polymers. J. Appl. Phys., V 69, No. 9, p. 6273.

Walker, R. G., 1991. High speed III-V semiconductor modulators. IEEE J. Quantum Electronics, V 27, No. 3, pp. 654–667.

Walker, R. G., 1995. Electro-optic modulation at mm-wave frequencies in GaAs/AlGaAs guided wave devices. IEEE Lasers and Electro-optics Society 1995 Annual Meeting (LEOS 1995), San Francisco, Oct–Nov 1995 (IEEE cat 95 CH35739), V 1, pp. 118–119.

Wheeler, H. A., 1942. Formulas for skin effect. Proc. IRE, V 16, pp. 412–424.

Yamada, S., and Minakata, M., 1981. Dc drift phenomena in $LiNbO_3$ guided-wave devices. Jap. J. Appl. Phys., V 20, p. 733.

Yi-Yan, A., 1983. Index instabilities in proton exchanged $LiNbO_3$ waveguides. Appl. Phys. Lett., B42, pp. 633–635.

Zhang B., Forouhar, S., Huang, S. Y., and Chang, W. S. C., 1984. C_2F_6 reactive ion beam etching of $LiNbO_3$ and Nb_2O_5 and their application to optical waveguides. J. Lightwave Tech., V LT-2, No. 4, p. 528.

Zhang, C., Ren, A. S., Wang, F., Dalton, L. R., Lee, S. S., Garner, S. M., and Steier, W. H., 1999. Polymer preparation, Am. Chem. Soc. Div. Polym. Chem., V 40, p. 49.

Chapter 7 | Optical Switching in Transport Networks: Applications, Requirements, Architectures, Technologies, and Solutions

Daniel Y. Al-Salameh*, Steven K. Korotky, David S. Levy, Timothy O. Murphy, Sunita H. Patel, Gaylord W. Richards, and Eric S. Tentarelli

Bell Laboratories, Lucent Technologies, Holmdel, New Jersey

1. Introduction

1.1 OPTICAL SWITCHING FOR OPTICAL CROSS-CONNECTS

Optical switches and cross-connects for fiber communication have been a topic of applied research and development now for more than two decades. Indeed, the history of these technologies parallels the events leading to the milestone deployments of other optical elements into long-distance fiber transmission for communications, such as single-mode optical fiber, injection lasers, optical amplifiers, and external modulators. Since the publication of *Optical Fiber Telecommunications III* [Kaminow and Koch, 1997], the coincidence of the maturity and performance of state-of-the-art optical switches and cross-connects and of the emerging need for high-capacity network elements for managing many terabits/second of long-haul traffic has resulted in optics being selected for several important roles in switching in the backbones of the world's communication infrastructure. Now, because of their natural ability to controllably redirect extremely large bandwidths with very high fidelity, optical switches find application in component-sparing, flexible-wavelength add/drop, and small- and large-port-count optical cross-connects. In this chapter we focus on the emerging core optical transport layer in the communication hierarchy, on optical layer cross-connects and the role they are beginning to serve in managing bandwidth in the converging global voice and data communication transport networks, and on the optical switching technologies considered for the engines of these critical network elements.

---

* Present address: JDS-Uniphase.

OPTICAL FIBER TELECOMMUNICATIONS.
VOLUME IVA

ISBN: 0-12-395172-0

1.2 EVOLUTION OF OPTICS IN DIGITAL TRANSPORT – OPTICAL LINE SYSTEMS

The transmission capacity of wired communication systems has been increasing since their inception. From telegraph systems carrying a few clicks per second a century ago, to the state-of-the-art, long-distance, dense wavelength-division multiplexing (DWDM) fiber systems carrying over one Terabit per second today, the capacity of wired long-haul systems has burgeoned at an average annual increase of approximately 1/3—or a doubling approximately every two years. There is no indication that this historical trend of wired transmission capacity growth will end any time soon. Indeed, innovation in optics has been a key enabler to sustain and accelerate this trend. Since the groundbreaking introduction of commercial single-mode optical fiber systems about 20 years ago, the average rate has increased to a doubling about every 18 months. Owing to increased competition, the short-term trend for the past five years has been nothing short of phenomenal with the pace of the carrying capacity of a single fiber doubling every year. If it were not for our immediate preoccupation with describing the events of the recent past and how to achieve and manage the next factor of two to four, the challenges and implications of sustaining this astonishing rate are themselves subjects that warrant considerable attention. As we shall see, each major increase of carrying capacity has been made possible by the often revolutionary introduction of optical components to perform system functions analogous to those carried out by the electrical and electronic systems of the past.

As illustrated in Fig. 1.1, new optical devices and capability have played a significant role in the evolution of wired long-haul communication transmission since the late 1970s, beginning with the introduction of the light-emitting diode (LED) and multimode optical fiber (MMF), and have been an essential ingredient of progress ever since. By encoding data as the presence or absence of light and using glass in the form of MMF as the transmission medium, single-channel capacities of 45–90 Mb/s and transmission distances of approximately 10 km were possible (Fig. 1.1(b)). Prior to that, digital communication was entirely the province of electronics, and transmission was carried over electrical coaxial cable at maximum bitrates of 565 Mb/s and distances of 1 km between regenerators, as well as by wireless transmission (Fig. 1.1(a)). By the mid- to late 1980s, single-mode fiber and injection lasers permitted transmission of 565 Mb/s to 2.5 Gb/s over distances of 100 km without regeneration. It is worth noting that the first use of optical switches in fiber communication was during this period; they were used for the sparing of injection lasers in submarine systems, as the expected life of a semiconductor injection laser—being newly introduced—was not confidently known to exceed the expected system life of more than 20 years.

After the displacement of electrical coax by optical fiber and optoelectronic transmitter/receivers, the next extraordinary advance in optical fiber

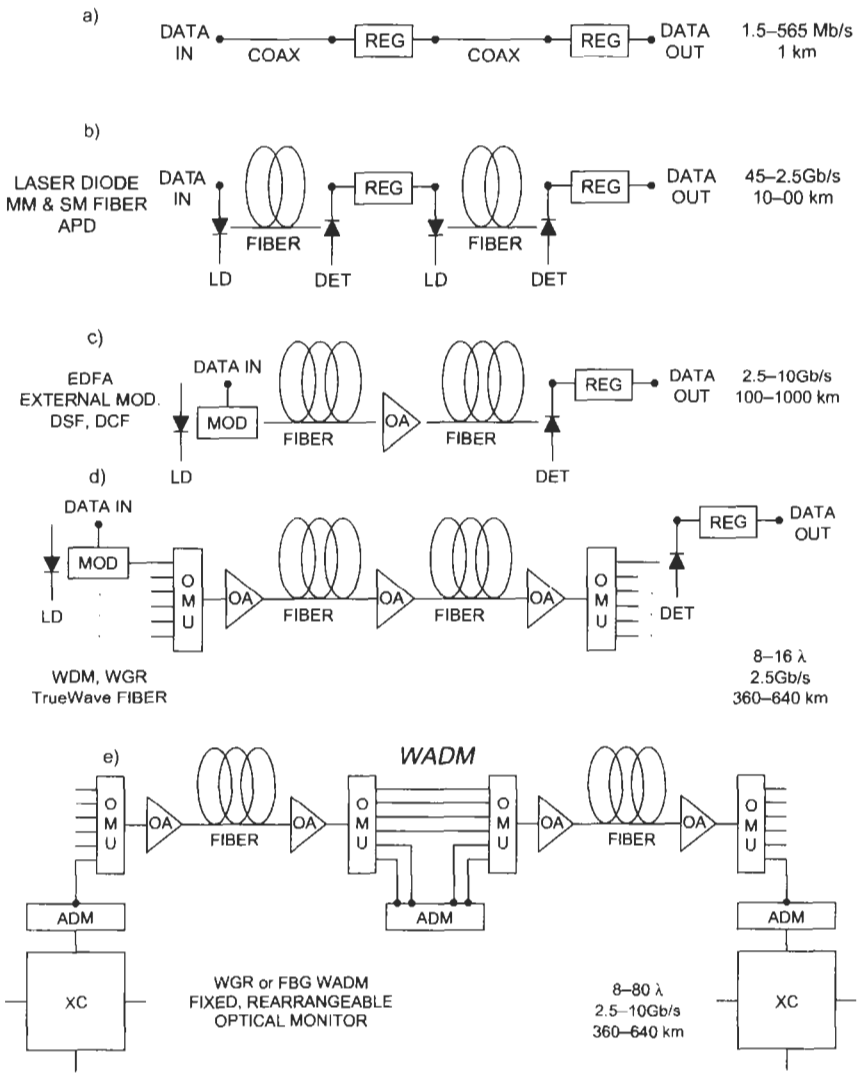


Fig. 1.1 Evolution of optics in digital transport: transmission.

communication occurred in the early 1990s with the introduction of practical optical amplifiers, Fig. 1.1(c). These optical amplifiers—in the form of erbium-doped fiber—obviated the need for frequent electronic regeneration. So effective were erbium-doped fiber amplifiers (EDFAs) that the transmission reach attainable before electronic regeneration was no longer determined by the optical loss of the transmission fiber. Rather, pulse dispersion and laser chirp became the limiting factor. This situation in turn drove the development

and deployment of more refined transmitters based on external modulation [Heismann, et al., 1997]. Using optical amplifiers and lithium niobate external modulators, transmission distances approaching 1000 km could be achieved at 2.5 and 10 Gb/s over standard and dispersion-shifted fiber, respectively.

Following closely on the heels of the introduction of the erbium-doped fiber amplifier and transform-limited transmitters was the breakthrough and disruptive introduction of commercial dense wavelength-division multiplexing (DWDM), Fig. 1.1(d), which took place in the mid-1990s [Kowalenko, 2000]. The DWDM system combines (multiplexes) multiple optical signals so that they can be amplified and transported over a single fiber as a group, and separates (demultiplexes) the individual wavelength channels at the receiving end. This capacity-enhancing capability was made possible by advances in optical waveguide technology—integrated optics or photonic integrated circuits, which permitted wavelength demultiplexing with close channel spacing and low crosstalk and is analogous to electronic multiplexer/demultiplexers. A significant increase in the number of wavelengths that could be transported was also made possible by improvements in erbium-doped fiber amplifier technology and optical filter technology. The combination of broadband optical amplification and DWDM has had a revolutionary effect on fiber communication, as it provided not only significant increases in system capacity but also a substantial reduction in cost. This is because using DWDM, a single optical amplifier could simultaneously amplify all the wavelength channels carried on an optical fiber, and the distance between repeaters was increased several fold in comparison to nonoptically amplified systems [Jackman, et al., 1999].

Initially, DWDM offered a more economical way to increase transport capacity without upgrading the existing fiber plant or terminal equipment by avoiding the signal degradation caused by dispersion at higher TDM bitrates. It also offered the flexibility to expand capacity within any portion of a network by making many more channels available per fiber [Hersey and Lumish, 1997]. To handle traffic growth, network providers have deployed such DWDM systems, which can transmit multiple channels simultaneously over one fiber, each wavelength carrying 2.5 or 10 Gb/s SONET or SDH channels over distances of ~700 km before requiring regeneration. Initial generations of DWDM systems that were brought to the market beginning in 1995 could transmit 2–16 wavelength channels simultaneously. Today's commercial systems can carry 128 channels 10 Gb/s—or more than 1 Tb/s—over unregenerated distances of several thousand kilometers. In the near future, the bitrate will increase to 40 Gb/s or higher per wavelength to support the anticipated increased usage.

Early DWDM line systems transported optical signals directly between electronic end terminals (ET), i.e., point-to-point, without access to wavelengths at intermediate locations. However, in the late 1990s, improvements in optical amplifier, multiplexer/demultiplxer, filter, signal monitoring, and fiber technology provided the capability to expand the DWDM capacity to 40–120 wavelengths and extend the transmission distances to 1000 km at bitrates of

2.5–10 Gb/s. With this came the need to access optical capacity at locations between major nodes. To provide intermediate access, linear optical wavelength add/drop multiplexers (OADM) were incorporated at sites between end terminals, as shown in Fig. 1.1(e). The first OADMs introduced were fixed types configured manually at installation. Now, a further step in the evolution of optics in digital transport is occurring with the implementation of automated flexible and rearrangeable OADMs, which allow network operators to remotely provision the addition and removal of individual wavelengths. These networking features require optical switches to provide the desired remotely programmable reconfigurability.

Today, long-distance global backbones, as well as regional networks, are virtually entirely based on fiberoptic cable. Among the few exceptions are free-space radio and optical systems, which are used to traverse terrain where running cable is prohibitively expensive. In Fig. 1.2 we illustrate a prototypical fiber transport network that is indicative of the major interexchange carrier (IXC) networks of North America. It consists of 100 nodes and 171 internodal segments, also referred to as edges, links, or sometimes routes, joining them. Some IXC networks are smaller (~50) and some are larger (~200) than this example. Today, the majority of long-distance traffic in these networks is routed via nodes at major population centers using 2.5 and 10 Gb/s channels carried on dense wavelength-division multiplexing (DWDM) over single-mode optical fibers. Soon the maximum time-division multiplexed rates will increase to 40 Gb/s, transmission distances of 1–10 Mm will be possible, and the number of wavelength channels carried will approach 200.

The use of optical switches at the heart of optical layer cross-connects in nodes of the fiber transport network, Fig. 1.2, is the latest example of the

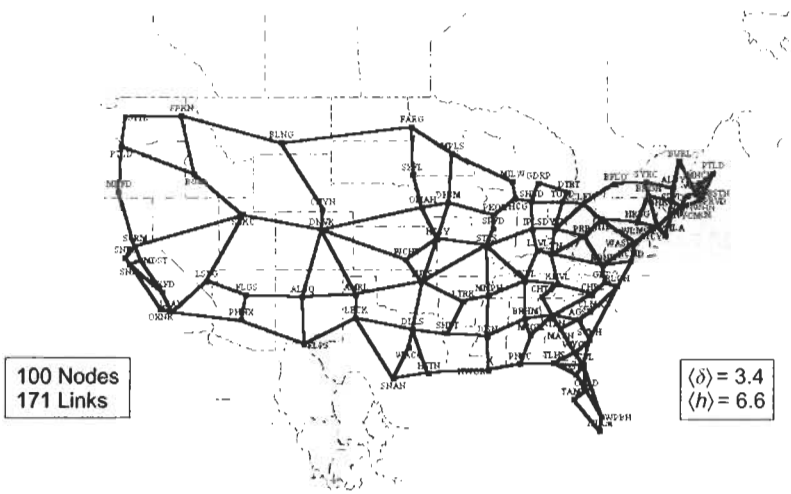


Fig. 1.2 Nodes and links of a prototypical fiber transport network.

ongoing replacement of electronics by optics in long-distance voice and data communication. With the advent of optical multiplexing/demultiplexing and optical switching, optics not only represents a path to increasing capacity but a radically new infrastructure for networking. Using optical layer cross-connects and flexible optical add/drop multiplexers to manage bandwidth in the core, it will be possible to more efficiently provision, rearrange, and restore high-bitrate channels carried on wavelengths. The first deployments of optical switches for managing bandwidth in long-haul fiber communication were 72 × 72 multimode optical cross-connects for 622 Mb/s in 1996 [Young, et al., 1998]. However, as late as 2000 the majority of the international R&D community considered single-mode OXCs with several hundreds of I/O ports unfeasible and gave them little attention. Today, a few short years after the challenge of developing large-port-count single-mode optical switch fabrics was undertaken in earnest 256 × 256 optical cross-connects for 2.5 Gb/s and higher data rates have been commercialized and deployed, and research prototypes having an astonishing 576 × 576 and 1296 × 1296 ports have been reported [Lee and Capik, 2000; Ryf, et al., 2001].

1.3 LONG-HAUL TRANSPORT, OPTICAL LAYER NETWORKING, AND OPTICAL CROSS-CONNECTS

Both traditional voice and contemporary data communications networks are structured hierarchically to provide affordable communication connectivity among the world's large and growing number of humans and machines. In much the same way that governments and transportation systems are structured into local, regional, national, and international infrastructures, so too the population of communication end users is typically divided into communities of increasing geographic scale. The reasons are fundamental. While all members of the population should have the opportunity to engage and communicate with any other member, the processing power and fan-in/fan-out capability of both biological organisms and man-made machines are limited, and the energy and cost of transporting goods and information increase with distance. By dividing the extended population into communities, dividing the communications between individuals into either intracommunity or intercommunity, and creating interconnected communities of communities, we can use the hierarchy to achieve complete communication connectivity among all the members of the population under the physical and cost constraints that are imposed.

Beginning with the end user and moving down the hierarchy, communications networks are arranged into local, access, metropolitan, regional or wide-area, and long-distance or global networks. For example, local communities of a few to several thousand persons can be served by single private branch exchanges (PBX) on building premises; groups of local communities totaling several tens of thousands of individuals are served by single

electronic switching systems located at local central offices; and collections of groups of local communities totaling several hundreds of thousands to approximately one million people are served by cross-connects at large offices. An analogous hierarchy exists in strictly data networks with Internet Protocol (IP) routers of appropriate sizes and capacities serving the edge, middle, and core of the corresponding network. The long-distance network at the base of voice and data hierarchies is referred to as the backbone, or core, and ultimately supports the long-distance communication among the end users. Because the purpose of the core network is to move digital information efficiently over large distances, the core network is also often referred to as a transport network. Similar to transportation systems, while the probability that any one individual makes a journey between an arbitrary origin and destination is small, the large number of individuals within a region ensures that there is a constant demand for transportation resources among all major hubs. One of the goals of transport networking is to size the transport capacity to the demand (average or peak), which dictates a high degree of resource sharing at the core of the network among the end users.

To be more concrete about the structure of the core network, we illustrate in Fig. 1.3 a small hypothetical hierarchical network having 216 access nodes, 36 metro nodes each serving 6 access nodes, 12 regional nodes each serving 3 metro nodes, and a backbone transport network of four nodes connected by five routes with each core node serving 3 regional nodes. If each access node were to support four local networks and each local network were to support

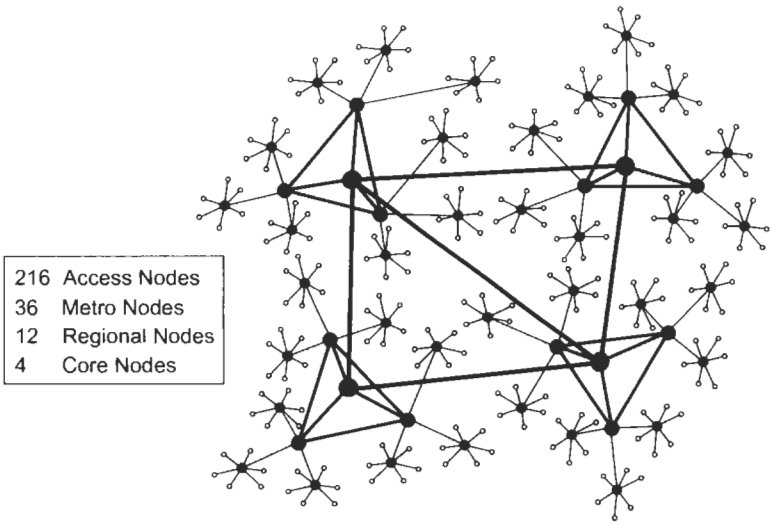


Fig. 1.3 Hierarchical network with backbone serving end users.

eight end users, then this miniature pedagogical network would provide complete connectivity among over seven thousand end users. In Fig. 1.4 we focus on a portion of the core and regional networks to show how the regional layer may be serviced by electronic switches, cross-connects, or IP routers, which themselves are supported by optical layer cross-connects in the backbone. The optical transport layer architecture may be further abstracted as shown in Fig. 1.5. There, nodes consist of an optical layer cross-connect that supports an extended region and substantial population represented by the client electronic terminal attached to the node. The optical layer cross-connects at the

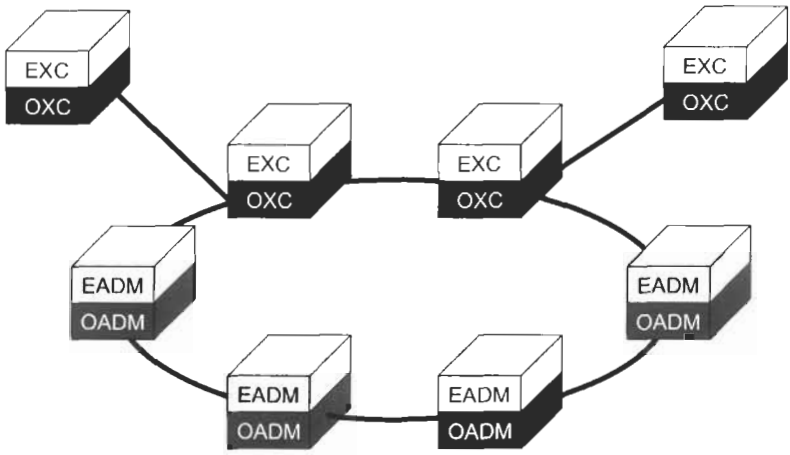


Fig. 1.4 Regional and core networks.

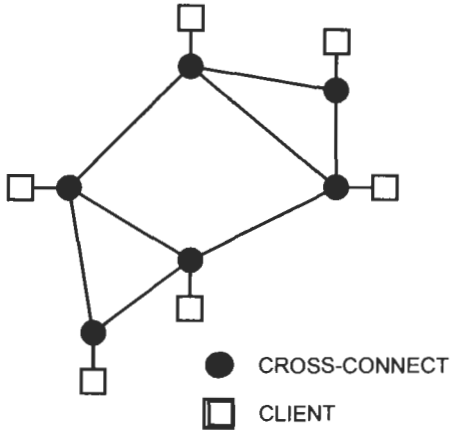


Fig. 1.5 Transport network architecture.

nodes are connected via a set of major routes, or links—in this example, eight in number.

Cross-connects are used to manage bandwidth in the core network. Consequently, to better appreciate what is required of an optical cross-connect, it is necessary to consider the basics of core networks. Referring to Figs. 1.2 and 1.5, the average number of two-way routes entering/leaving a node $\langle\delta\rangle$, also called the average degree of a node, is easily deduced by inspection and is $\langle\delta\rangle = 2L/N$, where N is the number of nodes in the network and L is the number of internodal edges, or links. The factor of two occurs because each two-way link is shared by two nodes. The number of routes linking nodes L in the core network is usually between one and one half to two times the number of nodes. This is understandable when one considers that for a ring the ratio of the number of two-way routes terminating, i.e., entering/leaving, at a node to the number of nodes is unity and for an infinite rectangular grid the ratio is two. Consequently, $\delta \sim 3.5$ is a typical value for the degree of the node of a mesh network.

The average degree of a node for the transport network illustrated in Fig. 1.2 is $\langle\delta\rangle = 3.4$. Like the real networks it mimics, this prototypical network was constructed with the condition that the failure of all connections carried over any single segment does not isolate any node from the other nodes. Consequently, the minimum degree of a core node is $\delta = 2$, as happens to be the case for Burlington, VT, Milwaukee, WI, Portland, ME, Seattle, WA, and Tampa, FL, among others in our example network. The largest degree of a node in this network is $\delta = 6$, for example the nodes at Denver, CO and St. Louis, MO.

Among the key attributes, or drivers, for cross-connects is the port count. The number of ports required on a typical cross-connect in a mesh network can be estimated by making a few basic but reasonable assumptions. These are that all the nodes of the network serve similar number of end users; that end users are equally likely to contact other users regardless of distance; that the traffic is ideally groomed when entering the backbone network; and that the unit of bandwidth currency in the core is an optical channel just large enough to accommodate node-to-node traffic. These conditions describe an optical layer network of *uniform* traffic. Note that as illustrated in Fig. 1.6, the average number of output ports is equal to the average number of inputs, as connections are usually duplex, i.e., two-way. Also, again referring to the figure, an $M \times M$ cross-connect, which has M inputs and M outputs (or M I/O ports), supports connections among M two-way links.

As the terminal, or client, at each node of a *uniform* network has an optical channel path to every terminal at each of the other nodes, the number of add (drop) ports originating (terminating) at each node is $A = N - 1$. Of course, every optical channel enters one or more additional cross-connects as it traverses the network between the terminating nodes, as indicated in Fig. 1.7. Consequently, besides the original add port, another input port is occupied on each additional cross-connect that is encountered. If we denote the number of

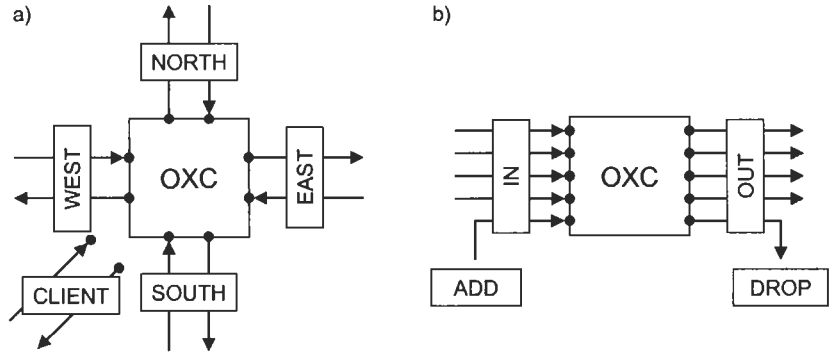


Fig. 1.6 Cross-connect views.

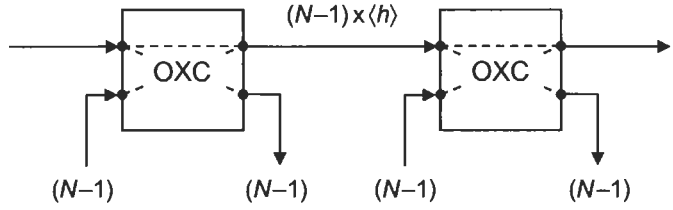


Fig. 1.7 Average number of ports on an optical cross-connect.

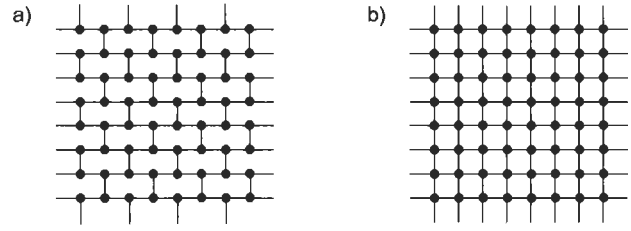


Fig. 1.8 Uniform triangular and square mesh networks.

hops averaged over all $N(N-1)/2$ node–node pairs of the network as $\langle h \rangle$, then in the absence of extra capacity for restoration the number of input (output) ports $\langle M \rangle$ required on the typical cross-connect is

$$\langle M \rangle = A + A \times \langle h \rangle = (N-1)(1 + \langle h \rangle) \tag{1.1}$$

The average number of hops for a closed-mesh network depends largely on the number of nodes in the network. To estimate $\langle h \rangle$ it suffices to consider a closed uniform square mesh consisting of N nodes, as depicted in Fig. 1.8(a). Such a network has four routes impinging on each cross-connect, except for those nodes located on the boundary. For a uniform square mesh, the average

number of hops between node pairs is $\langle h \rangle \sim \frac{2}{3}\sqrt{N}$. Planar closed-mesh networks having different average degree of node exhibit a similar dependence on the number of nodes. For example, in the case of a closed uniform triangular grid where the average degree of a node approaches 3 for large N, the average number of hops between node pairs is $\langle h \rangle \sim \frac{3}{4}\sqrt{N}$. The average number of hops for the rectangular-shaped network of 100 nodes in Fig. 1.2 is approximately $\langle h \rangle = 6.6$. Semi-empirically we find that the average number of hops of a network may be estimated to within 10% accuracy using $\langle h \rangle = 1.1/\sqrt{(N-1)/\delta}$. The diameter D of a network is defined as the minimum number of segments, or hops, traversed to connect the most widely separated nodes. For the network currently under consideration, the diameter is $D = 17$.

The above estimate for the number of ports on the average cross-connect was based solely upon the minimum number of optical paths to support all pairwise connections between terminals. Usually, extra capacity is deployed within the network for the purpose of network survivability and should be taken into account when sizing a cross-connect. The percentage of extra capacity varies depending upon the details of the network topology, the types of failures to be accounted for, and the amount of resource sharing that is implemented [Baroni, et al., 1999]. Typical values are in the range of 25 to 75% for mesh-based shared restoration of the scenario where all channels of an arbitrary single link may fail, which may occur when a conduit trench is accidentally breached. If we assume a 50% overbuild for shared restoration, the average number of input number ports occupied on a cross-connect is

$$\langle M \rangle = (N-1)\left(1 + \frac{3}{2}\langle h \rangle\right) = (N-1)(1+\sqrt{N}) \tag{1.2}$$

We observe that when N is large the average number of ports is approximately

$$\langle M \rangle \sim N^{3/2} \tag{1.3}$$

Thus, if the number of nodes in the network is approximately 30, then the average number of ports required is about 200. When N is about 100, then $\langle M \rangle \sim 1000$; and when N is about 200, then $\langle M \rangle \sim 3000$. Similarly, the average one-way traffic cross-section carried on the route between adjacent nodes counted in multiples of the basic internodal communication channel (bandwidth unit-of-currency) is $\sim N^{3/2}/\delta$ when N is large compared to unity. These estimates correspond to a typical node and link, and therefore the largest nodes and the links with the heaviest traffic may easily require more than twice the average values.

The above analysis can also be used to estimate the ratio of the amount of the traffic that is added and dropped (terminated) at a node, *add + drop*, to the amount of traffic that passes directly through a node without terminating, *thru*. For the uniform square mesh with 50% extra capacity for restoration, the

ratio of dropped to through traffic is approximately

$$(add + drop)/thru = 2/\sqrt{N} \tag{1.4}$$

Thus, for a mesh network of 25 nodes the ratio of add/drop to through channels is approximately 40% on average, and the percentage decreases as the number of nodes increases. Of course, this estimate is for the average node, and the percentage can be larger or smaller depending upon the details of the network demand and topology.

While not explicitly stated, so far we have focused on networks that are mesh-like in structure. However, optical switches also find application in flexible optical add/drop multiplexers of linear and ring networks. Whereas closed mesh networks in a plane are intrinsically two-dimensional, linear and ring networks are essentially one-dimensional. This difference in the physical topology has consequences for the requirements placed on the switches for a given network size. For example, the average number of hops in a unidirectional ring network is $\langle h \rangle = N/2$. Therefore, based on Eq. 1.1 the number of effective input/output ports required of a switch within a fully flexible add/drop multiplexer on a unidirectional ring network with full logical connectivity among all the terminals is

$$\langle M \rangle = (N - 1)(1 + N/2) \tag{1.5}$$

As an example, a basic architecture of an optical add/drop multiplexer for a unidirectional ring with 16 nodes requires an optical switch unit having $\langle M \rangle = 135$ input/output ports. Note, unlike a cross-connect in a mesh network, the switch fabric within a ring switch is generally not strictly nonblocking among all the I/O ports.

1.4 EVOLUTION OF OPTICS IN DIGITAL TRANSPORT—OPTICAL CROSS-CONNECTS

The characteristic of transport networks, reflected in Eq. 1.4, that a very large fraction of the data directed into a node is destined for other nodes, is a primary motivation for introducing optics into the set of tools for bandwidth management in the core. As illustrated in Fig. 1.9, by using a combination of electronic and optical bandwidth management at nodes of the backbone, only those channels that terminate or need to be electronically regroomed at a node would be directed by the optical cross-connect to the electronic bandwidth manager, which may take the form of an electronic cross-connect or an IP packet router.

Clearly, if the cost of redirecting the concatenated bandwidth of an optical layer channel is less using an optical switch than an electronic switch, the optical tools will be introduced into the architecture. Not surprising, the relative costs of the optical and electronic technologies depend upon the bitrate of the individual channels and their number. Influencing the costs are limitations

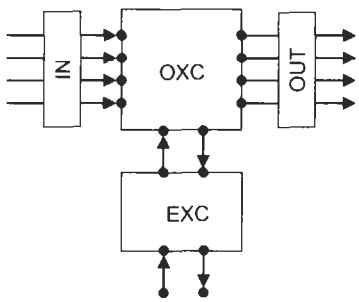


Fig. 1.9 Complementary optical and electronic bandwidth management.

imposed by the power requirements and physical sizes of the respective solutions, which include the switching fabrics, drivers, and backplanes, as well as external interfaces and controllers. Currently, optics is considered competitive when the unit of currency of bandwidth is in the neighborhood of 500 Mb/s–2.5 Gb/s.

The cost advantage and relative technology life of optics, which intrinsically has extraordinarily broad bandwidth, increases quickly with bitrate. This is because, as the bitrate of the incoming channel is increased, the effective number of ports of an electronic switch decreases as a consequence of the fixed operating bitrate of the fundamental switching elements, or cross-points. For example, if we consider a 256 × 256 electronic fabric working at a maximum of 2.5 Gb/s, then as the bitrate of the incoming channel is increased to 10 Gb/s, the effective number of ports decreases to 64 × 64. In this case, four 2.5 Gb/s ports are occupied and processed in parallel for each incoming 10 Gb/s channel. Similarly, if the bitrate of the incoming channel is raised to 40 Gb/s, the effective number of ports is further reduced to 16 × 16. In contrast, the effective number of ports of an optical switch fabric, such as those based on free-space beam-steering, can remain constant as the bitrate is increased beyond several hundred gigabits/second.

Another benefit of using the complementary strengths of two technologies to architect an optimized solution for bandwidth management, rather than accepting the tradeoffs imposed by any single technology, is that larger networks can be supported. For example, any generation of electronic switch or IP packet routers, which by definition process all bits and packets regardless of whether they are to be dropped or passed through, are bound by a total capacity measured in bits per second. Consequently, a given electronic technology used for bandwidth management in the backbone can support a network of some maximum number of nodes N_e. By introducing an optical cross-connect in combination with the electronic switch at each node, the optical switches can offload through traffic from the capacity-limited electronic switches. In this way, based on the assumptions leading to Eq. 1.4, networks of roughly a factor of $\sqrt{N_e}$ times larger can be realized with existing technology. Today's IP

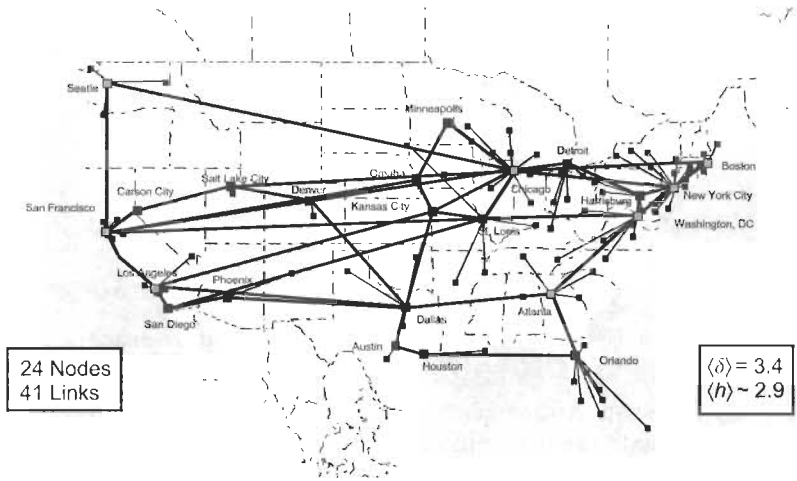


Fig. 1.10 Architecture of AT&T's IP Backbone. Adapted from Pease, 2001.

backbones consist of approximately 25 core nodes, as is illustrated in Fig. 1.10. Thus, the introduction of OXCs into the IP core would permit the network to increase in the number of nodes that could be supported to more than 100. According to Eq. 1.3, such OXCs would need approximately 1000 I/O ports. Recently, optical switch fabrics capable of such very large numbers of I/O ports have been reported, and soon optical layer cross-connects based on these fabrics are expected to become generally available and to play a significant role in the expansion of IP data networks [Ryf, et al., 2001].

Optical switch fabrics will complement electronic switch fabrics in applications such as optical layer cross-connects where and when optical switches are a lower-cost solution for routing and protecting express traffic. As we have discussed, because optical switch fabrics are virtually transparent to bitrate, this new role for optics in transport networking is inevitable and has already begun. However, the introduction of optical switch fabrics into cross-connects has further significant implications for the future capabilities and cost reduction of transport networks. This is because the introduction of optical fabrics in the transport layer will begin to bring long-haul fiberoptic communication another step closer to transparency to signal bitrate, format, and protocol. Such independence will one day lead to lower costs for reliably transporting data and to greater flexibility and simplicity for network operators and their customers to manage their networks and changing communication needs efficiently and quickly. For example, data that could be transported great distances from originating and terminating nodes without interruption would not unnecessarily encounter electronic processing. Also, in principle, multiple formats such as SONET and native ATM and IP could all be carried within a single network infrastructure, and terminal bitrates could evolve with fewer constraints.

However, it should be remembered that although optical switch fabric technology is a critical enabling vehicle to transparent networking, it is not a solution to all of the networking and transmission needs and issues that are encountered. First, the multiplexing and grooming of lower bitrate traffic into wavelength channels at the edge of the network is best accomplished by electronic means. Also, long-distance optical fiber communication systems are largely analog transmission systems, and consequently the quality of the transported signals (waveform fidelity, signal-to-noise ratio, and synchronization) degrade, or age, as they traverse the network. Presently, optoelectronic regenerators are used to periodically recreate and simultaneously rejuvenate data channels and thereby ensure all channels reach their final destination in good health. These regenerators, which briefly convert the data signals into the electronic domain to remove accumulated impairments, also provide an excellent and low-cost means of detailed signal performance monitoring and wavelength conversion—essential functions for the efficient management of the network. Thus, in the near term, optical switch fabrics will be used in conjunction with optoelectronic regenerators to achieve the increased levels of performance and cost savings that are sought. In the future, all optical methods for regeneration, wavelength conversion and interchange, and performance monitoring, which are actively being investigated, will be devised, improved, and refined in the challenge to supersede the capabilities and efficiency of their electronic predecessors and fulfill the vision of an all-optical transport network.

1.5 TRENDS IN CAPACITY REQUIREMENTS FOR TRANSPORT AND BANDWIDTH MANAGEMENT

Expanding voice service and the introduction of new data services like the Internet, broadband, video, etc. have increased demand for network capacity. That demand and the technology it has stimulated have driven recent transmission system capacities to double as the cost of capacity decreases by about fifty percent (Fig. 1.11). That same increased capacity and cost reduction trend has in turn further fueled demand in the optical networking market.

As a consequence of the increase in the demand for services, higher bandwidth of services, and longer distances of communication, has come the need to manage bandwidth in the core at ever higher bitrates. This is illustrated in Fig. 1.12, where we trace the evolution of cross-connect systems using Lucent Technologies' cross-connect products as indicative of the industry trends. The first cross-connect in the hierarchy shown Digital Access and Cross-Connect System II (DACS II) was introduced in 1981 and supports DS-1 (1.5 Mb/s) line rates and cross-connects traffic at DS-0 (64 kb/s) rate.

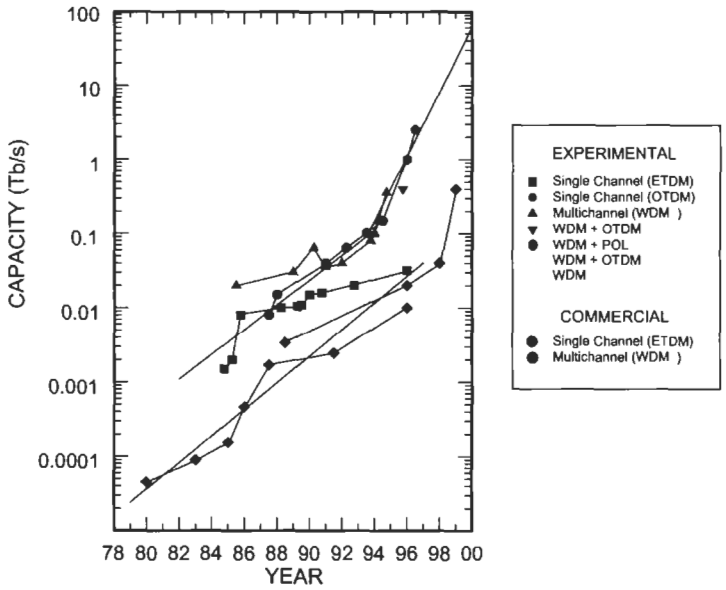


Fig. 1.11 Trends in lightwave transmission capacity. Adapted from Chraplyvy and Tkach, 1998.

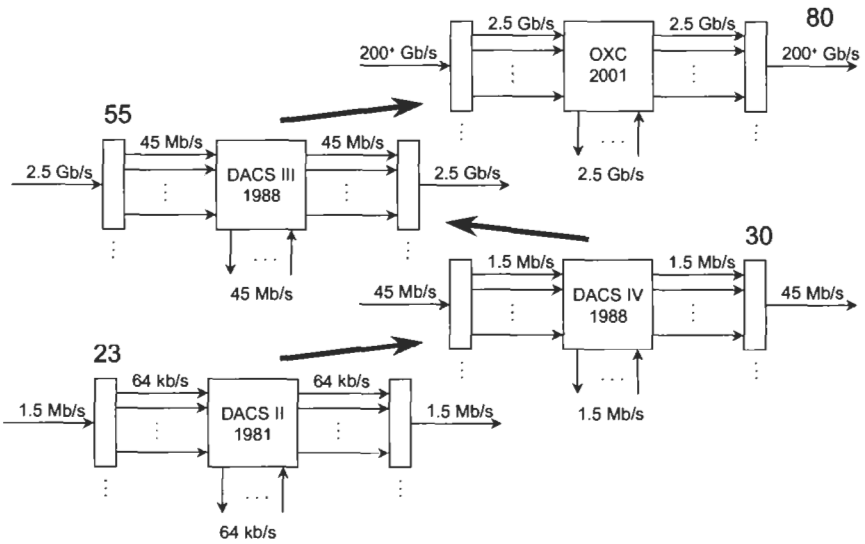


Fig. 1.12 Evolution of optics in digital transport: Networking-XCs.

The second, DACS IV-2000 was introduced in 1988 and is a wideband cross-connect that supports 512 DS-3 (45 Mb/s) line rate ports and cross-connects signals at the DS-1 (1.5 Mb/s) rate. The third, DACS III-2000—also introduced in 1988, is a broadband cross-connect that supports various line rates including OC-48 (2.5 Gb/s) and cross-connects signals as many as 2048 channels at the DS-3 (45 Mb/s) rate. Additionally, the WaveStar™ Bandwidth Manager (not shown), which was introduced in 1998, supports line rates up to OC-192 (~10 Gb/s) and can cross-connect as many as 4608 STS-1 (51.84 Mb/s) signals [Alegria, et al., 1999].

In examining the ratios of the typical line rates to the cross-connect rates from DACS II through DACS III-2000, we note that they remain within about a factor of two (23 to 55) even as the line rates and cross-connect rates have increased by roughly three orders of magnitude (1.5 Mb/s to 2.5 Gb/s) as the two new generations of systems were introduced. Arguably, this is because the numbers of nodes in the core, and consequently the number of ports on the cross-connects, has grown relatively slowly as the average long-distance bandwidth used by population centers has substantially increased. In the last decade and generation of transmission systems, the carrying capacity of optical fibers has been driven to increase by a factor of approximately 100. Based on past trends, we would therefore predict that the unit of bandwidth currency managed in the core by the current generation of cross-connects should now be about 2.5 Gb/s. Indeed, IXC network operators are actively seeking optical layer cross-connects with this feature, and equipment suppliers have recently introduced products with this capability.

1.6 TOPICS AND OUTLINE OF THIS CHAPTER

In this chapter we focus on optical switches for optical cross-connect applications, the achievement of which is among the most significant milestones in the evolution of optical fiber communications since the publication of the last volumes in this series.

While the topic of optical switches for optical add/drop multiplexers is a topic of considerable interest, coverage of this important application is not treated here, but may be found in other reviews [Giles and Spector, 1999]. Further information and discussion about optical switch technologies may be found in the previous edition of this series [Murphy, 1997], and elsewhere in this volume.

2. Features, Architectures and Applications

2.1 CROSS-CONNECT FEATURES

In this section, we describe some of the primary features of an OXC that we must consider for a given application. OXC size, rapid and strictly nonblocking

reconfiguration, bit rate and format transparency, capability to operate in a multivendor environment, and OAM&P features will play a major role in implementing and developing OXC architectures. The various architectures that support these features are described in more detail in Section 2.2.

As mentioned in the previous section, the size of an OXC can increase in size to hundreds of input and output ports depending upon the network bandwidth requirements and line system capacity (refer to Section 1.3 for a more detailed discussion). The provisioning and control of the OXC—basically a wavelength map of the entire optical network, needs to be constructed and maintained. In addition, nodes containing OXCs need to distribute the states of their local OXCs to other nodes so that applications such as mesh restoration and bandwidth-on-demand can be realized. Accordingly, a whole suite of communications protocols is being discussed in the standards bodies toward this end (e.g., MPLS, GMPLS, RSVP, etc.).

Bit rate and protocol transparency is a critical feature of an OXC. Bit rate and format transparency means that different bit rates and formats, such as analog signals and digital formats, can be carried on the transmission interface transparently. This feature is illustrated in Fig. 2.4. This type of OXC is capable of switching traffic as wavelengths, bands of wavelengths, or entire fiber. The bit rate and format transparency will provide network operators with a flexibility to provision and restore wavelengths in their networks quickly by not having to worry about unnecessary port blocking due to any specific properties of the OXC itself. As the fiber bandwidth capacity increases with traffic demand, the need for service providers to efficiently manage their networks in this manner will increase as well. This is especially true in networks consisting of equipment from more than one vendor.

Multivendor operation is another feature that may be used to distinguish OXC architectures. An opaque OXC architecture is tailored to multivendor operating environments because it ensures there is no connection blocking between the different networks and possibly eases requirements on the physical properties of the OXC itself. In this architecture the transmitter and receiver interfaces adjacent to the OXC from multiple line systems are designed to operate at only one wavelength. Examples of a multivendor OXC architecture are shown in Figs. 2.1 to 2.3. In the configuration described here, the OXC

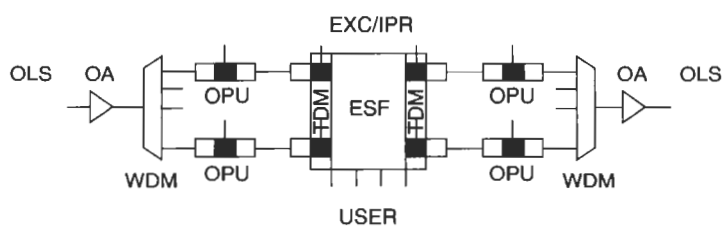


Fig. 2.1 Opaque OPU: Rx + BP + Tx providing 3R/PM/WC.

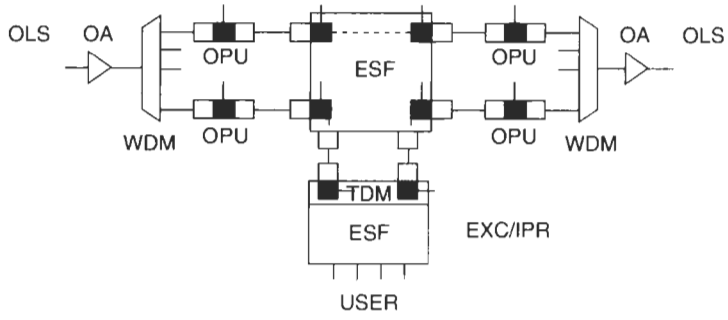


Fig. 2.2 Opaque OPU: Rx + BP + Tx providing 3R/PM/WC.

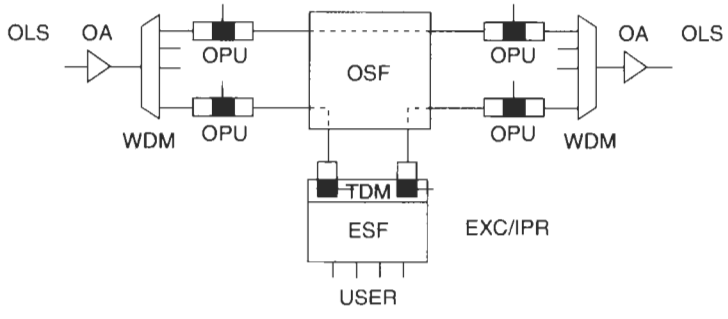


Fig. 2.3 Opaque OPU: Rx + BP + Tx providing 3R/PM/WC.

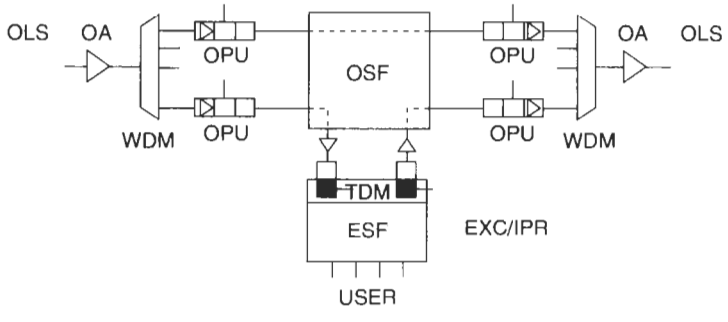


Fig. 2.4 Transparent OPU: OA + OPM + WC providing 3R/PM/WC.

receives signals using standard 1.3 μm receivers and transmits signals using standard 1.3 μm transmitters. An OXC that uses these standard interfaces could terminate signals from different vendors as shown with the only design requirement being that the OXC operate at 1.3 μm.

In addition to the switch fabric, another important component of an OXC is adequate OAM&P features and to have redundant switch fabric to maintain high reliability. Since OXCs handle large volumes of traffic, it will be valuable

to eliminate the switch as a single point of failure in the event of an internal fault in the switch. In addition, the OXC should be able to locate and detect internal faults at its input and output ports, which can be achieved by OAM&P capability. In the following section, we describe an OXC with opaque interfaces, which contain OEO conversion to provide bit rate monitoring functions and support some of the OAM&P capability.

2.2 CROSS-CONNECT ARCHITECTURES

At the heart of digital cross-connects (DCS) are large matrix switches used to aggregate, groom, and redirect traffic on telecom networks. Traditional DCS-type products usually handle SONET voice traffic as well as time-division multiplexed (TDM) leased lines. However, broadband cross-connect products are evolving and data networking technologies such as ATM and IP are gradually being added. From a switching perspective, these cross-connects are the next generation of current broadband digital cross-connects, also known as electronic cross connects (EXCs) and optical cross-connects (OXCs). To show the evolution of cross-connect systems, we illustrate an office view and explore various cross-connect architectures. The optimum architecture will be dependent on the application and network needs.

Figures 2.1–2.4 illustrate a cross-connect that can be realized using electrical or optical switching fabrics. As shown in Figs. 2.1, 2.2 and 2.3, a cross-connect is positioned at the intersection of optical line systems. Signals from optical line systems (OLSs) are demultiplexed, and the constituent wavelengths are passed through transmission interfaces that are opaque containing OEO conversion providing 3R regeneration, bit rate monitoring, and wavelength conversion functions at the node ingress and egress. The single wavelength signals are then cross-connected by the cross-connect, passed through OEO conversion again, and multiplexed before they enter the outgoing optical line systems. Architectures in which signals undergo OEO conversion are often referred to as opaque architectures.

Since opaque interfaces contain OEO conversion, a wavelength conversion is an automatically available feature. This provides the correct set of wavelengths at the output of the cross-connect for traffic that, in turn, has to be multiplexed onto the same output fiber. In addition, it helps simplify the restoration process because it will not block the signal in the wavelength dimension. In addition, the optoelectronic converters regenerate the transmission signal by retiming, reshaping, and thus ensuring the integrity of the signal when it exits the node. The cross-connect with opaque interfaces can provide performance monitoring functions with the information obtained from optical-to-electrical conversion. This will support SONET/SDH performance monitoring functions such as loss of signal (LOS), loss of frame (LOF), alarm indication signal (AIS), and bit error rate (BER). Such measurements are critical for fault detection and fault correlation, and are usually required for carriers to adhere to any existing service level agreements (SLAs).

However, OEO conversion entails high equipment cost and, typically, the cost of the interfaces dominates the cross-connect cost. In addition, since these converters retime and reshape the signal, they are bit rate specific. Thus, an increase in the bit rate will require an upgrade in the converter. Also, the lack of bit rate and protocol independence capability leads to an inability to interconnect signals between two fibers with different bit rate and protocol—ATM, IP, Ethernet, SONET, SDH, etc.

To avoid blocking within the cross-connect, the opaque interfaces need to be bit rate and protocol adaptive, meaning that they can automatically provision themselves to support multiple rates and protocols. Such interfaces are only now becoming available.

2.2.1 Cross-Connects with Electrical Switching Fabric

As shown in Figs. 2.1 and 2.2, in electrical fabric based systems carriers must turn light signals into electrons and then back to photons to switch wavelengths from one fiber to another. This requires optical–electrical conversion when the signal enters the cross-connect and electrical–optical conversion when the signal is routed out of the cross-connect. Compared to that as shown in Fig. 2.3, in optical fabric based systems the carriers can route wavelengths optically from one fiber to another. This removes the need for converting optical signals to electrical signals and then back to optical signals at intermediate nodes.

The EXC shown in Fig. 2.1 may contain an electrical switching fabric operating at, say, STS-1 (50 Mb/s) granularity. It converts the optical signal into an electrical signal and demultiplexes it down to STS-1 levels before switching it internally. At the intermediate nodes in the network, some wavelengths may need to be demultiplexed to the STS-1 layer to extract individual STS-1 streams. Therefore, this network element is clearly needed here to groom traffic at the STS-1 layer. However, according to industry studies, 70 percent or more of the traffic at any node is through traffic, which can be handled more efficiently with optical switch systems. As shown in Fig. 2.3, the optical switch fabric reduces the number of OEO conversions as a signal traverses the network, and has the potential to reduce system costs.

Typically, if the cross-connect cannot cross-connect or switch at the higher rate, then more ports are required on the cross-connect to handle the increased fiber throughput. An electrical cross-connect system will be size limited as data rates rise to 10 Gb/s. For example, a 40 Gb/s signal must be broken down to 16 individual 2.5 Gb/s streams to connect to the electrical switching fabric operating at 2.5 Gb/s. This capacity will absorb 16 ports in an electrical switch, which will degrade the switch fabric capacity as the bit rate will rise at the interfaces. For example, a 512-port cross-connect system with 2.5-Gb/s traffic becomes a 128-port system (512/4) with 10-Gb/s traffic and a 32-port system (128/4) with 40-Gb/s traffic. In general, in the electrical fabric, a higher bit rate signal needs to be broken down to the rate at which the electrical

switch fabric operates. This leads to greater complexity in the EXC, resulting in systems of larger footprint and great power consumption.

As a next step of cross-connect systems evolution, one other option that is currently available is the cross-connect with electrical fabrics and optical interfaces (see Fig. 2.2). Typically, this type of cross-connect has an electrical switching fabric operating at 2.5 Gb/s and switches and demultiplexes traffic in units of 2.5 Gb/s. In terms of general network applications, these cross-connects are still limited because they can switch OC-48 streams but without sub-rate grooming. In other words, they provide a functionality between cross-connects that can groom at the STS-1 rate and those that can groom at the wavelength level.

2.2.2 Cross-Connects with Optical Switching Fabric

In the future optical networks, as shown in Fig. 2.4. cross-connects may have an optical switching fabric and transmission interfaces that are transparent from the standpoint of bit rate and signal format and contain no OEO conversion. Bit rate independence means bit rates from 45 Mbps to 40 Gbps can be carried on an optical interface. Format independence means support of formats such as analog signals along with digital formats such as SONET, ATM, IP, and FDDI. However, an optical interface must meet certain constraints such as the optical wavelength of the fabric, the optical loss through the OXC system, and the optical requirements of the interfaces.

In this scenario, wavelengths are routed optically reducing the number of OEO conversions a signal experiences as it traverses a network. As carriers begin deploying more wavelengths in their networks, the need to manage capacity at the wavelength level (that is, at the optical layer) will increase. Managing wavelengths at the optical level should prove to be less costly than converting them to the electrical level. In addition, it will increase reliability and decrease operations costs, because it requires fewer network elements.

In the network core, a significant amount of traffic at any node can be treated as express traffic, which can be best handled by the network element that manages traffic in wavelengths. In the future, as traffic increases, there will be an increasing need to manage traffic in groups of wavelengths (bands) or even of entire system fibers. A cross-connect with an optical switching fabric is capable of switching traffic in all these ways.

However, because an OXC switches at the wavelength level without sub-rate grooming, it is most likely that they may be implemented in the hybrid office architecture scenario as shown in Figs. 2.2–2.4. Here an EXC is used for traffic termination and electronic re-grooming, e.g., at STS-1 level granularity, and an OXC is used for pass through traffic.

Typically, since the system cost of the use of o-e-o converters is much more than the cost of the OXC element itself, this architecture offers huge cost savings for service providers. However, this lack of o-e-o converters surrounding

the OXC prevents this solution from performing the essential performance monitoring function that is required. Presently, the o-e-o converters provide this. In the transparent OXC architecture, the PM function must be done elsewhere in the network or must be provided off-line at the OXC inputs until the technology and standards work address these challenges that are required to support transparent payloads. It is most likely that future networks will consist of a hybrid OXC architecture where client-side interfaces can be opaque and network-side interfaces can be transparent. This approach would have the full benefits of the optical switching fabric and the best of both opaque and transparent interfaces to achieve intelligent optical networking. Such a network is shown in Figs. 2.3 and 2.4. Note that the network-side interfaces can remain transparent as long as the loss budget can be met for the optical channel.

Another advantage of an OXC is providing a high capacity in a small footprint. Since the optical switch fabric operates at any optical layer bit rate, including 40 Gb/s and beyond, the switch fabric capacity will not degrade as the bit rates at the interfaces increases. For example, an OC-768 interface will take only a single port for switching in an all-optical switch fabric and the whole 40-Gb/s signal will be kept intact at the wavelength level. In addition, in an OXC system, the signals are routed optically from one fiber to another, which will reduce the number of electrical components needed in the system, and typically reduce the power consumption by an order of magnitude compared to an electronic cross-connect system.

2.3 OPTICAL CROSS-CONNECT APPLICATIONS

OXCs perform optical channel bandwidth management through grooming, automatic provisioning, and easy-to-administer network survivability. In addition, they will support network reconfiguration for new service setup and changes in traffic patterns.

2.3.1 Grooming

A basic form of traffic grooming can be accomplished by routing traffic to facilities external to the OXC that supply multiplexing and demultiplexing functions. These facilities may be SONET/SDH terminals, IP routers, or ATM switches. They can be connected locally at client interfaces or remotely through the optical line system (OLS). The functionality and interfaces are illustrated in Fig. 2.4.

2.3.2 Provisioning

OXCs allow fast connection setup initiated by manual user requests or by client devices connected to the network. The client devices support bandwidth-on-demand services such as optical dialtone, dynamic trunking, bandwidth brokering services, and optical virtual private networks VPNs.

2.3.3 Survivability

2.3.3.1 Protection

Today's fiber optic networks are being scaled into the Tbps regime and are responsible for transporting service traffic on a global scale. Accordingly, network outages can affect customers on a global scale as well. The sobering fact is that bad things happen to the optical network. Failures such as equipment failures or instabilities (e.g., temperature or vibration) and human mishaps (fiber cuts, excessive fiber bends, etc.), can and do occur with devastating consequences. Previously published data [Neusy, 1994] suggest that the typical FIT rates (defined as the number of failures in 10^9 operating hours) for transmitter and receivers is approximately 10,876 and 4300, respectively. With the large number of transmitters and receivers in long-haul networks, this means that all or part of the optical network is down several minutes a year due to equipment failures. In addition, with the mean time to repair (MTTR) [Neusy, 1994] typically 2 hours, the impact due to down time becomes even more severe.

As a result of the dire consequences of network outages, ways to protect service data have been developed that allow for minimal impact on the users of the optical network as a result of the failure. To date the time response (or protection time) criteria have been set by the SONET world (GR-1223), which allots 50 ms to restore data communication following failure detection. Note that SONET-based ADMs obtain protection in the electrical domain; however, it is believed that protection of the physical (optical) layer is more suited to an optical means due to its inherent bitrate transparency and protocol independence. Despite this, however, this 50 ms criterion has remained the target. It should be noted that the adherence to this criterion depends upon the type of protection involved and the geometry of the optical network.

- ***1 + 1 Protection:*** Perhaps the most common form of optical protection is obtained by transmitting a duplicate copy of the service traffic in a path diverse manner. Both copies converge at the location where the signal is dropped, and local processors determine which of the two copies of data is of higher quality. This technique is commonly referred to as 1 + I protection; an example of such protection during states of quiescence and protection is shown in Figs. 2.5 and 2.6, respectively. As a result of the autonomous nature of the decision event the protection times can be very short—typically only a few milliseconds. The act of protection itself is performed via an optical selection switch at the drop site.
- ***Span Protection:*** Other types of protection exist that offer selective protection from a smaller set of network failures. Protection against equipment failures in the normal service path is known as span protection and is shown in Fig. 2.7. This form of protection is ideal against simpler transmitter/receiver failures, amplifier failures, or single

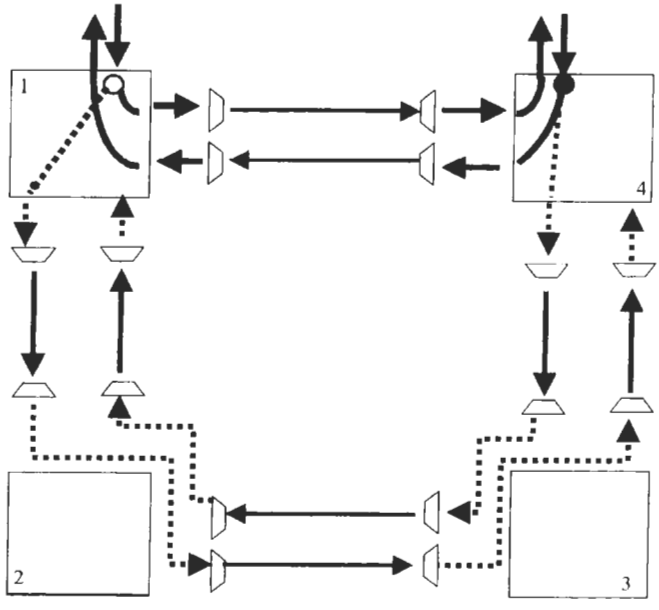


Fig. 2.5 Conventional 1 + 1 configuration in an optical network during quiescence. The solid lines are service channels and the dotted lines are protection channels. Note nodes 1 and 4 are dropping the service channel.

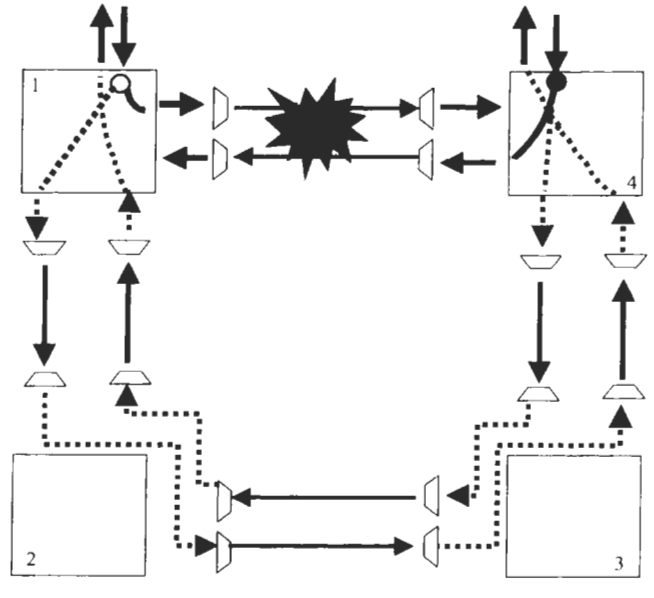


Fig. 2.6 Conventional 1 + 1 configuration in an optical network during a state of protection. The solid lines are service channels and the dotted lines are protection channels. Note nodes 1 and 4 are dropping the protection channel.

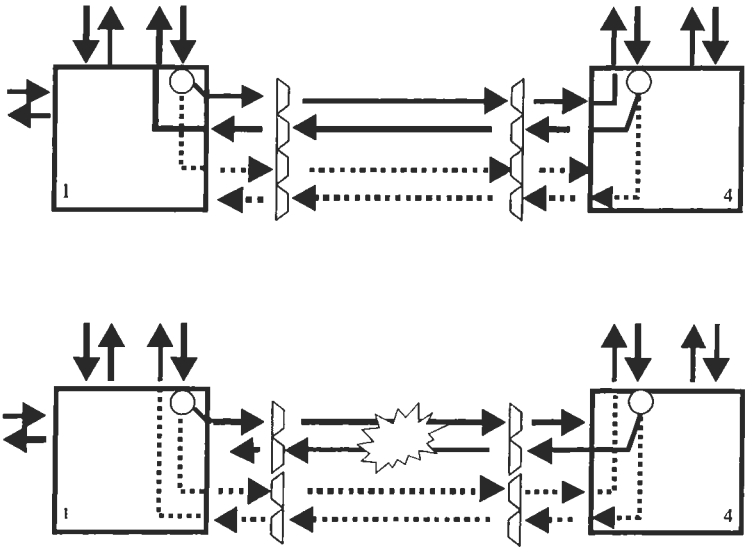


Fig. 2.7 Span protection in a 4-fiber network. Solid and dashed lines are the service and protection signals, respectively. Quiescence (top) and protection (bottom). Note the use of keep-alive splitters used during protection that allow for efficient rollback.

fiber cuts, but it is not sufficient to protect against multiple fiber cuts (e.g., a single sheath cut) that would require recovery via a diverse path.

- ***1 : 1 Protection:*** Both 1 + 1 and span protection require a protection channel, or reserve bandwidth, that is available at the instant of a network failure. The ability to access this unused, sometimes idle, protection channel on a preemptive basis is know as 1 : 1 protection. It can allow service providers to recover some of the operating costs associated with maintaining this channel. Such protection access is shown in Fig. 2.8.
- ***Shared Protection:*** Schemes such as 1 + 1 and span protection that require dedicated protection bandwidth for each service channel are inherently inefficient in the way they utilize the fiber spectrum, reducing the total service traffic capacity by 50%. Shared protection, where a single protection channel is used to provide protection over several service channel connections, has been shown to be much more efficient, cf. Fig. 2.8. In case of a failure anywhere in the network, the terminal nodes that are affected are allowed immediate access to the reserved protection channel. Those nodes not affected by the failure would be instructed to pass through the protection channel unaffected.

Even though the multiwavelength service channels travel together in the same physical layer, the nature of these independent wavelength

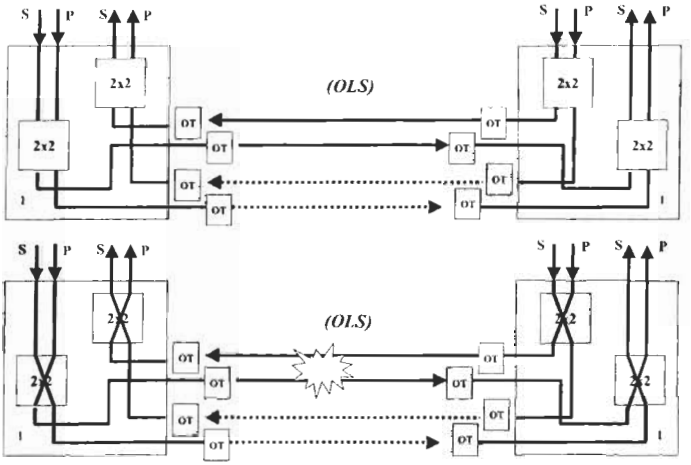


Fig. 2.8 Span protection with preemptive traffic in a 4-fiber network. The labels "S" and "P" indicate the service and preemptive signals, respectively. Quiescence (top) and protection (bottom).

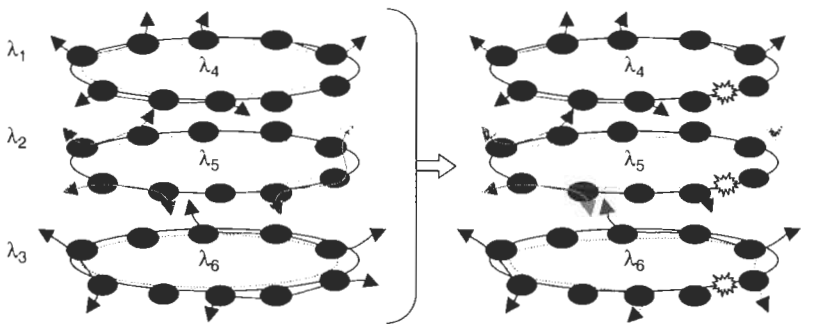


Fig. 2.9 Left: A DWDM fiber optic network utilizing shared optical protection. Each node contains MUX/DEMUX equipment that supports service channel communication over a WDM channel. Each WDM service channel (λ_1, λ_2, or λ_3) is reused around the ring and has a dedicated WDM protection channel associated with it (λ_4, λ_5, or λ_6). Right: Reconfiguration of the ring after the protection event occurs. Note that channels λ_2 and λ_3 are re-routed to the corresponding protection channel, whereas λ_1 is not affected by the failure. See also Plate 3.

shared protection rings allows for the virtual stacking of protected networks. This concept is shown in Fig. 2.9.

Note in the figure that each optical channel, shown as a unique color, represents a bidirectional connection between nodes on the ring network and that each optical channel (or wavelength) is reused within the ring so that one or more pairs of nodes can be connected (i.e., there is wavelength reuse).

A failure, or protection, event such as an amplifier failure, transmitter failure, or fiber cut can simultaneously affect one or more of the virtual rings shown in Fig. 2.9. The X shown in Fig. 2.9 (right) represents an event occurring on the network that could possibly affect all of the nodes. Each node affected by the event (and even those not directly affected) needs some type of reconfigurability in order to maintain operation; this functionality is provided by a processor-controlled optical switch fabric. After the reconfiguration occurs, all communication is restored via the protection channel. Note that all configuration changes at the nodes must be simultaneously coordinated among all nodes. Thus, in the example shown in Fig. 2.9, not only must the add and drop sites reroute their service channel to the protection channel, but other nodes must reconfigure to pass through the protection channel unhindered.

Shared protection ring networks have been built and tested [Chen, 2000; Levy, 2001]. Sample measurements of system behavior during a failure event in a 5-node network are shown in Figs. 2.10 and 2.11 [Levy, 2000]. A fiber cut is emulated by the application of a voltage impulse to an optomechanical switch placed between two nodes. A few milliseconds after the optical signal is lost, a terminal node affected by the failure communicates the event information to node controllers around the ring, and approximately 10 ms later all of the SOSFs located on the ring have reconfigured and the service channel communication is restored.

Postprotection activities by service providers who are responsible for the network operation include rollback to the service channel after the failure cause has been repaired. Such rollback must also be executed quickly in order to avoid inducing an additional outage for network

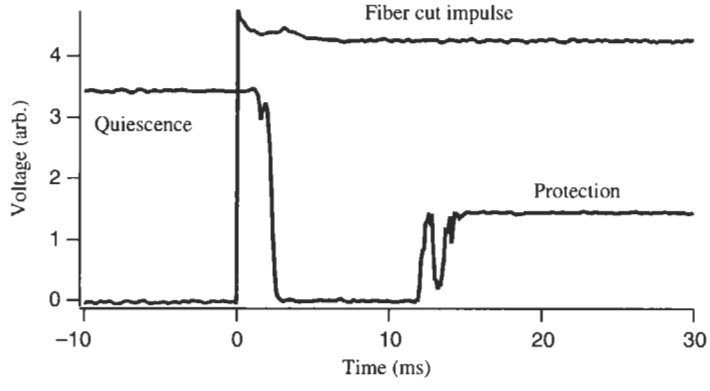


Fig. 2.10 The fast protection of service data to one of the two terminal nodes affected by a failure.

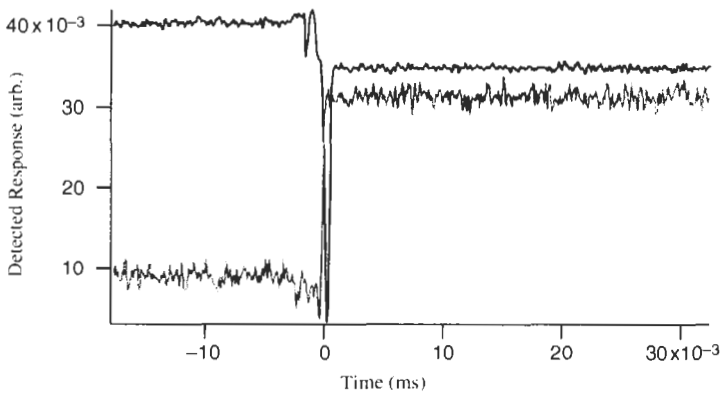


Fig. 2.11 Rollback to the service channel (pending failure source repair). Both lines represent each of the two terminal nodes initially in a state of protection.

customers. A typical measured rollback result [Levy, 2000] is shown in Fig. 2.11.

2.3.3.2 Restoration

An important OXC function is restoration. In the event of catastrophic failure, an OXC will be able to restore the network much more rapidly than it can today, since the switching will take place directly at the wavelength level.

Centralized and distributed algorithms are the two methods that exist today for restoration. In the centralized scheme, OXCs report failures to a centralized entity such as an operating system (OS). In this approach, all information is sent by the centralized entity via the maintenance channel, which makes this approach rather slow. The distributed algorithm is in general faster than the centralized approach. However, in the distributed algorithm, each OXC must have enough knowledge to perform the restoration alone. Thus, the degree of intelligence residing at each network node varies by restoration approach.

- ***Mesh Network:*** While ring topologies are useful for smaller implementations, mesh topologies are indispensable for the national, continental, and global long-haul networks that are of prime commercial interest today. In a ring network, there is little choice of backup path when a link fails, and there are few alternatives for signaling. In a mesh network, however, the high degree of connectivity is conducive to many possible traffic paths and signaling routes for the provisioning and restoration of a given demand. The more efficient use of bandwidth and the greater robustness in the case of multiple faults offset the greater complexity of the necessary route computation algorithms, signaling algorithms, and other issues.

- ***Network Node and Cross-Connects:*** Each node in the network must have the ability to receive and interpret control messages from another node and reconfigure its cross-connect accordingly. In the case of distributed restoration, each node must have the additional capability to send control messages to other nodes, while in the case of centralized restoration, it must be able to report faults to the central operating system. In general, each node must also have the ability to detect faults, though this ability can be restricted to certain nodes if a path-based restoration approach is used.

  The degree of intelligence residing at each network node varies with the restoration approach. For centralized restoration, each node is little more than a cross-connect controller, connecting input ports to output ports on a dumb patch panel pursuant to commands from a remote operating system. For distributed restoration, however, each node must possess a high degree of intelligence, capable of computing primary and backup paths and initiating the signaling sequence required to implement a lightpath. An intermediate scenario, in which selected nodes in the network can coordinate provisioning and restoration while others cannot, might be useful in reducing race conditions.
- ***Failure Scenario and Recovery:*** The two fundamental failure scenarios involve the failure of a node and the failure of a link. Each type of failure may disrupt any number of demands, depending on the quantity of traffic in the system. Following a failure, fault detection may be done either at the destination node of a disrupted demand or at some earlier point along the lightpath, typically the first node downstream from the failure. Because the former, which is typical of an AIS detection scenario, does not entail fault isolation, restoration is restricted to a path-based approach, unless some supplemental method is used to locate the fault once it has been detected. The latter, however, does entail fault isolation, so restoration can be either link-based or path-based.

  After the fault has been detected, the demand must be restored on a backup path if its quality of service so requires. The backup path can be precomputed at the time the demand is provisioned or can be computed at the time of the fault. The former is usually preferable, as the restoration process need not be delayed by computation, but it becomes necessary to police all channels reserved for the backup path to ensure their availability at the time of a fault. In the case of $1 + 1$ protection, the backup channels are never used for any other purpose, while in the case of $1:1$ protection, the backup channels are available for use by another demand; if those channels become used, however, a new restoration path must be precomputed for the original demand. In the case of fault-time computation, this continual monitoring of the

backup path is unnecessary, and there is no possibility that a backup path will have to be computed more than once per fault event; the resulting savings in computing and signaling resources make fault-time computation attractive for demands with a lower quality of service or for networks with a lower incidence of faults.

Once the backup path has been determined, the provisioning of the backup path can begin. In a centralized approach, the central operating system learns of the fault from the detecting node, performs any necessary path computations, and signals to all relevant nodes to make the required cross-connections. In a distributed approach, the procedure is initiated by the detecting node, but this node does not necessarily lie on the restoration path and may therefore have no cross-connections of its own to reconfigure. Nonetheless, the detecting node signals to other nodes, which in turn signal to other nodes, until all pertinent nodes have received instructions to reconfigure their cross-connections, thereby restoring the demand.

If a precomputation approach is used, a new backup path must be precomputed at this time in case of a second fault. Moreover, new restoration paths must also be determined for any other demands whose precomputed restoration paths used channels now in use by the demand just restored. A typical fault therefore causes a flurry of signaling and computation even after the broken demand has been restored. The situation may be conceptually more complex if the affected demands have varying levels of service, but the resulting level of activity is reduced insofar as the demands with lower quality of service may not require precomputation at all.

2.4 SOSFs, PROTECTION, AND OXC INTERNETWORKING

Many switch fabric technologies have been employed in protection applications, such as MEMS [Lin, 1999], semiconductor optical amplifiers [Takeshita, 1997], integrated optics [Uehara, 1997], and discrete optomechanical switches [Miyazaki, 1999; Levy, 2001]. It is important to note that only limited connectivity is needed in most of these applications and that a strictly nonblocking cross-connect provides superfluous connectivity. Simpler and more specialized fabrics can therefore provide a level of optimization sufficient for performance in the stringent central office environment.

The concept of shared protection need not be limited to optical networks with simple ring geometries. As the need for connectivity increases, collections of ring networks form larger mesh networks, with quality-of-service guarantees remaining a critical feature. Local connectivity and the overall demand for communication bandwidth is dictated by many factors, including the size of the local population center as well as local industry types. Thus, any mature network might well employ both large and small OXC fabrics appropriately.

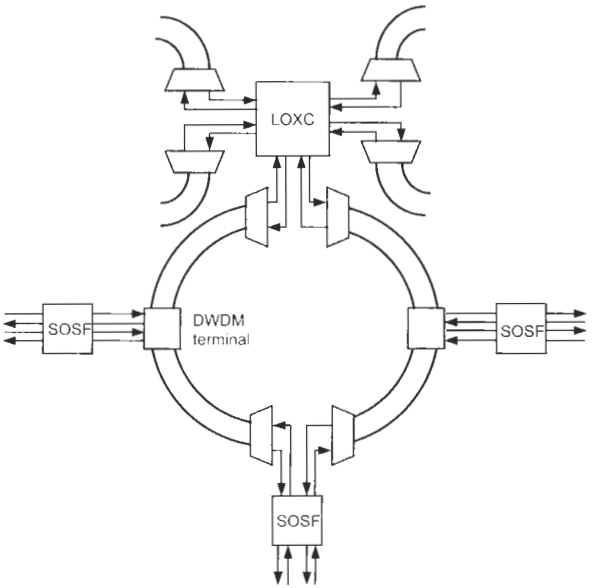


Fig. 2.12 An intermingling of ring networks using small OXC fabrics and large OXC fabrics at mesh points.

In such mesh networks, where the smaller, constituent ring-type networks meet, larger OXCs can participate in such OADM and protection functions as well. Such interconnected behavior is sketched in Fig. 2.12. Although perhaps the smallest unit for SOSFs is the 8 × 8 fabric discussed earlier, the inevitable trend towards larger mesh networks means that larger OXCs will be needed in the future, and the constraints that make them viable are challenging to meet, as discussed below.

2.5 CONCLUSION

Traditional voice and data networks have been designed to provide a high degree of survivability to localized network outages such as failures in the transmission interfaces or cabling. Client equipment features such as dynamic routing SONET-type automatic protection service have been used to help standardize the quality of service carriers that end users have come to expect. To date, systems based upon optical layer transport still place the responsibility of surviving transport layer failures onto the incumbent electronic equipment (e.g., IP routers, SONET, ADMs). However, as the shape and size of optical networks increase and become more complex, greater survivability and functionality will be demanded of the optical transport layer itself, and optical cross-connects of all shapes and sizes will play a critical part in obtaining this goal.

As optical transport architectures evolve from point-to-point, to rings, to mesh-type architectures, OXCs will enable optical channel bandwidth management functions through wavelength grooming, automatic rapid provisioning, and faster restoration. To support these functions, we have explored various OXC architectures. Some networks and office architectures will implement a hybrid architecture scenario where electrical cross-connects are used for traffic termination and re-grooming, and an optical cross-connect is used for pass-through traffic. Others may use the all-optical cross-connect architecture with both opaque and transparent interfaces, where client-side interfaces could be opaque to have OEO functionality, and network-side interfaces could be transparent. In either case the technology and standards work need to address regeneration and monitoring challenges that are required to support transparent payloads, so that the intelligent optical network can be achieved.

The design of any OXC and the underlying switch technology requires careful consideration of most or all of the key properties discussed herein, but it is the device's operating environment that ultimately dictates the component performance requirements. An important directive for optical switch fabric designers is that the selected switch technology must be compatible with the type and arrangement of various network and edge equipment, something that has proven difficult to do for most commercialized technology options presently available. It is important, however, to note that in some cases the type of application is more critical in determining the technology of choice and may make one choice more viable than another. For example, applications requiring OXCs with small or moderate port counts (i.e., less than 256 × 256) may choose waveguide technology, but applications needing OXCs with larger port counts may require free-space technologies. In considering the use of any switch technology, price and performance parameters such as insertion loss, switching speed, port count, etc. continues to be critical tradeoffs in what seems to be a never-ending battle for system simplicity and efficiency.

Acknowledgments

We are grateful to Neil Jackman and Benny Mikkelsen, from whose work we have drawn to complete our own.

3. Requirements and Technologies of Optical Switch Fabrics

Although cross-connects can be located in various places in an optical network, they are customarily located within the central office (CO) due to the physical convergence of the fiber transport layer, optical performance limitations, size constraints, and control issues. A typical placement of optical switch

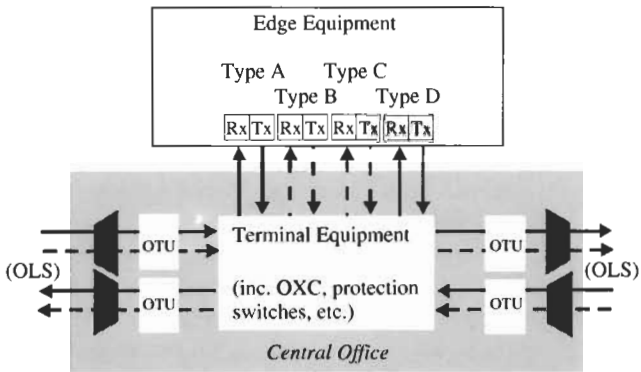


Fig. 3.1 A small OXC in a 4-fiber ring configuration. The reconfigurability that the fabric allows is among the 8 input and 8 output ports shown here. Note that the type of transmitter and receivers located at the edge can vary. The dark circles are optical taps provided for performance monitoring. The dashed lines are protection channels that the client edge equipment can utilize on a preemptive basis. The solid lines are service channels.

fabrics in a CO is shown in Fig. 3.1, where an Optical Cross Connect (OXC) is positioned directly between the client's edge equipment and the network's optical transponders units (OTU).

As discussed elsewhere in this chapter, the OXC can be used to make and break connections between ports on the optical network and the service inputs from the client edge equipment and in doing this the OXC must be optically compatible with both the edge equipment as well as the line system equipment. As will be discussed below, insuring complete compatibility among the wide variety of optical performance parameters is very difficult.

3.1 OPTICAL PERFORMANCE PROPERTIES OF OXCs

Signals traveling through the CO encounter degradations that can be caused by performance parasitics that are both internal and external to the OXC. In addition, most available optical switching technologies have an optical performance that is dependent upon the specific internal path taken through the fabric. Factors such as fiber (or waveguide) bend loss, propagation loss, splice loss, polarization-dependent loss, switch element loss, and the placement of crossover junctions make it hard to design a switch fabric where all paths perform equally. In addition to internal path-dependent performance variations, there are external variations as well. For example, for an OTU-to-OTU connection within the central office the path losses may be negligible, but for connections to OTU equipment that originate at the network edge, path losses due to patch panels, propagation losses, or dispersion effects may accumulate.

In general, these losses are small (usually around 1 dB), but they need to be taken into consideration.

In the CO depicted in Fig. 3.1, the properties of the transmitters and receivers in the OTUs and the edge equipment are critical in determining the allowable system penalties due to polarization-dependent loss (PDL), wavelength-dependent loss (WDL), temperature-dependent loss (TDL), crosstalk, and so on. However, the presence of both internal and external path-dependent performance variations in the OXC makes the characteristics for a given port-to-port OXC connection difficult to predict; when specifying the performance specifications for the entire device, it is usually easier to define the OXC specification as that specification for the most severely affected path.

3.1.1 Port Count

Small-scale optical cross-connects are ideal for a variety of per-channel applications, such as small channel count programmable optical add/drop multiplexers (OADMs), protection switches, and performance monitors. Such devices can be found in hundreds of central offices today. Large matrix OXC fabrics, on the other hand, are typically deployed at regional or long-haul COs where a large number of wavelengths or fibers converge. In some offices there will be a combination of the two OXC types, where small per-channel switches will reside next to the larger matrix switch fabrics for applications such as protection switching.

3.1.2 Scaling

Ideally, OXC technologies should allow vendors to increase the capacity of their switches in line with carrier requirements in an incremental or scalable fashion. However, such technologies do not really exist today. Many switch technologies can be used to create very effective 8×8 or 32×32 switches. For example, discrete 1×2 and 2×2 switches can be used to form strictly nonblocking switch matrices, tree structures, or crossbar arrangements of theoretically arbitrary size. However, these matrices grow on the order of N^2, where N is the number of inputs and outputs, so optical impairment such as loss quickly add up. In general, most OXCs of this kind are confined to a single physical plane and thus can lead to large package sizes. Large $N \times N$ OXCs can also be fabricated with larger $M \times M$ devices ($2 < M < N$) in various multistage arrangements. However, this approach usually results in a size requirement (e.g., footprint) that is prohibitive in constrained environments such as central offices and involves extensive fiber interconnections between stages. In general, most fabric technologies that make very good small port count OXCs do not have the necessary scaling behavior. Accordingly, to date it has not proved cost-effective to use a technology that can scale to large port counts and use it in small port count applications.

3.1.3 Granularity

Some optical switches can be little more than provisionable patch panels, linking together all of the wavelengths in a fiber to any number of other fibers. However, by adding signaling and performance capabilities to the optical switching system, one can provide critical autonomous protection switching capabilities, such as the dynamic rerouting of optical service paths around network failures. In general, system design and control become much more complicated if individual wavelengths must be separated from a group and switched onto a different fiber, as may be the case with optical add-drop multiplexers or other per-channel optical cross-connects.

It is worth noting that the majority of today's optical switches do not intrinsically perform wavelength conversion and cannot simply route traffic from a certain wavelength on one port to a different wavelength on another port. For this to occur the signals passing through the OXC need to undergo an OEO conversion somewhere in the network. Currently this is done by routing the signal out of the switch, converting it to an electrical signal at a transponder unit, and then retransmitting the signal at the new wavelength back into the switch. Although great strides have been made in this area and some enabling technologies have appeared (e.g., wavelength conversion in semiconductor optical amplifiers), all-optical wavelength converters are still not widely utilized in optical networks.

3.1.4 Switching Time

The complete optical state-to-state switch time is the time measured from the leading edge of the input control signal to the leading edge of the optical output signal, once the optical output is stable (stability herein meaning that the output has settled to at least the 90% point on an upward transition, or to at most the 10% point on a downward transition). For reference, the reader is referred to Bellcore specification TR-NWT-001073. In general the required switching speed depends upon the application. There are presently three general switching applications in optical networks: packet switching, protection, and restoration, each of these applications has a different switch time requirement.

Optical packet switching eliminates the speed limitations typically found in router or switch backplanes and is a way to expand overall throughput capability. Such switching needs to occur on a per-packet basis and timescale, which can be in the nanosecond regime at faster bitrates. In this application individual packets are forwarded to different output port destinations. To many, optical packet switches are referred to as "optical routers," but note that any real packet router must read the destination address of the packet and process it accordingly, so these devices are not strictly all-optical.

Protection switches have a less stringent switch time requirement than packet switches; however, note that for protection applications the type of data and the associated service level agreements primarily dictate the overall network requirement on the required optical switch times. Most protection switches available today operate on the order of 1 to 10 milliseconds. For autonomous protection switching this switch time is satisfactory. In more complicated protection schemes that involve complex internodal signaling, the total protection time of the protection switch itself is small compared to overall system response time but should be kept as small as possible.

Fast protection switches are usually used in networks of limited geometry (rings or small mesh-type networks) where the signaling and control is tractable. Failures in much larger-scale mesh networks in which there are both a large number of nodes and a large number of wavelengths involved present a much more complicated task to restore traffic. Because of the more complicated nature of the network, such "restoration" type applications typically have reaction times measured on the hundreds of milliseconds to seconds. Although the optical switch transition itself is usually a small part of the overall restoration time, much larger switch times are typically acceptable.

3.1.5 Switching Frequency and Symmetry

The frequency of switching also can have an impact on the design and associated technology choice of an OXC. In some applications such as protection switching, devices are called into service only on very rare occasions, possibly never. In this case, maintaining port-to-port connectivity with little or no external control is a requirement. Some switch technologies are better suited than others at maintaining their switch positions. In general, switch technologies that involve analog control mechanisms are more susceptible to drift and control problems. Reliability of such technologies is a concern for all network engineers, but there is not enough field data yet to produce a high level of confidence in them.

Symmetry in the switch times between the on-off and off-on transitions is also important. Rise and fall times are both critical when evaluating switch technologies for active switch implementations such as packet switching. If these times are not equal, then a switch may go out of phase during a period of switching, with unpredictable results. In addition, such a switch used for optical protection applications may induce measurable service failures upon reversion.

3.1.6 Optical Power Budget

Optical networks depend sharply on the precisely defined system engineering rules which guide them. The optical power budget, defined as the difference in receiver sensitivity minus the minimum transmit power, is one of the most

critical in that it determines the quality (and thus the cost) of the transmitter and receiver interfaces supported in the network as well as the losses and other various parasitic phenomena that can exist between them. In principle, if the introduction of a lossy optical switch requires the addition of an optical amplifier to overcome the additional path loss, then that cost must also be considered. There are two main causes of loss in an optical switch: the coupling of the fiber to and from the switch fabric and the loss introduced by the switching element itself.

The properties of the transmitters and receivers in the network inherently serve to set the upper bound on the optical performance characteristics of the OXC. For reference, the optical interface specifications for SONET-based OC-48 and OC-192 interface equipment are listed in Tables 3.1 and 3.2, respectively. Specifications for Gigabit ethernet interface equipment are listed in Table 3.3.

For a given application type, the difference between the minimum transmitter power, $T_{\min}$, and minimum receiver sensitivity, $R_{\min}$, determines the maximum allowable insertion loss. This figure, called the optical power budget (OPB), is also shown in Tables 3.1–3.3. The data indicates that both SONET

Table 3.1 **SONET OC-48 Interface Specifications***

Application	*SR*	*IR1*	*IR2*	*LR1*	*LR2*	*LR3*
Wavelength (μm)	1.3	1.3	1.55	1.3	1.55	1.55
$T_{\max}$ (dBm)	−3	0	0	3	3	3
$T_{\min}$ (dBm)	−10	−5	−5	−2	−2	−2
$R_{\max}$ (dBm)	−3	0	0	−9	−9	−9
$R_{\min}$ (dBm)	−18	−18	−18	−27	−28	−27
OPB (dB) $= T_{\min} - R_{\min}$	8	13	13	25	26	25

* GR-253-CORE, 2000. Table by Peter Mitev. Note that additional path penalties are not included (see text).

Table 3.2 **SONET OC-192 Interface Specifications***

Application	*SR1/SR2*	*IR1/IR3*	*IR2*	*LR1/LR3*	*LR2′*	*LR2″*	*VR1/VR3*	*VR2*
λ (μm)	1.3/1.55	1.3/1.55	1.55	1.3/1.55	1.55	1.55	1.3/1.55	1.55
$T_{\max}$ (dBm)	0	2	2	13	1	17	13	13
$T_{\min}$ (dBm)	−4	−1	−1	10	−2	10	10	10
$R_{\max}$ (dBm)	0	−3	−3	−3	−9	−3	−9	−9
$R_{\min}$ (dBm)	−12	−13	−14	−13	−26	−14	−24	−25
OPB (dB) $= T_{\min} - R_{\min}$	8	12	13	23	24	24	34	35

* GR-253-CORE, 2000. Table by Peter Mitev. Note that additional path penalties are not included.

Table 3.3 **Gigabit Ethernet Interface Specifications***

Application	*1000B-LX*
λ (μm)	1.3
T_{max} (dBm)	−3
T_{min} (dBm)	−11
R_{max} (dBm)	−3
R_{min} (dBm)	−19
OPB (dB) = $T_{min} - R_{min}$	8.0

* IEEE 802.3, 2000. Note that additional path penalties are not included (see text).

SR and gigabit ethernet optics have the smallest OPB of 5 and 8 dB, respectively. This type of interface equipment is commonly utilized due to its lower cost. After an additional system margin of 1 dB is allocated for various possible path penalties (e.g., fiber loss and patch panel loss), then the maximum allowable OPB (thus, OXC insertion loss) becomes 7.0 dB. Of course this limitation can be overcome through the use of higher-quality transmitters and receivers, and for simple cross-office type applications a single-channel amplifier can be used as well. However, the use of amplifiers will lead to increases in the cost per port, which the usual metric.

A given path in an OXC can have variable loss due to path length differences. For example, differences in the number of switch elements located in a given path through the switching fabric will contribute to variability in the loss, as well as will any performance variations in the elements themselves. These path-dependent losses must be considered when determining the overall OXC performance and its use in restrictive OPB environments.

3.1.7 Crosstalk

Signal crosstalk is an issue when small leakage fields from other signal sources are allowed to add in single fiber (e.g., amplifiers and WDM filters). The effect of crosstalk on the overall performance of optical networks has been found to be the cause of substantial increases in the signal bit-error rate [Goldstein, 1994, 1995].

When optical signals enter an OXC via a port on the fabric, some of the optical power from the other input ports will leak into the designated path due to performance limitations of the switch fabric. Because the OXC has inputs from different sources (each of the same or different wavelengths, polarizations, and bitrates) the crosstalk contributions can be either coherent or incoherent [Agrawal, 1997]. Incoherent crosstalk normally falls outside the range of the detector's bandwidth and usually has a small power penalty. With the use of complex OXC components in optical networks, the nature and extent of the crosstalk contribution can be very difficult to predict in advance.

Crosstalk effects can become more pronounced when there exists an optical power divergence between the various inputs to the switch fabric. As discussed earlier, Tables 3.1–3.3 show that the spread in transmitted powers at SONET SR OC-48, SR OC-192, and Gigabit Ethernet can be as large as 11 dB. This power divergence must be taken into account when determining the crosstalk requirement on the particular OXC.

In general, OXCs come in numerous shapes and sizes and can contain a wide variety of cross-points, discrete switch elements that are crosstalk contributors. Clos [1952] has shown that the number of cross-points in large, $N \times N$ cross-connects can be minimized by arranging the switch fabric into a small number of switching stages n. Here, each stage is a smaller scale $M \times M$ cross-connect that can also be a crosstalk contributor. Thus, when designing n-stage cross-connects, n crosstalk contributions must be accounted for. In general, smaller n fabrics are more advantageous because of other parasitic properties, such as insertion loss, that also depend upon n.

Ultimately the presence of crosstalk serves to degrade the quality of the service signal, resulting in transmission disruptions in the optical network. The measured, crosstalk-induced OC-192 system power penalty is shown in Fig. 3.2.

As seen in the figure, a copolarized crosstalk contributor has a greater power penalty over that of a cross-polarized crosstalk contributor by about 2 dB. This is because of the superposition of electric fields in the copolarized case. Analysis of polarization-dependent crosstalk effects has indicated that systems with randomly polarized fields (such as those within the typical optical network) tend to behave in a worst-case fashion [Goldstein, 1995]. Such effects must be considered as well in the design of any OXC.

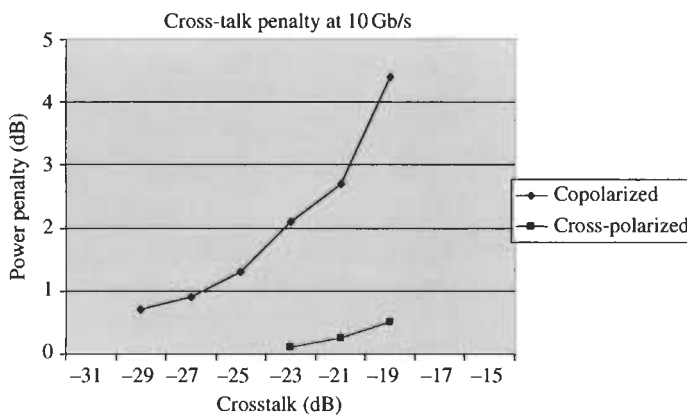


Fig. 3.2 The system power penalty versus crosstalk at a bit-error rate of 10^{-10} at OC-192 (data taken by Haisheng Wang).

3.1.8 Polarization Mode Dispersion

Polarization mode dispersion (PMD) occurs in optical networks because fiber optic signals are composed of two orthogonal modes of polarization that travel in the fiber at different speeds [Namihira, 1989]. The differences in propagation delays between the two modes and the conversion between the two modes that can occur along a length of fiber make it difficult to interpret the optical signal at the receiver, typically resulting in a measurable degradations in bit-error rates [Poole, 1991; Khosravani, 2001]. In general, PMD is usually a significant issue only at high bitrates (>10 Gbps) and long fibers (>~60 km), and it is considered by many system designers to be a gating factor to transmission of 40 Gbps signals over long distances in conventional fiber. As a result, any optical component (such as an OXC) that can contribute to overall PMD accumulation is a potential problem. For the most part PMD is an issue in optical cross-connects that are based on birefringent waveguides or materials, such as lithium niobate. For these devices, some sort of PMD compensation should be used. Most mirror-based switch technologies (e.g., optomechanical, MEMS) have negligible PMD performance characteristics (e.g., <1 ps). Polarization mode dispersion is less of an issue in the limited environment of the central office than it is in the network line system, on backbone.

3.1.9 Wavelength Dependence

Optical cross-connects must operate at different wavelengths, depending upon where they are situated and the nature of the transmitters that surround them. Edge equipment typically communicates through the OXC switch fabric directly to the optical line system (OLS). As shown in Fig. 3.1 the edge equipment communicates directly to the OLS OTUs responsible for wavelength conversion to the OLS. Thus, these transmitters and receivers at both the edge and OLS must be wavelength-compatible. In addition, the edge equipment must have OLS-compatible transmitter and receivers in cases where the OLS OTUs are removed. This scenario is typically called one of "compatible optics."

3.1.10 Power Consumption

The allowed power consumption of an optical switch fabric varies by system and ultimately depends upon the application. In general, switches in optical networks are inactive for most of their life and should have negligible power consumption when not being used in order to decrease overall costs of operation. Optical switch fabrics that draw their power from common telecommunication shelf backplanes typically have power draw maxima of about 100 watts or less. Associated heat dissipation by the OXC and/or neighboring equipment can lead to substantial increases in the OXC's ambient

temperature. Thus, an OXC that operates over a wide temperature range and with low heat dissipation (or small heat dissipation times) can be advantageous as well.

3.1.11 Physical Size and Fiber Interface

The number of shelves and bays an OXC requires is a crucial consideration in determining the best technology for a given application. Obviously, the greater the size of the OXC the more a carrier will have to pay for the facilities in which it is housed. The fiber interface is also an important concern in terms of reliability and cost. A free-space fiber array has to be extremely stable or an optical signal can be entirely lost. In general, fiber management between single- and multistage OXCs is a great concern.

3.1.12 Latching

From the point of view of the switch designer it is useful to have a switch action that is stable; i.e., the switching function saturates into the desired states. Ideally, there should be no need to provide feedback to hold a particular switch state, and ideally no need to bias the switch. Sometimes switches that behave in this manner are called "digital." From the point of view of the network engineer who is concerned about network survivability, however, a stronger form of latching is required: the switches must hold their states in the event of a disruption of power. Devices with switches that function, for example, by current drive would not latch under this definition if a power failure interrupted the current. Redundant power supplies can make this feature unnecessary. Also, latching switches can aid in reducing the power consumption of OXCs when not in use, an issue that can become important in older, smaller central offices.

3.1.13 Reliability

Guidelines for the kind of performance and reliability testing that need to be met for particular classes of components are published in documents provided by the various standards bodies. These include performance criteria that should be met under a wide variety of environmental conditions. Some of the operational and nonoperational tests last as long as 5000 hours; therefore it is important to ensure that the testing procedure and test equipment itself is reliable so that the data obtained are as accurate as possible. Both operational and nonoperational tests help assure that the switches will perform when needed. Stiction tests (not widely found in the standards bodies) are also suggested for OXC technologies that have moving parts.

The text under heading 2.5 on pages 326–327 should have been placed on page 367 as the conclusion to Chapter 7. The publisher regrets this error.

OPTICAL FIBER TELECOMMUNICATIONS,
VOLUME IVA

ISBN: 0-12-395172-0

3.1.14 Nonlinear Effects

Various nonlinear materials can induce parasitic penalties on the signals traveling through them. Such penalties originate from well-known phenomenon such as four-wave mixing, self-phase modulation, and cross-phase modulation. Most OXC technologies available today do not have these problems, because most approaches to date have involved predominately refractive linear materials. Also, most switching is presently done on a per-wavelength basis, so cross-channel effects do not occur. However, in order to adhere to the increasingly stringent CO operating environment, complex OXCs that include inline amplification (e.g., SOAs) may be used to nullify (or minimize) the loss of the switching elements within the OXC itself. These technologies provide new performance concerns when used in optical networks.

3.1.15 Other OXC Functionality

It is important to describe, albeit briefly, other performance characteristics and functionalities that are desirable as well. For example, variable optical attenuators built into the fabric can be used to control the output power levels leaving the fabric. In the case of "compatible optics," discussed earlier, where the signal output from the fabric goes directly into the line system, this is critical in controlling OSNR and crosstalk penalties that can be exacerbated further downstream.

In addition, power splitters are needed not only for optical power taps for performance monitoring but also for "keep alive" signals used to continuously drive protection lines. Such signals can quiet unnecessary equipment alarms and can assist in faster turn-on equipment times during a protection event. Integrating splitters into the OXC can also decrease the footprint and overall path losses of the switch fabric as a whole. Such integration is not feasible in discretely packaged devices (e.g., 1×2 and 2×2 optomechanical switches), but is easily done in integrated optic or microoptic switch fabrics [Lin, 1999].

3.2 TECHNOLOGIES

Switch technologies can be classified by the general optical phenomena utilized to direct the optical signal between the various input and output ports. There are numerous modalities of optical switching, such as thermooptic, electrooptic, acoustooptic (bulk), optomechanical, and micromechanical. Each of these technologies can usually be executed in more than one material system, the use of which usually depends upon costs, the degree of complexity of the overall photonic integrated circuit, and the level of integration needed.

Switching media can be roughly divided between free space, waveguide, or a hybrid of both. In some free-space approaches, an input waveguide in the form of a fiber can be mechanically moved to line up with a selectable

output fiber. Other such approaches have the fibers fixed but have bulk optic mirror or prisms redirect the light via mechanical motion to specific output ports. The use of MEMS-type devices, where microscopic rotating mirrors are used to direct signals between input and output mirrors, has aided in the miniaturization of conventional technologies. In waveguide-based approaches, light is coupled into a substrate that contains multiple waveguides, and various optical phenomena are utilized to nonmechanically change the angle of the light to redirect the signal between the waveguides.

Although many types of switch technologies are available [Jackman, 1999], in the following discussion we primarily focus on advances in bubble switching, liquid crystal, silica-on-silicon, polymer, and MEMS-type devices, due to demonstrations of their utility in OXCs of significant size and functionality.

3.2.1 Bubble Switching

Bubble-type switches have a waveguide that intersects at an angle with a trench that is filled with an index-matched liquid. At the bottom of the trench lies a heater that is turned on or off, causing a phase change in the liquid and thus a change in refractive index. This refractive index change thereby increases the reflectance at the waveguide-trench interface, creating a selectable mirror, thus a switch. A simple sketch of a bubble switch is shown in Fig. 3.3. In this device

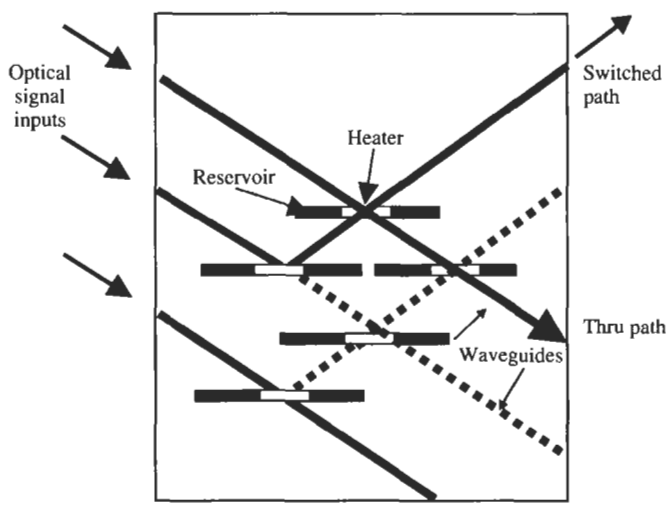


Fig. 3.3 A bubble switch-type optical switch fabric. Here the circles represent the trench region in which heaters serve to modify the refractive index of the material in the trench, thereby modifying the total internal reflection that occurs at the waveguide–trench, or liquid, interface.

an optical signal transverses many fluid-filled trenches and reflects off of the waveguide-bubble interface.

A 32 × 32 bubble switch-based OXC has been demonstrated (in silica on silicon) [Agilent, 2000]. Overall (path dependent) insertion losses are less than 8 dB with switch times less than 10 ms. In this switch fabric technology the characteristics of the trench (e.g., sidewall smoothness and trench width) and the waveguide width are crucial in determining the diffraction loss encountered for signals traversing the trench. Bubble switches in other material systems have been demonstrated as well (www.opticalcrosslinks.com). A similar approach uses a thermal capillary switch element that moves liquid in and out of a trench [Makihara, 2000].

Note that in this nonlatching technology the bubbles need to be maintained for the duration that the switched path needs to be maintained, which can be a very long time in some optical networks. This, combined with a necessary hermetic, current-controlled environment, is needed to control the phase of the liquid in the trench.

3.2.2 Liquid Crystals

The liquid crystal (LC) material technology commonly found in display monitors for computers can also be used to make large-scale optical switches. In such devices the nature of the birefringent LC is used as a polarization beam router, selectively switching the polarization of signals passing through. Once the polarization has been changed, the signal is sent through a birefringent material (e.g., calcite) where the walk-off angle combined with the material thickness is used to physically displace the beam at the output. Using a matrix of these basic switching units can be used to form an $N \times N$ OXC. The fundamental beam-steering mechanism used in a typical LC switch is sketched in Fig. 3.4. For example, a 64 × 64 Benes OXC using spatial light modulators has been demonstrated [Noguchi, 1991, 1997]. The measured insertion loss was

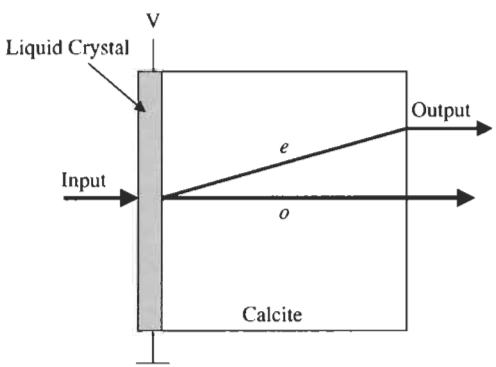


Fig. 3.4 Beam steering in an LC directional element.

found to be 9–10 dB, with a PDL and isolation of 0.21 and 25.9 dB, respectively. Note that the authors used a half-wave plate to eliminate the polarization effects (e.g., PDL, PMD) typically encountered in LC-based devices.

Liquid crystals are also the key enabling element in recent advances in holographic optical switching. Such devices use a hologram recorded onto a liquid crystal spatial light modulator to steer the input beam to the desired output port. A 1 × 8 device based on a reflective ferroelectric liquid crystal (FLC) placed over a silicon spatial light modulator was recently demonstrated, and a 3 × 3 OXC using a 2-FLC SLM design was also constructed and tested [Crossland, 2000]. Although these devices suffered very high insertion losses and large crosstalk between input ports, it is believed that these losses are independent of the number of ports. Also, the fact that these devices do not need continuous analog control makes them particularly attractive in terms of overall system engineering.

3.2.2.1 Silica-on-silicon

Silica-on-silicon (SOS) waveguide technology has been widely developed, and such devices can have low propagation, bend, crossing junction, polarization dependent, and fiber attachment losses. Switching in SOS is performed using a temperature-induced refractive index change, called the "thermo-optic effect." Recently a 16 × 16, cross/bar OXC was demonstrated using a dual Mach-Zehnder 2 × 2 switch element [Goh, 2001]. Such an architecture is well-suited for thermo-optic devices because only one switch element per input needs to be active, thereby reducing overall power consumption.

The promise of SOS technology is not just in OXC-type applications, but it also holds promise toward the goal of obtaining larger-scale integration with other passive components (e.g., arrayed waveguide gratings and splitters). For example, OXCs can be integrated between a MUX and DMUX to obtain a reconfigurable OADM. To date such greater integration has been difficult to obtain because of fabrication problems in addition to thermal management (i.e., thermal crosstalk or thermal uniformity) difficulties. Also, operation at both 1.3 μm and 1.55 μm is difficult with this technology because an adiabatic, or digital switch, is hard to implement due to wide spacings required between switch element waveguides. The heaters used in SOS operation typically require large separation.

3.2.2.2 Polymer

Polymer-based waveguide devices have the promise of low loss, low power consumption. In addition, wide operating wavelength ranges are possible since digital switch elements can be used. A SNB 16 × 16 tree structure has recently been reported with 6 dB of loss and 30 dB extinction [Rabbering, 2001]. This OXC has shown a great improvement in loss over past polymer devices. Single switch element extinction ratios have been shown to be better with this

technology than with many other guided-wave technologies. Improvements in material reliability and an ability to integrate multiple components on the same chip will result in this technology being useful for many central office optical switching applications.

3.2.3 Microelectromechanical Systems (MEMS)

Various MEMS technologies can also be used to realize relatively large switching fabrics with sub-millisecond switching speed. This technology combines free-space interconnections with the advantages of monolithic integration on silica platforms. The switching concept can be based on moving mirrors rotated around micromachined hinges. There are two basic types of optical cross-connects to date: 2D- and 3D-type OXCs. The 2D OXCs are usually N^2, crossbar-type fabrics, where N is the number of input ports and N^2 is the number of mirrors. The operation of the mirrors in these devices is strictly digital; they have only two positions: up (signal pass) or down (signal reflect) (see Fig. 3.5). A 2D bidirectional 8×8 switch employing such a switching principle has been reported [Lin, 1998]. The insertion loss was ~3 dB with individually optimized fiber-to-fiber couplings, and the isolation ratio was greater than 60 dB. An important issue for such switches is mirror size, position, and angle, which are basically fixed during the lithographic process. Also, in such OXCs the worst-case path loss is usually critically limited by beam divergence, in addition to the mirror losses.

Larger 3D MEMS fabrics have the greatest potential for large port-count OXCs. In such fabrics the mirrors are allowed to pivot in two dimensions, but the arrangement of the mirrors in a 3D manner allows for full nonblocking connectivity with a smaller number of mirrors (usually only $2N$ mirrors are needed). A 3D MEMS optical cross-connect is pictured in Fig. 3.6.

Note that each port-to-port connection undergoes two reflections; thus there are $2N$ mirrors needed to support N connections. The benefits of this

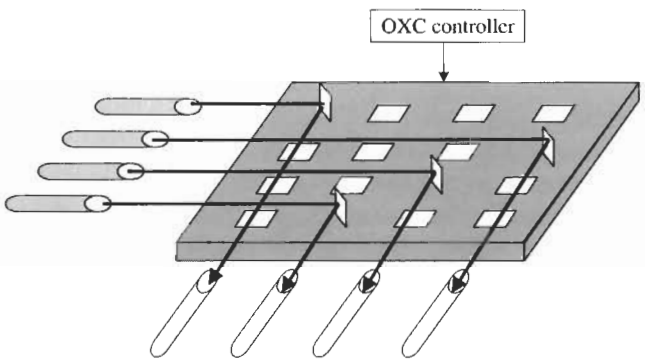


Fig. 3.5 A two-dimensional 4×4 MEMS cross-bar optical cross-connect.

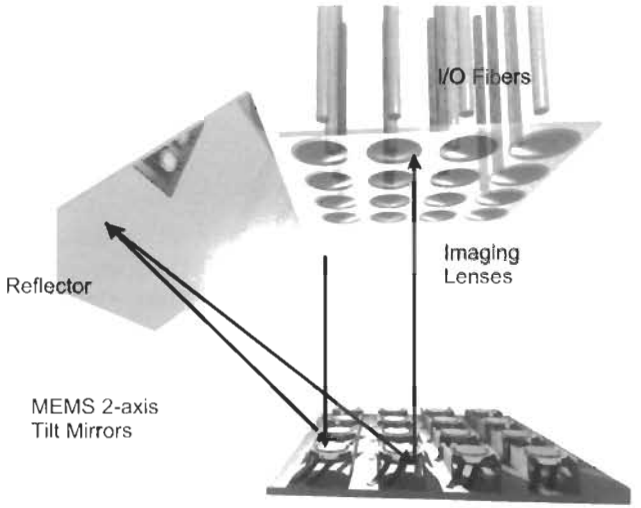


Fig. 3.6 A picture of a 3D MEMS optical cross-connect. Light is collimated across the switch fabric via a lens array. The use of the reflector reduces the package size and the number of chips needed to perform the switching function. See also Plate 4.

MEMS approach are that the losses are not as path-dependent as the 2D design and are relatively independent of the total port count. However, this technology introduces a formidable control problem in that each mirror needs two analog voltages to control its 2D rotation; and since there can be a large number of mirrors on a given fabric, finding cheap and fast methods of control of the entire fabric has become a serious development issue.

3.3 SUMMARY OF SWITCH TECHNOLOGIES

Typically, the free-space beam-steering type switch's, e.g., 3D MEMS, switching speed is an order of milliseconds and has fairly low loss for a large size switch module. The large size switch module provides an ability to implement single-stage architecture that can be easily scaled into hundreds or thousands of ports, but introduces difficult control electronics complexities that may make it less attractive for small port count applications. On the other hand, the waveguide-type switch's switching speed can be on the order of nanoseconds, allowing it to be used in fast packet switching, restoration, and protection type applications; and it also may have simpler control electronics. However, the use of waveguide-type switches may yield higher system losses when used in a multi-module architecture to build a large port count OXC.

In general, applying individual switch component technologies, not as discrete switches but rather for uses in larger matricies of sufficient sizes to be useful optical network cross-connect applications, has encountered extremely difficult manufacturing and performance issues. It should be noted that the

reverse holds true as well: that component technologies optimal for large-scale OXCs are not necessarily most efficient when scaled down to small port count, blocking or nonblocking OXCs. Although each technology has unique performance characteristics specific to the optical phenomenon utilized, it is ultimately overall manufacturability, size, and both component and system cost that usually make one approach more viable than another.

4. Optical Cross-Connects

4.1 OPTICAL SHARED PROTECTION RING APPLICATIONS

4.1.1 Introduction

The architecture of a typical electrical mesh network is shown in Fig. 4.1(a). Each router has a direct electrical connection to all others in the network. User data that are sent into the network enter the local router and are transmitted via the appropriate output port. Note that the connectivity of this network is complete, or nonblocking. In other words, all data entering the network at one router can be directed towards any other router on the network.

Such user connectivity can also be obtained in a fiberoptic ring through the use of multiplexed optical wavelengths, as shown in Fig. 4.1(b). The network shown has only a single transport medium, which interconnects all of the network users. The users are usually not directly connected to the optical network, but rather are connected to some piece of edge equipment such as a router.

Routers typically have a large number of low-speed ports for connections to the users (e.g., 10/100BaseT) and a larger number of high-speed optical output ports for connections to the network (e.g., 1000BaseT, OC-48, OC-192).

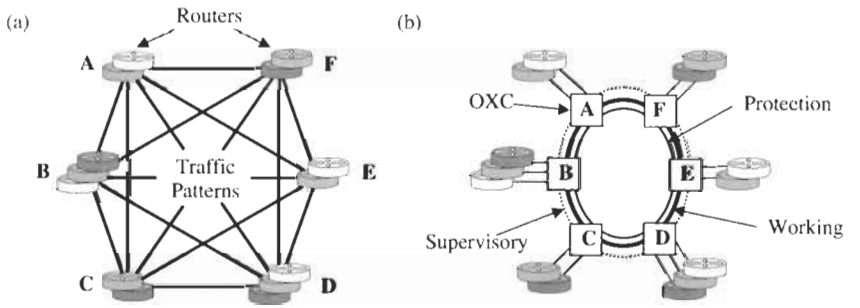


Fig. 4.1 (a) An electrical network showing full nonblocking connectivity; (b) A DWDM fiberoptic ring network showing the connectivity that is supported. The nodes, depicted as squares, contain MUX/DEMUX equipment.

Connectivity between routers or edge equipment is obtained by communication over individual wavelengths carried in the fiber. Inputs from users on the network enter the local node or central office and are typically multiplexed into a higher bit-rate carrier frequency (e.g., TDM, as done in SONET) or transmitted directly into the ring (e.g., strict IP over WDM). In either case, point-to-point connectivity is obtained through a common wavelength used for adding and dropping data to and from the ring.

Although it is possible to build a strictly nonblocking WDM fiberoptic network, the majority of today's WDM fiberoptic networks do not support strictly nonblocking connectivity due to the nature of existing traffic patterns and the equipment costs of building such networks. Service data destined for a particular user may have to undergo wavelength and even bit-rate translation while being carried from node to node. Despite this, however, the ability to reconfigure the service channels that are added and dropped at each node allows enhanced connectivity on the network. To date, this functionality, sometimes called optical add/drop multiplexing (OADM), can be obtained by using one of several techniques, such as tunable lasers and filters, nonlinear optical wavelength conversion, and optical switch fabrics.

In addition to the required connectivity, the ability to maintain service traffic during various types of network failures is crucial as well. Many kinds of problems can occur in an optical network, including transmitter and receiver failures; temperature-induced instabilities in optical components, such as variation of a laser's central frequency; polarization effects, such as PMD and PDL; and fiber breakage. Some of these failure types, such as a fiber cut, result in a loss of signal transmitted to the client; others, however, induce system penalties through bit-error rate degradation. For providers who offer service level agreements, the ability to measure all these failure types by some method of performance monitoring becomes crucial.

In SONET-based networks, performance monitoring and service channel protection have traditionally been responsibilities of the local SONET add/drop multiplexer. Such networks have traditionally operated in the TDM domain and have been designed to support bidirectional line-switched ring and 1 + 1 protection. However, non-SONET types of optical layer protection have recently been developed, in which reserved optical bandwidth, as in Fig. 4.1(b), is allocated to carry service channel data during a failure event. Note that these protection channels are used in different ways depending upon the failure type and, ultimately, how the particular type of protection capability is provisioned. The way that this reserved bandwidth is exploited during a failure event is critical.

As described, two important elements critical to overall WDM network performance are the ability to ensure the required connectivity and the ability to guarantee service traffic by quickly detecting and reacting to the numerous types of network failures that can occur. Of the many possible approaches to these problems, optical switching shows the most promise due to cost,

scalability, and protocol and bitrate independence. Optical $N \times N$ cross-connect fabrics for OADM or protection applications can be either large or small, depending upon how many users or wavelengths are being added and dropped, but in general small-scale optical cross-connects are particularly suited for these applications. Such fabrics can have their performance optimized because of their limited scope and sometimes limited connectivity. The next subsection will focus on the operating environment, design, and performance specifications of small-scale optical switch fabrics (SOSFs) used for OADM and optical protection applications.

4.2 SMALL LITHIUM NIOBATE SWITCH MODULES

Lithium niobate (LN) photonic switch arrays were among the forerunners of contemporary all-optical cross-connects and OADMs. One early prototype network consisted of 32 × 32 wavelength selective OXC (WSXC) network elements built from enhanced 4 × 4 switch arrays [Murphy et al., 1997] as well as 48 × 48 WSXCs made from crosstalk-reduced 4 × 8 and enhanced 12 × 6 switch modules [Chen, 1999]. More recently, a single-module 16 × 16 has been produced [Murphy et al., 2000].

The enhanced (EN) 4 × 4s use 1 × 2 digital Y switches optimized for 1.55 μm performance to construct polarization independent, single-chip, strictly non-blocking (SNB) tree structures. These devices have been enhanced to increase performance and system functionality, including limited 1 × 2 bridging. After signal duplication by a 1 × 2 passive splitter, a signal can be connected to either of two outputs. The other duplicated signal can connect to either of the other two outputs of the 4 × 4. Hence, after bridging the device consists of two separate SNB 4 × 2s. Also, the device has the ability to disconnect or dump any signal from the network. Thus the module has limited 8 × 12 connectivity. Each 4 × 2 is capable of being in one of 21 states. An extra switch column at the center of the device is added to reduce crosstalk another level beyond a dilated structure. Instead of 24 switch elements and four switch columns of a regular 4 × 4 tree structure, the array consists of 44 1 × 2 switch elements and 5 switch columns plus an additional column of 1 × 2 passive splitters. The Y switch element uses the optimal shaping and other waveguide engineering procedures to achieve the best combination of polarization-independent extinction, device length, switching voltage, and loss.

Like the EN 4 × 4, both of these devices are based on the 1 × 2 switch element. SNB performance is also achieved with an active splitter-active combiner tree structure. The 4 × 8 consists of four 1 × 8s connected to eight 4 × 1s. Three columns of 1 × 2s are required to build a 1 × 8, and two columns are needed for the 4 × 1; so, along with a crosstalk reduction (CR) column the CR 4 × 8 consists of six switch columns totalling 84 switch elements. With ganging, only 32 control voltages are required. As with all other devices discussed here, instead of having a massive waveguide interconnect between the

middle two columns, the switch elements are arranged so the more complicated waveguide interconnects are after the first switch columns or before the last switch columns, thereby minimizing the number of waveguide intersections in each path through the array. Path-loss equalization is achieved with dummy intersections in the paths with the least number of waveguide intersections and with variation in radius of curvature of the S-bends. Such CR 4 × 8 modules have crosstalk performance similar to that of the EN 4 × 4; typical parameters are type loss of 11 dB, 1 dB PDL, and 0.5 ps of PMD.

The EN 12 × 6 consists of a strictly nonblocking (SNB) 12 × 6 with integrated 1 × 2 passive splitters, signal disconnect ports, a crosstalk reduction stage, and a variable attenuator at each output port. It contains 240 1 × 2 switch elements and nine switch columns plus one passive splitter column. Ganging results in 72 control voltages required. There are 72 switches at the center of the module for crosstalk reduction. After ganging, they are controlled with 6 voltages, while crosstalk cubed for all active connections is maintained. Including the dump ports, the device has limited 12 × 18 connectivity. A typical set of EN 12 × 6 modules has an average loss of 15.1 dB ± 1.1 dB, 1.2 dB PDL, and crosstalk averaging 43 dB ± 6 dB.

Four columns of 1 × 2s are required to build a 1 × 16, so, the 16 × 16 consists of eight switch element columns. Each 1 × 16 requires fifteen 1 × 2s, so the entire 16 × 16 consists of 480 switch elements. All 1 × 2s in a given column of a 1 × 16 are controlled with a single voltage because at most one switch in that column will direct a light signal. Hence, each of the first four columns of the array requires one voltage control for each input, and similarly, each of the last four columns of the array requires one voltage control for each output. Therefore, the 16 × 16 requires 128 control voltages, obtained by multiplying 8 columns by 16 inputs or outputs. This results in a package requiring 256 pins, as compared to 960 if each switch element were controlled individually. Unlike switch arrays made with 1 × N free space switches where N crosstalk terms can accumulate at a given output, binary tree architectures have, at most, a single dilated crosstalk term from each combiner stage at an output. Hence, the worst case crosstalk of the fully populated 16 × 16 is four times the crosstalk of the individual elements squared.

There are 256 paths through a single switch, one path for each input/output pair. For the minimum loss polarization, the typical optical power loss is 12.7 ± 0.7 dB, while for the maximum loss polarization, the optical power loss is 14.8 ± 1.3 dB. This loss could be reduced by 5 dB if thin films were used at the end faces instead of angle polishing and quartz waveplates. The 1 × 2 switch elements account for 4 dB of the remaining loss. Typical extinction ratios are 12 dB for the minimum extinction ratio polarization and 17 dB for the maximum extinction ratio polarization. Typical PDL is 2 dB.

4.2.1 Architectural Aspects of Crosstalk Control in Guided Wave Switching Arrays

Figure 4.2 illustrates a 12 × 6 switching array, composed of twelve 1 × 6 fan-out trees and six 12 × 1 fan-in trees. Each of the fan-out trees is associated with one of the twelve input channels, and each of the fan-in trees is associated with one of the six output channels. The fan-out trees are assembled from stages of active 1 × 2 switches and the fan-in trees are assembled from stages of 2 × 1 switches. Each leaf of a 1 × 6 fan-out tree is connected to a leaf on a 12 × 1 fan-in tree. No two leaves of a given fan-out tree are connected to the same fan-in tree. This provides connectivity from each input to each output of the switching array.

Control signals, not shown in the figure, are also coupled to each stage of active switches to control the connection states of the various 1 × 2 and 2 × 1 switches. To minimize the space taken up by control signals, ganging is typically used to control each stage of switches of a given fan-out tree or fan-in tree. That is, all switches in a given stage of a given tree are switched by the same control signal.

The area of interconnection between fan-out trees and fan-in trees is an advantageous location for the placement of crosstalk reduction devices. This is because crosstalk signals may potentially first combine with active signals in the 2 × 1 switches in the leftmost stage of the fan-in trees.

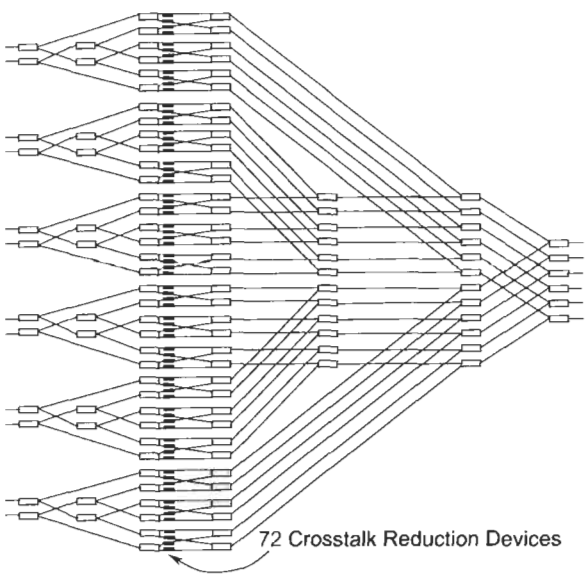


Fig. 4.2 Schematic of a 12 × 6 switch with crosstalk reduction devices.

4.2.2 Generalized Switching Arrays

An N-input by M-output switching array in the form of a tree architecture utilizing 1×2 and 2×1 switches, has $\lceil \text{Log}_2 N \rceil + \lceil \text{Log}_2 M \rceil$ stages, where $\lceil X \rceil$ denotes the smallest integer greater than or equal to X. Thus, the 12×6 tree architecture shown has $\lceil \text{Log}_2 12 \rceil + \lceil \text{Log}_2 6 \rceil = 4+3 = 7$ stages. Additionally, there are $N \times M$ interconnections between the N $1 \times M$ fan-out trees and the M $N \times 1$ fan-in trees. These $N \times M$ interconnections provide sites for the placement of $N \times M$ crosstalk reduction devices. Figure 4.2 also illustrates the placement of 72 crosstalk reduction devices in a 12×6 tree architecture. Conventionally, a total of $N \times M$ control voltage signals are used to individually control the $N \times M$ crosstalk reduction devices. We will demonstrate that many fewer than $N \times M$ control voltage signals are required.

4.2.3 Crosstalk Propagation

The fan-out trees on the inlet (left) side of a tree-architecture switching array are effectively crosstalk generators. Every fan-out tree with an active inlet signal will produce some level of crosstalk on each of its output leaves. Since output leaves on fan-out trees are connected to input leaves on fan-in trees, many of these input leaves will carry crosstalk signals. (Some of the input leaves may not carry a crosstalk signal because not all of the inlets on the switching array may be active.)

Crosstalk impairment occurs when a crosstalk signal enters a fan-in tree on one of its input leaves and subsequently combines with an active signal at a 2×1 switch in some stage of the fan-in tree. The amount of impairment caused by a single crosstalk signal depends on (1) the level of the crosstalk signal as it emerges from the fan-out tree (2) the amount of reduction (or attenuation) of the level of the crosstalk signal effected via a crosstalk reduction device placed between the fan-out tree output leaf and a fan-in tree input leaf, and (3) the amount of reduction (or attenuation) in the fan-in tree of the level of the crosstalk signal, before it combines with the active signal.

Two factors strongly contribute to the values of (1) and (3) above. They are the extinction ratios of the various 1×2 and 2×1 switches and the organization of the control of the 1×2 and 2×1 switches. Similarly, two factors that strongly contribute to the value of (2) are the extinction ratios of the crosstalk reduction devices and the organization of the control of the crosstalk reduction devices.

This section is concerned with architectural aspects of crosstalk control. We are not attempting to reduce crosstalk impairment via improved device extinction ratios. Our operative assumption regarding extinction ratios is that negligible crosstalk impairment will result only if an active signal is extinguished (or "knocked down") at least three times before it is allowed to combine as crosstalk with some other active signal. Our switch array architecture and control scheme are designed with the intention of guaranteeing this result.

The tree-architecture switching array is advantageous in reducing crosstalk impairment, even without any crosstalk reduction devices. Any crosstalk signal reaching an output leaf on a fan-out tree will have been extinguished (or knocked down) at least once. That is due to the very nature of the generation of a crosstalk signal in a 1×2 switch with a finite extinction ratio. The unused output port of the 1×2 switch carries a leakage (crosstalk) signal due to this fact. Similarly, when a crosstalk signal attempts to combine with an active signal on a 2×1 switch, the crosstalk signal will be knocked down once more (as determined by the extinction ratio of the switch) before impairing the active signal. Thus, in such an architecture, only one additional knockdown is required by a crosstalk reduction device.

We are interested in minimizing the number of control voltages required for both signal routing and crosstalk reduction. It will be shown that ganging can work in both cases. It is possible to address this problem either from the perspective of the inlet fan-out trees or the outlet fan-in trees. We consider the inlet side first.

4.2.4 Crosstalk Control—Fan-out (Input) Tree Perspective

In Fig. 4.3, we provide an example of crosstalk propagation that results from a ganged approach for controlling each stage of switches in a 1×32 fan-out tree. The nodes in the figure are to be interpreted as 1×2 switches. The active signal enters from the left and is subsequently routed by 1×2 switches at each stage toward the desired output port. Such a routed active signal is indicated by a dark solid line.

A 1×2 switching element is not a perfect digital switching element, in the sense that when one of the two output ports is selected and connected to the single input port, a small portion of the input port light (leakage) is manifested as crosstalk at the other (nonselected) output port. This crosstalk can propagate in a network and potentially combine with other active signal paths at subsequent switching elements. This type of crosstalk and its propagation are indicated by dark dashed lines.

At each stage of switching, the active signal is routed by a 1×2 switch to the desired output port. Additionally, as just described, a portion (knocked down version) of the active signal is transmitted to the other output port and is denoted as level-1 crosstalk.

Let us now consider the crosstalk implications of controlling the fan-out tree via a ganged approach. Such a method results in all switches in a given stage of the fan-out tree being in the same switch state, i.e., the state required to route the active signal toward the desired output. In the leftmost stage, the first 1×2 switch encountered by the active signal routes that active signal toward the desired output while simultaneously producing level-1 crosstalk at the other output port of the 1×2 switch.

In the second stage of switches, the active signal is further routed toward the desired output and, again, level-1 crosstalk is generated on the other output port of the 1×2 switch. We additionally need to consider what is going on at the other 1×2 switch in this stage. That switch is receiving level-1 crosstalk from the first stage. Because of ganged control, the state of this switch is the same as the other 1×2 switch in the second stage, namely the one that is routing the active signal. Thus, this switch will route level-1 crosstalk to one of its output ports and will generate level-2 crosstalk on its other output port. In general, a signal that has been knocked down n times will be referred to as level-n crosstalk.

Thus, each successive stage of switching in a fan-out tree routes exactly one active signal, routes potentially numerous crosstalk signals of various levels, and creates additional numerous crosstalk signals. More specifically, a given 1×2 switch in a given stage either receives an active signal, which it routes, and creates level-1 crosstalk on its other output port, or receives a level-n crosstalk signal, which it routes, and creates a level-$(n + 1)$ crosstalk signal on its other output port.

In Fig. 4.3, each output leaf of the 1×32 fan-out tree is labeled with a number corresponding to the crosstalk level emerging at that leaf. The leaf labeled with a zero is carrying the active signal (a signal knocked down zero

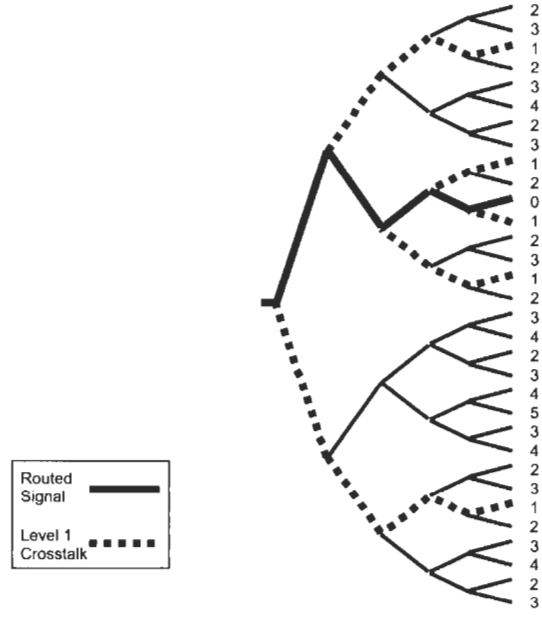


Fig. 4.3 Crosstalk propagation with identical (ganged) control for all switches in a given stage.

times). Thus, we may observe that there are one level 0 signal, five level-1 crosstalk signals, ten level-2 crosstalk signals, ten level-3 crosstalk signals, five level-4 crosstalk signals, and one level-5 crosstalk signal. In general, for an S-stage fan-out tree with 2^S leaves, and for $0 \leq n \leq S$, there will be $S!/(S-n)!n!$ leaves containing a level-n crosstalk signal.

As stated earlier, our requirement is to knock down a signal three times before it is allowed to combine as crosstalk with some other active signal. Any crosstalk from a fan-out tree will be knocked down one additional time when it combines with an active signal at a 2×1 switch on a fan-in tree. Thus, our control of crosstalk reduction devices will not have to consider level-2 or lower crosstalk, since it will be reduced to level-3 or lower before combining with an active signal at a 2×1 switch. So our basic problem in controlling crosstalk will be with level-1 crosstalk signals.

Our approach, as mentioned earlier, is to place crosstalk reduction devices at the intersections of the various fan-out trees and fan-in trees (equivalently, at the output leaves of each fan-out tree). A crosstalk reduction device has only two states. It either passes or blocks the optical signal, based on the voltage status of a control lead.

We know there are a total of $N \times M$ crosstalk reduction devices required for an N-input by M-output switching array. Thus, for any large tree architecture network, the number of control voltage signals for the crosstalk reduction devices can potentially impose significant burdens on both chip real estate and wiring patterns. We will demonstrate that it is possible to gang the control of the crosstalk reduction devices, such that the smaller of N or M control voltages will suffice to control the entire collection of $N \times M$ devices.

4.2.5 Organizing the Control of Crosstalk Reduction Devices

If we are to be able to control all of the crosstalk reduction devices of N $1 \times M$ fan-out trees with only N control voltages, we will have to be able to control all of the M devices of each tree with a single control voltage. Consider once again the 1×32 fan-out tree in Fig. 4.6. We wish to control all 32 crosstalk reduction devices for this tree with a single control voltage. For any particular connection, we only care about the devices carrying level-0 (active) and level-1 (crosstalk) signals. We need to be able to pass the level-0 signal while simultaneously blocking the 5 level-1 signals, and we do not care whether we block or pass signals of level-2 and beyond.

For this to be possible, we must be able to partition the 32 crosstalk reduction devices into two groups, A and B, such that when all devices in Group A are in the block state, all devices in Group B are in the pass state, and vice-versa. Referring to the 1×32 fan-out tree in Fig. 4.4, let us arbitrarily assume that the level-0 leaf is in Group B. Then, the level-1 leaves will have to be in Group A, for if we pass the level-0 signal, we must block all of the level-1 signals. Now the Group A (Group B) assignment is static for any particular

fan-out tree. Each specific connection we choose will result in a particular level-0 signal leaf and five level-1 signal leaves. In every case, the level-0 signal leaf must be in one group and the five level-1 signal leaves must all be in the other group.

If a valid partition exists, it must exhibit the following attribute. Any connection to a leaf in Group A generates level-1 signals only at leaves in Group B and, conversely, any connection to a leaf in Group B generates level-1 signals only at leaves in Group A. Does such a partition exist? The answer is yes, and we will now proceed to explain why and indicate how to find such a partition.

4.2.6 Destination Addresses of Level-0 and Level-1 Signals

In a fan-out tree composed of stages of 1×2 switches, there is a natural way to uniquely label (give destination addresses to) each of the leaves of the tree. A label corresponds to the route followed from the input on the left side of the tree to a particular leaf on the right side of the tree. By convention, in a 1×2 switch, we use the name "output 0" for the upper output port and "output 1" for the lower output port.

Consider the routing of a particular input signal to some output leaf in a 1×32 fan-out tree. We introduce the notion of a binary address label $S_1S_2S_3S_4S_5$. Each S_i in this label represents a binary digit having either the value 0 or 1. This label is used to route an input signal as follows. If S_i has a value of 0, the signal is routed to the upper port (port 0) of the 1×2 switch in the i^{th} stage. Similarly, if S_i has a value of 1, the signal is routed to the lower port (port 1) of the 1×2 switch in the i^{th} stage. After the sequential routing according to S_1 through S_5, we arrive at a leaf that we will simply name $S_1S_2S_3S_4S_5$. Thus, the destination address (or leaf label) corresponds directly to the route that was followed to reach the leaf from the input node.

The 1×32 fan-out tree in Fig. 4.4 exemplifies this address scheme. The leftmost (first) stage in this figure contains a single 1×2 switch. The next stage to the right contains two 1×2 switches: one labeled 0 and one labeled 1. These labels correspond, respectively, to the routing taken from the first-stage switch to reach these second-stage switches; i.e., the upper port (port 0) on the first-stage switch is connected to switch 0 in the second stage and, similarly, the lower port (port 1) on the first-stage switch is connected to switch 1 in the second stage. This labeling process continues in a stage-by-stage fashion as we move from left to right. For example, port 0(1) on switch 0 in the second stage connects to switch 00(01) in the second stage, and port 0(1) on switch 1 in the second stage connects to switch 10(11) in the second stage.

Continuing this process, we find that the label of a switch in a given stage is constructed by first writing the label of the switch in the previous stage to which it is connected and then adjoining a single bit (0 or 1) to the right end of said label. A 0 (or 1) is adjoined if the path from the switch in the previous stage was via the upper (or lower) port of that switch. Thus, the labels of the

switches in each stage increase in size by one bit from the previous stage. As can be observed in Fig. 4.4, each of the 32 output leaves is labeled with a unique 5-bit label.

Now consider an input signal routed to $B_1B_2B_3B_4B_5$. This routed signal generates level-1 crosstalk at the stage-one 1×2 switch on port $\overline{B}_1$, where $\overline{X}$ denotes the binary complement of X, i.e., $\overline{X} = 0$ if $X = 1$, and vice versa. Because of the ganged control for each stage, this level-1 crosstalk will be subsequently routed in stages 2 through 5 on output ports B_2 through B_5, respectively. Thus, this level-1 crosstalk arrives at output leaf $\overline{B}_1B_2B_3B_4B_5$.

The input signal will also generate level-1 crosstalk in stage 2, at a 1×2 switch on output port $\overline{B}_1$. Now since this crosstalk was not generated until stage 2, it was routed identically to the active signal in stage 1. And because of ganged control for each stage, this level-1 crosstalk will be routed in stages 3 through 5 to output ports B_3 through B_5, respectively. Thus, this level-1 crosstalk arrives at output leaf $B_1\overline{B}_2B_3B_4B_5$.

In general, an input signal will generate level-1 crosstalk at each stage. The routing for the crosstalk is identical to that for the input signal except for the stage at which the crosstalk is generated. For the stages before the one at which the crosstalk is generated, this is true because the eventual crosstalk is still part of the input signal. For the stages after the crosstalk is generated, the

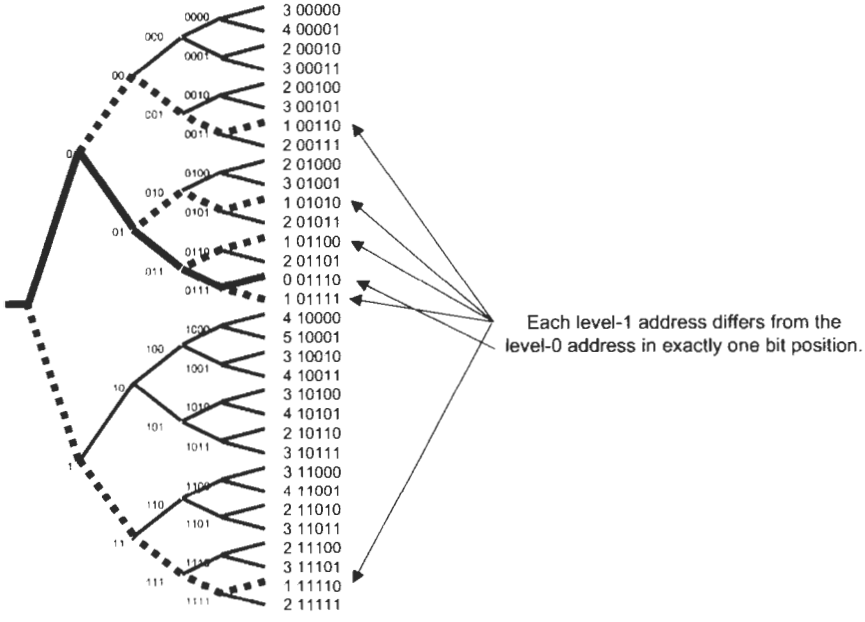


Fig. 4.4 Comparison of destination addresses for a routed signal and level-1 crosstalk.

crosstalk is routed identically to the input signal because of the ganged control for each stage.

In an S-stage fan-out tree, there will be S level-1 crosstalk signals generated. If $S_1 \ldots S_s$ represents the output leaf address of the input signal, then the S level-1 crosstalk signals will appear at S output leaf addresses given by complementing one of the S_i address bits. Thus, the leaf address for each level-1 crosstalk signal differs in exactly one bit position from the input signal destination leaf address.

4.2.7 Group Assignments for Output Leaf Destination Addresses

We are now in a position to determine how to partition leaf addresses into two groups, A and B. We shall use the 1×32 fan-out tree as an example in our approach. The generalization from this example will be straightforward.

As described in the previous section, we will identify the 32 output leaves with 5-bit address labels. Thus, the 32 leaves will have labels 00000, 00001, ..., 11111. Let us arbitrarily choose 01110 as a destination address for an input signal. According to the previous section, we can produce the five addresses where level-1 crosstalk will emerge, by successively complementing single bits in the address 01110. The resultant addresses are 11110, 00110, 01010, 01100, and 01111.

Let us arbitrarily assign address 01110 to Group A. Then the other five addresses (11110, 00110, 01010, 01100, and 01111) must all be in Group B. We know that each address in Group B differs in exactly one bit position from the address in Group A. Now consider what happens when one of the addresses in Group B becomes the destination address for an active signal. This Group B address now generates five additional level-1 crosstalk addresses, all of which must be in Group A. For example, Group B address 11110 will generate the following five Group A addresses; 01110, 10110, 11010, 11100, 11111.

Notice that one of these addresses (01110) is the same as the original Group A address that we started with. This occurs because 11110 and 01110 generate each other when their leftmost bits are complemented.

The other four Group B addresses will each generate five Group A addresses. Including the first Group B Address 11110, we have:

Group B → Group A
11110 → 01110, 10110, 11010, 11100, 11111
00110 → 10110, 01110, 00010, 00100, 00111
01010 → 11010, 00010, 01110, 01000, 01011
01100 → 11100, 00100, 01000, 01110, 01101
01111 → 11111, 00111, 01011, 01101, 01110

Let us examine these addresses we have for Groups A and B to see if any patterns emerge.

Our original Group A address was 01110. It contains three 1s. In the above 15 Group A address, it can verified that they contain either one, three, or five 1s. Why do they all have an odd number of 1s? The reason is easily understood by considering the method that generates the Group A and Group B addresses.

For a given Group A address, Group B addresses are generated by complementing exactly one bit of the Group A address. If this process changes a 0 to a 1, the number of 1s in the Group B address will be one larger than in the Group A address. If the process changes a 1 to a 0, the number of 1s in the Group B address will be one smaller than in the Group A address. Either way, if the original Group A address had an even number of 1s, the generated Group B addresses will have an odd number of 1s. And if the original Group A address had an odd number of 1s, the generated Group B addresses will have an even number of 1s.

Now, when the generated Group B addresses are used to subsequently generate Group A addresses, a similar result will occur, i.e., a Group B address with an odd (even) number of 1s will generate Group A addresses with an even (odd) number of 1s. And so the overall result will be that the parity of the number of 1s (or 0s) will be different for the Group A and Group B addresses.

Thus, we have the following assignment rule for placing addresses in Group A or Group B: arbitrarily assign all of the addresses with an even number of 1s (or 0s) to one group and assign all of the addresses with an odd number of 1s (or 0s) to the other group. Let us now examine this assignment policy as regards our ability to control all of the crosstalk reduction devices of a given fan-out tree with a single voltage.

4.2.8 The Two Groups and Their Control Implications

The implication of a single control voltage for controlling both Group A and Group B is that when all of the Group A devices are in the pass state, all of the Group B devices are in the block state, and vice versa. Thus, a single voltage will suffice only if any signal routed to a Group A address generates level-1 crosstalk only at Group B addresses, and conversely, any signal routed to a Group B address generates level-1 crosstalk only at Group A addresses.

This will be true because the parity of the number of 1s (0s) in the Group A addresses is different than the parity of the number of 1s (0s) in the Group B addresses. And we demonstrated in the above section that any routed signal generates level-1 crosstalk only at addresses that have one more or one fewer 1 (or 0) than the active signal destination address. Thus, a single voltage control can indeed simultaneously pass the routed signal and block all S of the associated level-1 crosstalk signals in a $1 \times M$ (for $M = 2^s$) fan-out tree. This means that N control signals (one for each fan-out tree) will suffice to control level-1 crosstalk in the network.

Earlier we asserted that the number of required control voltages was the smaller of N or M. This is because we can also control crosstalk from the perspective of the fan-in trees. We now investigate this approach.

4.2.9 Crosstalk Control—Fan-In (Output) Tree Perspective

For this discussion, we still assume that there are $N \times M$ crosstalk reduction devices placed at the intersections of the N $1 \times M$ fan-out trees and the M $N \times 1$ fan-in trees. What is different is that we now will organize the control of the devices according to fan-in tree associations, instead of fan-out tree associations. To facilitate this analysis, we concern ourselves with a different view of crosstalk propagation.

Consider the 32×1 fan-in tree in Fig. 4.5. Except for the routed active signal, let us assume that all paths from the left contain level-1 crosstalk signals. In most networks, this is not possible. Many of the paths from the left will contain level-2, -3, etc. crosstalk signals. However, if we can control crosstalk for a "beyond worst-case" assumption, we will certainly be able to control it for the less demanding possible situations.

Since the crosstalk reduction devices are no longer being organized and controlled from the perspective of the input (fan-out) trees, level-1 crosstalk

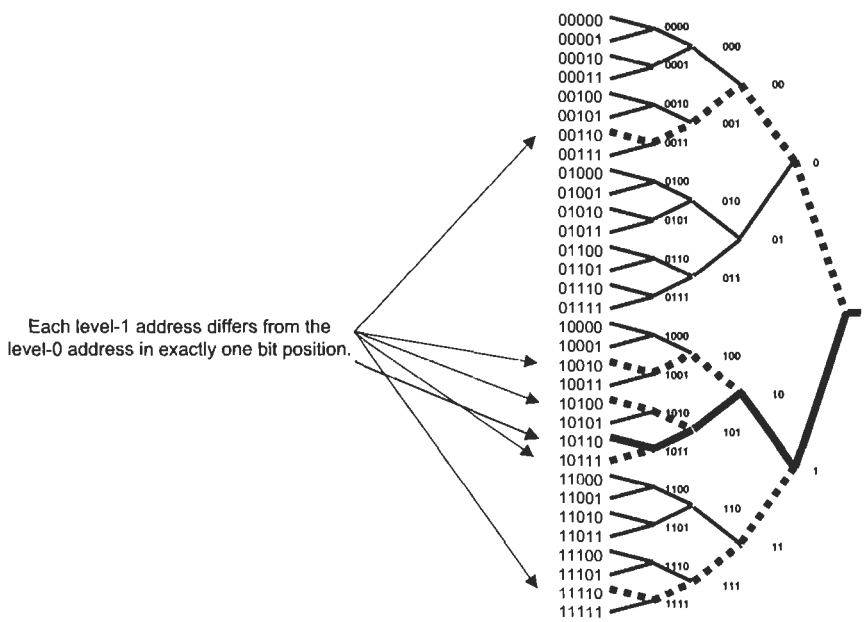


Fig. 4.5 Organizing the control of crosstalk reduction devices from the perspective of the output tree.

may be arriving at the fan-in tree on any path. Some of these level-1 crosstalk signals must be controlled, and some may be ignored. Let us examine why this is true and see which level-1 crosstalk signals are of concern.

4.2.10 Effect of Ganged Per-stage Control

Recall that our requirement for crosstalk control is that all level-1 crosstalk be reduced to level-2 or less before reaching a 2×1 switch containing an active signal. This guarantees that the crosstalk will be reduced to level-3 or less before it combines with an active signal. The crosstalk signals on many input leaves will be knocked down at least one additional time before reaching a 2×1 switch carrying an active signal. But some crosstalk signals will not be further knocked down before reaching a 2×1 switch carrying an active signal. We need to be able to identify those that are further knocked down and those that are not.

In Fig. 4.5, each of the 32 input leaves is labeled based on the routing from that leaf to the outlet on the right side of the tree. At each stage in the tree, a route is characterized by whether the path enters the upper port (0) or lower port (1) of the encountered 2×1 switch. We will adopt the convention that the rightmost bit of the address label corresponds to the first switch encountered moving from right to left, the second bit from the right corresponds to the next switch encountered, and so on. Thus, for example, the leaf labeled 10110 implies that the route from that leaf to the outlet of the tree is upper port (0), lower port (1), lower port (1), upper port (0), lower port (1).

In general, a leaf labeled $B_1B_2B_3B_4B_5$ will first encounter a B_5 port, then a B_4 port, then a B_3 port, and so forth. In a fashion similar to the labeling in the fan-out tree, this scheme also produces labels for the switches in the stages between the input leaves and the outlet of the tree. Employing this concept, we note that leaf labeled $B_1B_2B_3B_4B_5$ will first encounter port B_5 on switch $B_1B_2B_3B_4$, then port B_4 on switch $B_1B_2B_3$, then port B_3 on switch B_1B_2, then port B_2 on switch B_1, and finally port B_1 on the 2×1 switch connected to the outlet of the tree.

This means that if two leaves have identical address bits B_1 through B_i, they will share switches B_1, B_1B_2, $B_1B_2B_3$, ..., through $B_1B_2 \ldots B_i$. Let us first consider two leaves whose labels differ only in bit position 5; i.e., leaves labeled $B_1B_2B_3B_4B_5$ and $B_1B_2B_3B_4\overline{B}_5$. These two leaves are connected to input ports B_5 and $\overline{B}_5$, respectively, on switch $B_1B_2B_3B_4$. So, if $B_1B_2B_3B_4B_5$ is carrying the active signal, then $B_1B_2B_3B_4\overline{B}_5$ will have to be receiving its signal from a crosstalk reduction device that is in the blocking state, to insure that no more than level-2 crosstalk is carried by $B_1B_2B_3B_4\overline{B}_5$.

Now, because of ganged control, all of the switches in a given stage of the tree will be in the same switching state as that used for the routed active signal. Suppose the active signal appears on leaf $B_1B_2B_3B_4B_5$. We have already

noted that leaf $B_1B_2B_3B_4\overline{B}_5$ shares switch $B_1B_2B_3B_4$ with $B_1B_2B_3B_4B_5$. These two leaves also share access to switch $B_1B_2B_3$ as do leaves $B_1B_2B_3\overline{B}_4B_5$ and $B_1B_2B_3\overline{B}_4\overline{B}_5$. What happens to the crosstalk carried by these latter two leaves?

The routed active signal is accepted (passed) on ports B_5 and B_4 on the first two switches it encounters. Because of ganged control, this means that the crosstalk on $B_1B_2B_3\overline{B}_4B_5$ will be passed on port B_5 on the first switch it encounters and will be blocked (knocked down) on port $\overline{B}_4$ on the second switch it encounters. And the crosstalk on $B_1B_2B_3\overline{B}_4\overline{B}_5$ will be blocked on both ports $\overline{B}_5$ and $\overline{B}_4$ on the first two switches it encounters. This means that the crosstalk on leaf $B_1B_2B_3\overline{B}_4B_5$ will be knocked down only once before combining with the active signal on switch $B_1B_2B_3$, whereas the crosstalk on leaf $B_1B_2B_3\overline{B}_4\overline{B}_5$ will be knocked down twice before combining with the active signal on switch $B_1B_2B_3$. So the crosstalk on $B_1B_2B_3\overline{B}_4\overline{B}_5$ does not have to be controlled with a crosstalk reduction device, but the crosstalk on $B_1B_2B_3\overline{B}_4B_5$ will have to be knocked down (blocked) by a crosstalk reduction device to insure that level-1 crosstalk on $B_1B_2B_3\overline{B}_4B_5$ will be reduced to level-3 before combining with the active signal.

4.2.11 Partitioning into Two Groups

This logic can be extended, and the conclusion is that the leaves that must have their crosstalk knocked down by a crosstalk reduction device are those whose addresses differ in exactly one bit position from the address of the leaf with the active signal. The key observation is that (because of ganged control) the crosstalk signal on a given input leaf will be additionally knocked down a number of times equal to the number of bit positions in which its address differs from the address of the active signal leaf. This is because those bit positions represent input ports on 2×1 switches that will be set to accept a signal on their other ports and hence block (knock down) the crosstalk signal.

Our assumptions about crosstalk propagation, why certain leaves need to be controlled, and why others do not are not the same for fan-in trees as they are for fan-out trees. But the results are similar. We have determined that the leaves that need to have their crosstalk controlled are identically characterized for both fan-out and fan-in trees. Thus, the logic applied in partitioning the leaves into two groups for the fan-out trees will apply identically for the fan-in trees. We repeat the rule here for completeness.

Arbitrarily assign all of the addresses with an even number of 1s (or 0s) to one group and assign all of the addresses with an odd number of 1s (or 0s) to the other group.

One should be careful to note that the addresses referred to in the two cases are not the same addresses. When considering fan-out trees, the addresses refer to output leaf addresses on fan-out trees. And when considering fan-in

trees, the addresses refer to input leaf addresses on fan-in trees. Furthermore, the respective sets of crosstalk devices controlled with single voltage pairs are different for the two cases. For the fan-in (fan-out) tree perspective, the set of crosstalk reduction devices controlled with a single voltage pair are those connected to a particular fan-in (fan-out) tree.

4.2.12 Concluding Remarks

We can add a few observations that tend to generalize the results presented here. The common thread will be the use of a tree architecture, but the parameters of the problem can be varied.

Our primary examples utilized trees whose numbers of leaves equaled powers of two. This does not have to be the case. One simply views a tree with an arbitrary number of leaves as a power-of-two tree with some unequipped leaves and branches. This does not change any of the logic employed in reaching our main results.

We were very specific in identifying how switch control was ganged in our discussions. We assigned the label 0 to upper ports and the label 1 to lower ports. The upper and lower designations do not actually matter. As long as we know which ports pass and which ports block signals for a given control voltage state, we will be able to determine how to control the network and partition the crosstalk reduction devices into controllable groups.

An $N \times N$ tree architecture network will contain $2N(N-1)1 \times 2$ switches plus N^2 crosstalk reduction devices, resulting in a total of $3N^2 - 2N$ active elements to be controlled. Such a network will have $2 \log_2 N$ stages of switching elements and one stage of crosstalk reduction devices. Because of ganged control, each of these $2 \log_2 N + 1$ stages can be controlled with N voltage pairs. Thus a total of $2N \log_2 N + N$ control voltage pairs will suffice to control $3N^2 - 2N$ active elements. This dramatic reduction in the number of control voltages, as compared to number of active elements, tends to increase the maximum size of switching arrays that can be fabricated.

We have already noted that the crosstalk reduction devices may be grouped and partitioned from either the input or output perspective. This provides additional flexibility in the design of the switch control. If the number of inlets and outlets are equal, the total number of required control voltages is the same for either choice. But if the network has an unequal number of inlets and outlets, then the number of control voltages required for crosstalk reduction can be chosen as the lesser of the number of inlets or outlets. This can provide further savings and flexibility.

Finally, it should be mentioned that the control algorithms to be used in such tree-structure architectures can be straightforward and quite elegant. This facilitates the potential for very high-speed control of such networks.

4.3 OPTICAL CROSS-CONNECT (OXC)

4.3.1 Introduction

The OXC will play significant roles in future optical networks due to an evolving need for lower cost and more flexible network architectures. As networks evolve from point-to-point to rings to eventually mesh type architectures, OXCs will enable optical channel bandwidth management functions through wavelength grooming, automatic rapid provisioning and faster restoration, since switching will take place at the wavelength level.

One of the important OXC function is performing restoration, specifically mesh type restoration. Typically, networks based on a mesh topology are inherently more flexible and less costly than ring-based designs because they can grow more easily in previously unplanned ways by growing mesh segments compared to ring planning. In addition, mesh core typically saves 30–60% of restoration bandwidth compared to rings, where 100% spare capacity is needed. The OXC can be used in a mesh configuration for more efficient use of bandwidth and to achieve scalable optical networks.

To support these functions, we have explored various OXC architectures in this section. The optimum architecture will be very much dependent on the application and network needs. Optical cross connects of all shapes and sizes will play a critical part in building complex optical networks.

4.3.2 Wavelength Selective Cross-Connect (WSXC)

The WSXC accepts multiwavelength signals from a transport facility, and may also accept single wavelength signals at the client network interfaces. It demultiplexes multiwavelength signals into single wavelength signal and cross-connects individual wavelengths without wavelength interchange. Typically, the WSXC accepts a specific set of wavelengths, and does not allow arbitrary cross-connections of different wavelengths. This makes it a blocking cross-connect, but with respect to all specific wavelengths that it is designed to accept, it is strictly nonblocking. As illustrated in Fig. 4.6, the number of optical-switch-units or the number of switch layers used in this type of OXC corresponds proportionately to the number of WDM wavelengths carried on each fiber. In addition, the size of the switch at any given layer is proportional to the number of input fibers carrying WDM traffic. As an example, a switch of an $M \times M$ size will be required at each layer for M number of input fibers. This architecture is typically being used for the small port count OXC making 2-D based small optical switches as a preferred technology.

A WSXC approach using 4×4 $LiNbO_3$ modules has been demonstrated for a 32×32 OXC in 1997, and using 6×6 $LiNbO_3$ modules has been demonstrated for a 48×48 OXC in 1999 [Johnson, 1999].

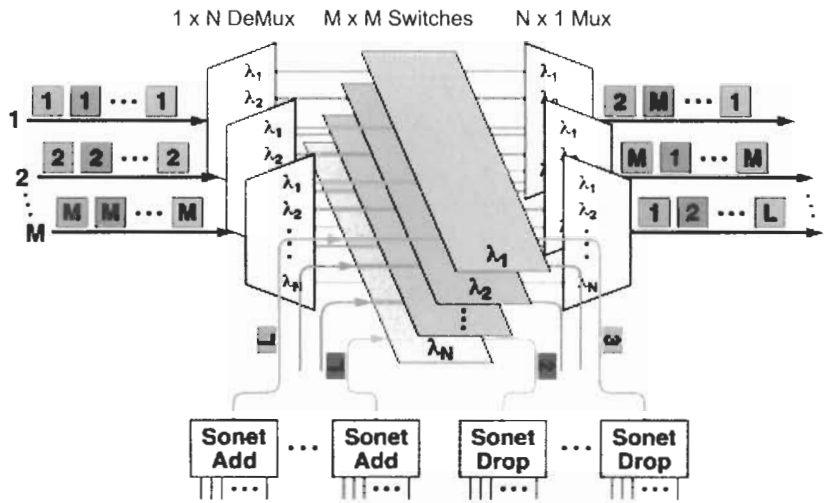


Fig. 4.6 A typical wavelength selective cross-connect.

4.3.3 Strictly Nonblocking Cross-Connects

The nonblocking cross-connects may be subdivided into two categories: rearrangeably nonblocking (RNB) and strictly nonblocking (SSNB). The distinction between these two categories is important from a practical perspective, since a rearrangeably cross-connect may require reconfiguration of an old connection to establish a new one. Compared to that, strictly nonblocking cross-connects have cross-connect fabric that is strictly nonblocking for all (one-way or two-way) point-to-point cross-connections. The term nonblocking means that a cross-connection request will not be denied due to lack of a path through the fabric, when the desired input and output ports are available. The term strictly means that there is always a path from any idle input to any idle output, and it is not necessary to do any rearrangement of paths in order to achieve this capability.

4.3.3.1 Multi-stage Architectures

The multi-stage architecture consists of a collection of basic switch elements that are connected together in a particular topology in order to build a larger switch fabric. The three-stage Clos architecture shown in Fig. 4.7(a) is a common multistage crossbar topology. As shown, m is the number of output ports and n is the number of input ports on the input chip. N is the total number of ports into the fabric. This architecture is rearrangeably nonblocking if $m = n$ and strictly nonblocking, if $m \geq 2n - 1$.

In the multistage Clos architecture, the first stage, an input stage requires the output bandwidth that is approximately twice its input bandwidth.

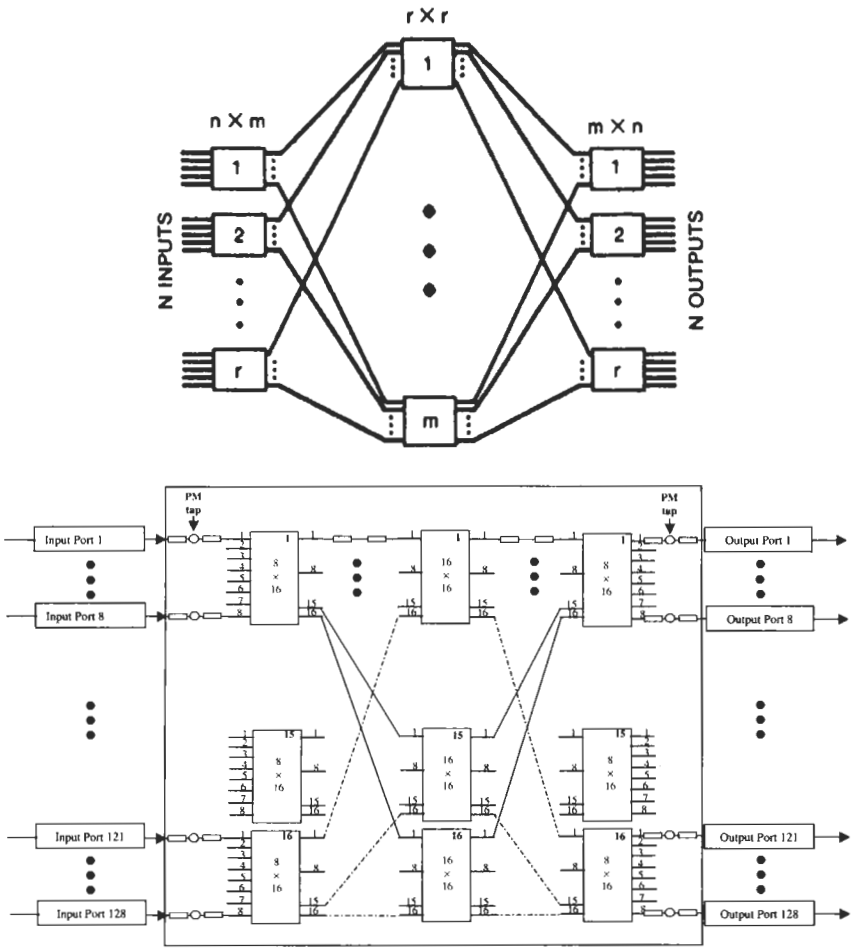


Fig. 4.7 (a) Three-stage Clos architecture. (b) Multistage Clos OXC with 128 × 128 port count.

Conversely, twice as many input ports are required on the final stage, an output stage. In addition, the number of crossbars in the center stage will also be equal to m.

A multiple $M \times N$ building block modules will be required to be implemented in the multiple stages to realize switch fabrics of port count 128×128, 256×256, 512×512 or 1024×1024. For example, a 128×128 OXC would be constructed in a three-stage Clos architecture from sixteen 8×16 switches, sixteen 16×16 switches, and sixteen 16×8 switches, as shown in Fig. 4.7(b). Likewise, to realize an OXC of size 512×512 using three stages would require thirty-two 16×32 switches, thirty-two 32×32 switches, and thirty-two 32×16 switches. Building a Clos-based switch fabric of 1024×1024 ports size would

require a crossbar chip of at least 64 × 64. This size of OXC will require a total of 192 64 × 64 switches. This is double the number of chips and the switch module size compared to the 512 × 512 OXC architecture. There is a fundamental trade-off between the size of the basic switching elements and the amount of interconnect required between switching elements to produce a multi-stage nonblocking fabric. The larger the basic switching element, the less interconnect required. In any case, an additional number of wires needed on the backplane to interconnect stages of the switch fabric will make Clos-based switches of more than 1024 ports impractical.

4.3.3.2 Single-stage Architectures

The single-stage architecture shown in Fig. 4.8(a) has a switching fabric that can interconnect signals between two ports. The signals from any input port can be connected to any output port in a single-stage to establish a cross-connection.

The switch used to construct a large all-optical switching fabric within a single stage is typically based on the "analog" or "three-dimensional (3D)"

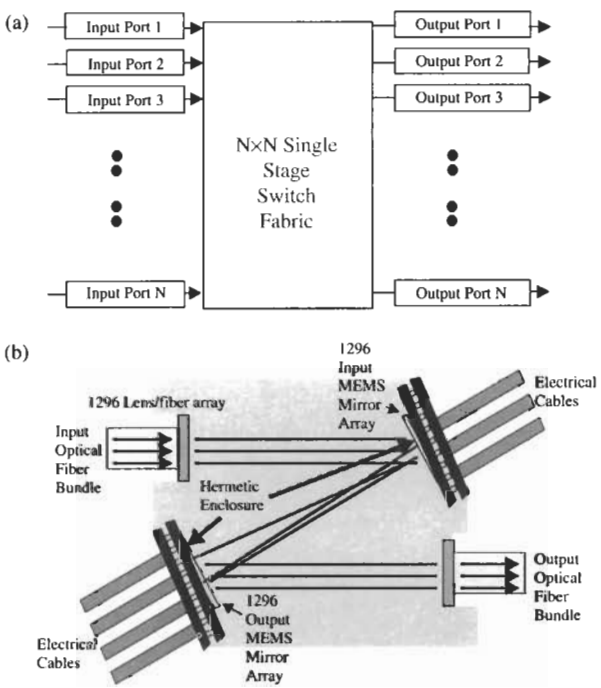


Fig. 4.8 (a) Single-stage architecture; (b) Optical layout of 1296 × 1296 port switch fabric.

approaches, since it scales well into very large port counts and offers an advantage of compact single-stage fabric. In the 3D architecture, the number of switch elements scales as $2N$, where two arrays on N mirrors are used to connect N input to N output ports, since each port-to-port connection undergoes two reflections.

A specific example may help to illustrate this point. Figure 4.8(b) illustrates a large single-stage 1296 × 1296 cross-connect, where an optical switching is performed using MEMS (Micro Electro Mechanical System) technology. The MEMS fabric is housed as an array of microscopic mirrors in the 3D architecture [Ryf, 2001]. As shown in Fig. 4.8(b), each single-crystal silicon MEMS mirror array is comprised of $36 \times 36 = 1296$ mirrors and is sealed behind glass windows in a ceramic package, and within a hermetic enclosure. Since the micro-mirror must have multiple possible positions to steer lightwave signals from each input port to N output ports, each mirror is rotated around micromachined hinges, which allows 2-axis tilt motion. In the example shown here, a large 1296 × 1296 cross-connect build with a 3D design requires only $2N$, 2,592 mirrors compared to a 2D design requiring N^2 mirrors. The 3D design minimizes the number of switch elements and the number of interconnects compared to a 2D design with N^2 switch element. The MEMS technology is explained in more detail in the technology section and illustrated in the Fig. 3.6.

The single-stage architecture build with a 3D design allows a large port count with low loss and good loss uniformity, making it highly desirable for optical architectures. In 2000, a number of vendors realized that 3D MEMS were essential in building large port-count OXCs, and made 3D MEMS a preferred choice.

4.3.4 Prototypes and Performance

An approach using servo-controlled beam steering has been prototyped for sizes up to a partially populated 576 × 576 cross-connect [Lee, 2000]. Two frames, each forming a 24 × 24 grid, are placed parallel to each other as opposite faces of a long empty box whose other four sides are solid; this box and the two end frames are collectively known as the optical frame assembly. An optical cluster, consisting of a collimator and the two servomotors that control the collimator's orientation, is mounted at each grid position on the two frames, with the fibers extending away from the cage. Each two-servomotor unit permits independent steering of the beam in orthogonal degrees of freedom.

The overall length of the unit is dictated by the range of motion of each optical cluster. The clusters are oriented so as to minimize the length of the assembly; clusters at the center of the frame are aligned with the length of the box, while those at the corners are angled toward the center of the assembly. The width of the unit is determined by the widths of the clusters themselves.

The resulting footprint for the early 576 × 576 assembly was 20 inches by 20 inches by 62 inches. Newer generations of servomotors have allowed the cluster width to decrease, however, so a 1024 × 1024 cross-connect could occupy the same footprint.

Performance results show an excellent insertion loss of 0.7 dB, as would be expected due to the limited number of components involved, especially the absence of mirrors. The isolation, measured at 100 dB, is also extremely good. These values, as well as the return loss of −70 dB and the PDL of 0.02 dB, are expected to be unchanged with the transition to larger port counts.

The switch speed, however, is measured to be 1.6 seconds for the worst-case scenario of reorienting from a port on one edge of the frame to a port on the far edge of the frame. Because the servomechanism operates independently in each degree of freedom, there is no additional delay for reorienting from one corner to the opposite corner; rather, the switching time is determined by the larger orthogonal component of the distance to be traveled. Future generations with faster servomechanisms are expected to bring the worst-case switching time down to 100 ms, but this is still quite long. As one would expect, the servomechanisms reorient much more slowly than MEMS, due to their greater inertia. The mean switching speed for random port reconnections is 0.47 times that of the worst-case scenario, though intelligent allocation of ports could be used to reduce the number even further.

The beam-steering approach pursued by Astarte Fiber Networks and Texas Instruments [Laor, 1999] is similar in many ways, though it is moveable micromirrors, operated by a closed-loop servo control system, that redirect each beam of light independently in two orthogonal axes. A fixed plane mirror in each module is used to fold the light path. Many of the details, such as the use of two fibers per optical module, are similar to those in the macroscopic beam-steering approach discussed above.

This method is also prototyped as a 576 × 576 OXC, permitting a ready comparison. The insertion loss is several times higher, 4.8 dB, as expected due to the use of micromirrors and beam-folding mirrors. The crosstalk is very good, generally less than −80 dB. The worst-case switch time is 10 ms, though by the same reasoning as given above, the mean switch time would be less than half that number. The use of micromirrors and beam-folding mirrors leads to a much smaller footprint.

The first large port count MEMS prototype, a fully provisioned 112 × 112 cross-connect, was created at Lucent Technologies [Neilson, 2000]. The micromirrors are electrostatically actuated and are trained for all 12,544 possible connections. As in the previous case, a fixed plane mirror is used to fold the light path and reduce footprint, yielding a very small switch fabric even for large port counts.

Typical switch times of 5–10 ms are reported, making it a strong candidate for fast restoration applications. Its insertion loss, 7.5 ± 2.5 dB, is the highest among the free-space beam-steering prototypes discussed here; but like

those other approaches, the insertion loss should not increase as port count is increased. The crosstalk is measured to be below −50 dB.

An additional test performed on this prototype highlights the advantages of all-optical cross-connects in handling high bit rates. Two 160 Gb/s signals are time and polarization interleaved to generate a 320 Gb/s TDM signal, which is then passed through a 1 : 112 passive splitter, allowing the same traffic to be passed through each input and output port. The resulting aggregate capacity of 35.8 Tb/s is passed through the cross-connect simultaneously.

An even larger single-stage cross-connect, 1296 × 1296 as shown in Fig. 4.8(b), has been produced using the same MEMS technology [Ryf, 2001]. The insertion loss, 5.1 ± 1.1 dB, has been reduced, while the switching time remains on the order of 5 ms. An aggregate capacity of 2.07 Pb/s has been obtained by sending forty 40 Gb/s DWDM channels into each port. This further illustrates the promise of such devices for switching extremely large quantities of data.

4.3.5 Mesh Restoration Demonstrations

The servo-controlled beam-steering approach has been used as the switch technology in one of the earliest distributed provisioning and restoration demonstrations using an all-optical cross-connect. A single cross-connect is partitioned to represent three independent nodes in a triangular topology. A pseudorandom bit stream is used for the SONET data transmitted from the source node to the destination node, initially along the link directly connecting them. When that link is subjected to a simulated fiber cut by the use of an automatic attenuator, the loss of signal is detected, initiating restoration along a path using the third node as an intermediate node. This demonstration shows that the servo-controlled beam-steering approach is viable for basic applications, though the restoration time is slow, as expected. The dominant contribution to the restoration time is the servomechanical switch time itself, rather than ancillary considerations such as software.

A similar but more elaborate demonstration has been performed using a 128 × 128 prototype of the Lucent MEMS switch fabric [Agrawal, 2000]. A four-node mesh network with five bidirectional links is constructed using a combination of Lucent 40 G™ and 400G™ optical line systems, supporting 16 and 80 wavelengths, respectively. The network is provisioned with bidirectional 2.5 Gb/s demands between every pair of nodes. The four OXCs required for this topology, two 11 × 11 and two 9 × 9, are obtained by partitioning the MEMS prototype; a total of 40 input and 40 output ports are therefore used.

The system architecture for a typical node is illustrated in Fig. 4.9. WDM signals from adjacent nodes are demultiplexed and regenerated by OTUs

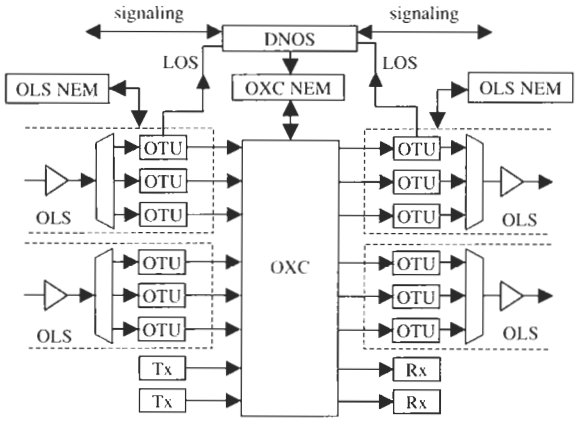


Fig. 4.9 Node architecture in an optical mesh network with distributed restoration.

before entering the OXC. The OTUs provide wavelength translation, signal regeneration, and performance monitoring. Similarly, the OXC output is passed to the OTUs, multiplexed, then transmitted to the adjacent nodes. When a fault occurs on a particular channel, the loss of signal is reported to the appropriate distributed network operating system (DNOS). A 10/100BASE-T Ethernet signaling channel is used to communicate between adjacent nodes. A fault is introduced by a mechanical switch on a link carrying multiple demands, and restoration is initiated collectively by copies of the DNOS running on all relevant nodes. A SONET test set, generating a pseudorandom bit stream, is used to measure the service disruption time between fault and restoration. No bit errors are observed before the fault or after restoration. The restoration time is measured to be 41 ± 1 ms; this restoration time would be about 15 ms faster due to parallel switching if multiple cross-connects were used rather than partitions of a single cross-connect.

Centralized restoration in a long-haul mesh network, with one 256×256 Lucent MEMS LambdaRouter™ switch partitioned into five logical nodes through the use of a software translator, has been publicly demonstrated in Shin Yokohama, Japan, immediately following the SubOptic Conference in May 2001 [Yamamoto, 2001]. Thousands of kilometers of cable are used for the transmission infrastructure between nodes, both on KDD-SCS submarine links and on Lucent Technologies terrestrial links. Five demands, representing platinum and bronze levels of service at OC-192 and OC-48 bit rates, are provisioned using a centralized operating system; all of the platinum trails undergo automatic restoration in response to arbitrary single-wavelength and cable cut hardware faults.

Abbreviations, Acronyms, and Terms

A/D	Add/drop
ADM	Add/drop multiplexer
AIS	Alarm indication signal
ATM	Asynchronous transfer mode
BER	Bit-error rate
BWM	Bandwidth manager
CO	Central office
DACS	Digital access cross-connect system
DARPA	Defense advanced research projects agency consortium
DCS	Digital cross-connect system
DS-3	Digital signal level 3 with transmission rate of 44.74 Mb/s
DWDM	Dense wavelength division multiplexing
EDFA	Erbium-doped fiber amplifier
ELXC	Electronic layer cross-connect
ESF	Electrical switch fabric
FAS	Fiber-array switch
FBS	Fiber-bundle switch
FEC	Forward error correction
FLC	Ferroelectric liquid crystal
FSBS	Free-space beam-steering
IO	Integrated optic
IP	Internet protocol
IXC	Interexchange carrier
LC	Liquid crystal
LN or $LiNbO_3$	Lithium niobate
LOF	Loss of frame
LOS	Loss of signal
MONET	Multiwavelength optical networking
OADM	Optical add/drop multiplexer
OAM&P	Operations, administration, maintenance, and provisioning
OC-N	Optical carrier digital signal rate of $N \times 51.84$ Mb/s
OEO	Optical-to-electrical-to-optical
OLS	Optical line system
OLXC	Optical layer cross-connect
OPB	Optical power budget
OPU	Optical port unit
OSF	Optical switch fabric
OSNRs	Optical signal-to-noise ratios
OTU	Optical translator unit

OXC	Optical cross-connect
PBX	Private branch exchange
PDL	Polarization-dependent loss
PIC	Photonic integrated circuit
PMD	Polarization mode dispersion
SDH	Synchronous digital hierarchy
SOA	Semiconductor optical amplifier
SONET	Synchronous optical network
SLM	Spatial light modulator
STS-N	Synchronous transfer signal of rate $N \times 51.84$ Mb/s
TDL	Temperature-dependent loss
TDM	Time-division multiplexing
UDWDM	Ultra-DWDM
WDL	Wavelength-dependent loss
WDM	Wavelength-division multiplexing
WSXC	Wavelength-selective cross-connect

Symbols

A	Number of add/drop I/O ports on a cross-connect
δ	Degree of a node
D	Diameter of a network
h	Number of hops between nodes in a network
I/O	Input/output
L	Number of links (edges) in a network graph
M	Number of I/O ports on a cross-connect
N	Number of nodes in a network graph

References

Agilent Photonic Switching Platform: N3565A 32 × 32 Photonic Switch, technical specification, http://www.agilent.com/

G. Agrawal, *Fiber-Optic Communication Systems*, John Wiley and Sons, 1997.

N. Agrawal, et al., "Demonstration of distributed mesh restoration and auto-provisioning in a WDM network using a large optical cross-connect," Optical Fiber Communication Conference, paper PD37, 2000.

C.A. Alegria, S.S. Bergstein, K.A. Dixson, and C.J. Hunt, "The WaveStar™ bandwidth manager: The key building block in the next-generation transport network," *Bell Labs Tech. J.*, Vol. 4, 1999, pp. 42–57.

D. Al-Salameh, M. Fatehi, W. Gartner, S. Lumish, B. Nelson, K. Raychaudhuri, "Optical Networking," *Bell Labs Tech. J.*,Vol. 3(1), January–March 1998.

ANSI T1.105.106-1996, GR-1377-CORE, Issue 5, 1998; GR-253-CORE, and IEEE 802.3. Note that 1000B-SLX is presently in the process of being standardized.

ANSI T1.105.01-1998, "Synchronous optical network (SONET)-automatic protection switching."

S. Baroni, P. Bayvel, R.J. Gibbens, and S.K. Korotky, "Analysis and design of resilient multi-fiber wavelength-routed optical transport networks," *J. Lightwave Technol.* Vol. 17, 1999, p. 743.

"SONET bidirectional line-switched ring equipment generic criteria," GR-1230-CORE, 1998.

Michel W. Chbat, et al., "Toward wide-scale all-optical transparent networking: The ACTS optical pan-European network (OPEN) project," *IEEE Journal on Selected Areas in Communications*, Vol. 16(7), September 1998.

A. Chen, R.W. Irvin, E.J. Murphy, R. Grencavich, T.O. Murphy, and G.W. Richards, "High performance lithium niobate switches for multiwavelength optical networks," Photonics In Switching '99, paper PFA2, 1999, p.113.

L. Chen, C. Shi, M. Shinta, H. Kamata, S. Nakamura, J. Chen, and M. Cvijetic, "A novel bi-directional wavelength path switched ring (BWPSR): principle and experiment," Optical Fiber Communication Conference, paper TuQ4, 2000.

A.R. Chraplyvy and R.W. Tkach, "Terabit/second transmission experiments," *IEEE J. Quantum Electron.*, Vol. 34, 1998, pp. 2103–2108.

C. Clos, "A study of non-blocking switching networks," *Bell Sys. Tech. J.*, 1952.

A. Crossland, et al., "Holographic optical switching: the 'ROSES' demonstrator," *J. Lightwave Tech.*, Vol. 18, 2000, pp. 1845–1854.

D. Forbes, G. Cannel, E. Pittuck, and N. Baker, "Optical shared protection ring performance," Proc. ECOC, p. II-52-3, 1999.

R. Giles and M. Spector, "The wavelength add/drop multiplexer for lightwave communication networks," *Bell Labs Tech. J.*, Vol. 4, pp. 207–229, 1999.

Goh, postdeadline paper, OFC, 2001.

E.L. Goldstein and L. Eskildsen, "Scaling limitations in transparent optical networks due to low-level crosstalk," *IEEE Photon. Tech. Lett.*, Vol. 7, 1995, p. 93.

E.L. Goldstein, L. Eskildsen, and A.F. Elrefaie, "Performance implications of component crosstalk in transparent lightwave networks," *IEEE Photon. Tech. Lett.*, Vol. 6, 1994, p. 657.

E. Goldstein, L. Eskildsen, C. Lin, and Y. Silberberg, "Polarization statistics of crosstalk-induced noise in transparent lightwave networks," *Photon. Tech. Lett.*, Vol. 71, 1995, pp. 1345–1347.

E.L. Goldstein, J.A. Nagel, J.L. Strand, R.W. Tkach, "National-scale networks likely to be opaque," *Lightwave*, February 1998.

W. Goralski, *SONET: A Guide to Synchronous Optical Networks*, McGraw-Hill, 1997.

F. Heismann, S.K. Korotky, and J.J. Veselka, "Lithium niobate integrated optics: Selected contemporary devices and system applications," in *Optical Fiber Telecommunications IIIB*, edited by I.P. Kaminow and T.L. Koch, Academic Press, NY, 1997.

Steven Hersey, et al., "Wanted more fiber capacity," *Telephony*, March 1997.

N.A. Jackman, S.H. Patel, B.P. Mikkelsen, and S.K. Korotky, "Optical cross-connects for optical networking," *Bell Labs Tech. J.*, Vol. 4, 1999, pp. 262–281.

S.R. Johnson, et al., "MONET DC network," *Bell Labs Tech. J.*, Vol. 4(1), March 1999.

R. Khosravani, and A. Wilner, "System performance evaluation in terrestrial systems with high polarization mode dispersion and the effect of chirping," *Photon. Tech. Lett.*, Vol. 13, 2001, pp. 296–298.

Optical Fiber Telecommunications III A&B, edited by I.P. Kaminow and T.L. Koch, Academic Press, NY, 1997.

K. Kowalenko, "IEEE honors Hewlett-Packard chief scientist, Lucent, and Sun," *IEEE Institute*, Vol. 24(11), November 2000.

H. Laor, et al., "Performance of a 576 × 576 optical cross connect," Proc. National Fiber Optic Engineers Conference, Vol. 1, 1999, p. 278.

B.H. Lee and R.J. Capik, "Demonstration of a very low-loss, 576 × 576 servo-controlled, beam-steering optical switch fabric," Proc. European Conference on Optical Communication, v. 4, 2000, p. 95.

D.S. Levy, Y. Kwon, and D.Y. Al-Salameh, "Optical layer shared protection using an IP-based optical control network," Optical Fiber Communication Conference Proceedings, Tu08-1-3, 2001.

L.Y. Lin, et al., "High-density connection-symmetric free-space micromachined polygon optical cross-connects with low loss for WDM networks," Proc. OFC'98, PD24, San Jose, February 1998.

L. Lin, E. Goldstein, and R. Tkach, "Free-space micromachined optical switches for optical networking," *IEEE J. Sel. Top. Quant. Elec.*, Vol. 5, 1999.

M. Makihara, et al., "Non-blocking $N \times N$ thermal-capillary optical matrix switch using silica waveguide technology," LEOS 2000, pp. 643–644.

"Media Access Control (MAC) Parameters, Physical Layers, Repeater, and Management Parameters for 1000 Mb/s Operation (Type 1000BASE-X)," IEEE 802.3z, 1998.

T. Miyazaki, T. Kato, and S. Yamamoto, "A demonstration of an optical switch circuit with 'bridge and switch' function in WDM four-fiber ring networks," *IEICE Trans. Commun.*, Vol. E82-B, No. 2, 1999.

E.J. Murphy, "Photonic Switching," in *Optical Fiber Telecommunications IIIB*, edited by I.P. Kaminow and T.L. Koch, Academic Press, NY, 1997.

E.J. Murphy, et al., "Enhanced performance switch arrays for optical switching networks," ECIO 1997, paper EFD5, pp. 563–566.

T.O. Murphy, et al., "A strictly nonblocking 16 × 16 electrooptic photonic switch module," ECOC 2000, v. 4, p. 92.

Y. Namihira, Y. Horiuchi, K. Mochizuki, and H. Wakabayashi, "Polarization mode dispersion measurements in an installed optical fiber submarine cable," *Photon. Tech. Lett.*, Vol. 1, 1989, pp. 329–331.

D.T. Neilson, et al., "Fully provisioned 112 × 112 micro-mechanical optical crossconnect with 35.8 Tb/s demonstrated capacity," Proc. Optical Fiber Communication Conference, paper PD12, 2000.

M. To and P. Neusy, "Unavailability analysis of long-haul networks," *IEEE J. Selected Areas in Communications*, Vol. 12(1), pp. 475–489, 1994.

K. Noguchi, "A rearrangeable multi-channel free-space optical switch based on multistage network configuration," *J. Lightwave Tech.*, Vol. 9, 1991, pp. 1726–1732.

K. Noguchi, "Optical multichannel switch composed of liquid-crystal light-modulator arrays and bi-refringent crystals," *Electron. Lett.*, Vol. 33, 1997, pp. 1627–1629.

R. Pease, "Optical-router battle heats up as OC-192 enters the core," *Lightwave*, March 2001, PennWell, Co., Tulsa, OK, p. 1.

C. Poole, R. Tkach, A. Chraplyvy, D. Fishman, "Fading in lightwave systems due to polarization-mode dispersion," *Photon. Tech. Lett.*, Vol. 3, 1991, pp. 68–70.

F.L.W. Rabbering, "Polymeric 16 × 16 digital optical switch matrix," postdeadline paper, ECOC, 2001.

Richards, "Architectures for multi-stage crossconnects," Optical Fiber Communication Conference, paper WI3, 1999.

R. Ryf, et al., "1296-Port MEMS transparent optical cross-connect with 2.07 Petabit/s switch capacity," Proc. 2001 Conf. Optical Fiber Commun. (Anaheim), Optical Society of America, 2001, post-deadline paper PD28.

C. Shi, L. Chen, M. Shinta, H. Kamata, S. Nakamura, J. Chen, and M. Cvijetic, "Optical layer protection of ATM traffics using a novel 4-fiber WDM ring with an ability to carry extra traffic," Optical Fiber Communication Conference, paper FE3, 2000.

T. Shiragaki, N. Henmi, H. Takeshita, H. Shimomura, and S. Hasegawa, "Bi-directional wavelength-path switched ring," *NEC Res. and Develop.*, Vol. 40, 1999, p. 22.

"SONET OC-192 Transport System Generic Criteria," GR-1377-CORE, Issue 5, 1998.

"Synchronous optical network: Physical Layer Specifications," ANSI T1.105.06-1996.

"Synchronous optical network (SONET) transport systems: common generic criteria—GR-253," Issue 2, December 1995.

H. Takeshita, S. Takahashi, T. Kato, S. Kitamura, H. Harano, and N. Henmi, "A demonstration of an optical cross-connect system for a self-healing optical network," IOOC-ECOC, 1997, p. 335.

D. Uehara, K. Asahi, Y. Nakabayashi, M. Yamashita, C. Konishi, and S. Fujita, "Highly reliable and economical WDM ring with optical self-healing and 1 : N wavelength protection," IOOC-ECOC, 1997, p. 22.

S. Yamamoto, N. Toda, L. Baxter, and K. Raychaudhuri, "Next Generation Submarine Networks Integrating Terrestrial Systems," SubOptic 2001, paper T3.5.4.

M. Young, H. Laor, and E.J. Fontenot, "First in-service network application of optical cross-connects," Proc. 1998 Conf. Optical Fiber Commun. (San Jose), Optical Society of America, 1998, post-deadline paper PD23.

Chapter 8 | Applications for Optical Switch Fabrics

Martin Zirngibl

Bell Laboratories, Lucent Technologies, Holmdel, New Jersey

1. Introduction

The explosion in bandwidth demand over the backbone network has led to a bottleneck in switching information. Indeed, the capacity of fiber transmission lines has been growing at a much faster rate over the past few years than the throughput capacity of switching nodes. Commercially available fiber transmission systems now carry more than 1 Tb/s of information per fiber. The largest electronic cross-connects have barely reached 1 Tb/s of total throughput. Now imagine 10 optical fibers terminating in a switching node and each one carrying eventually 1 Tb/s or more traffic. It is easy to see that switching nodes with at least 10 Tb/s throughput capacity are needed in order to handle this type of fiber capacities.

Until now, switching information has always been done electronically. Although optical switching techniques have been known for a long time, electrical switching has been the preferred approach because of the low cost and technological maturity of electronic switching gear. As we will show in this chapter, electronic switching has run into a scalability problem, not so much because of the switches themselves, but because of the transport of data within the switch, the so called backplane problem. We are now witnessing the appearance of the first commercial optical cross-connects (OXC), which get around the backplane problem by using optical transport and switching technology. These OXCs have the potential of throughputs of the order of several tens of Tb/s. One of the main points of this chapter is that packet switches, which by their nature are more complex than cross-connects, will run into the same scalability issues, and will eventually need some sort of optical switching technology as well, albeit one that is much faster than the switching technology employed by OXCs.

We will start out by defining the different functionalities of switching nodes and say a few words about "transparent" nodes. We will then discuss the limitations of electrical switching nodes but also show some of the technologies that make electronic switches more scalable. Finally, we will describe some of the optical switching technologies for XCs and elaborate in what kinds of architecture they are the preferred approach. In the last part of this chapter, fast optical switching technologies are discussed. In particular, we will present what

OPTICAL FIBER TELECOMMUNICATIONS, VOLUME IVA

ISBN: 0-12-395172-0

we believe is the most promising fast optical switching technology: wavelength switching.

2. Networks and Switching Nodes

2.1 CIRCUIT SWITCHING VERSUS PACKET SWITCHING

As is well known, the two main methods of routing information in a network are circuit switching and packet switching (KESHAV97). In a circuit-switched network, a connection with a fixed bandwidth is first established between two end points before the information is sent. Once the connection is set up, the bandwidth is available until the connection is broken. Switching nodes for circuit-switched networks are usually called cross-connects. A cross-connect has to be able to set up connections between any open input and output port. A circuit-switched network is managed by an overlay network that tells the cross-connect the connections that have to be established.

In packet-switched networks, data packets with a destination address are sent from switching node to switching node. Each node stores and forwards packets, not unlike a mail office. A switching node for a packet-switched network thus needs to be capable of reading address headers, storing packets, and routing them to the correct location. Switching nodes for packet-switched networks are usually called packet switches or routers. There is no overlay network in a packet-switched network to manage it. All the routing and networking information is contained in the packet itself, and the packet switch makes autonomous decisions on how to route individual packets.

There has been a vigorous debate in the technical community about whether the future network will be circuit-switched or packet-switched. Packet-switched networks use bandwidth more efficiently, but their switching nodes are more complex. There seems to be an emerging consensus that the future network will have a circuit-switched core network and a packet-switched access network (Fig. 1). The bandwidth of the circuit (connection) in the backbone would correspond to the capacity carried by a wavelength in a wavelength division multiplexed (WDM) transmission system. Typically, a wavelength carries a data rate of 2.5, 10, or 40 Gb/s. Thus, the future backbone will be not unlike today's SONET/SDH systems but with a much larger circuit size (in today's SONET a circuit is 50 Mb/s). In this network scenario, packet switches on the edge of the backbone aggregate the traffic to high data rates.

There have been numerous studies to determine where the boundary between the packet-switched and circuit-switched network should be (DNPBL2001). There is certainly no scientific answer to this question. The cost of transmission bandwidth versus the cost of switching will ultimately determine this boundary. If transmission bandwidth were free, the network would become a circuit-switched network with everybody having a hard-wired

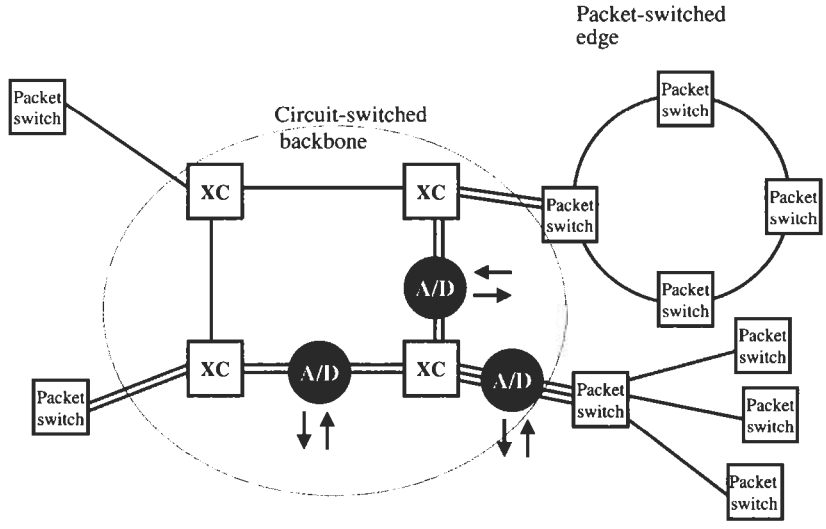


Fig. 1 Emerging architecture for IP-dominated service network. At the edge, packet switches aggregate the traffic to large size circuits that are routed via the backbone network.

point-to-point link to everybody else. Obviously, this is not possible. On the other hand, if packet switching were very inexpensive but bandwidth scarce, packet switches would be used in the backbone to squeeze the last bit of efficiency out of the transmission bandwidth.

It should be noted that between nodes, traffic is always encoded in some sort of a framed signal. There are no "real" optical packets flowing over the network, and thus the optical fiber transmission system for a packet-switched or circuit-switched network is pretty much the same.

2.2 DEFINITION OF VARIOUS SWITCHING NODES

To facilitate the discussion in this chapter, we will now define four different types of switching nodes that may be used in the network scenario outlined above. The nodes are schematically displayed in Fig. 2. For purpose of illustration, we show these nodes being connected to WDM transmission systems, but except for the node in Fig. 2a, these nodes are independent of the transmission systems they are interconnecting.

The wavelength-selective cross-connect (WSC) in Fig. 2a routes optical channels without any optoelectronic conversion. It is thus said to be optically transparent. By the same token, the wavelength of the optical channel will not be altered, a blue channel will stay blue; however, it may be switched from one fiber to another. Since no two wavelengths on a single fiber can be the same, such a cross-connect is blocking in the wavelength domain.

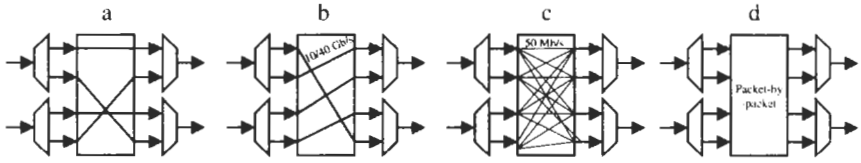


Fig. 2 Four types of switching nodes in a telecommunication network: a) Wavelength Selective Cross-connect (WSC), b) Wavelength Interchanging Cross-connect (WIC), c) Digital Cross-connect switch (DCS) and d) a packet-switch.

In a wavelength-interchanging cross-connect (WIC) in Fig. 2b, an entire wavelength channel is routed. The outgoing wavelength bears no relation to the incoming wavelength; thus there is wavelength interchange through optoelectronic-optic (OEO) conversion. The outgoing sets of wavelengths could be a completely different set from the incoming ones. A WIC doesn't even have to be attached to a WDM system; each port may be attached to a single-channel fiber transmission system running in a short wavelength region. It should be noted that a WIC is often referred to as optical cross-connect (OXC), although, as we shall see later, the WIC may be built with pure electronic switching technology. Thus the term "optical" emphasizes the black box functionality, i.e., the routing of optical channels rather than the underlying technology. We will use the term OXC only for a WIC with an optical switch fabric.

The node in Fig. 2c is commonly referred to as digital cross-connect switch (DCS). DCSs have been widely deployed in SONET and SDH networks. In a DCS, an incoming optical channel at, say, a bitrate of 10 Gb/s is first time-division demultiplexed down to the SONET granularity of STS-1 or 50 Mb/s. Each of the subchannels of 50 Mb/s is then individually cross-connected through the DCS. At the output port, the individual channels are time-division multiplexed (TDM) to form a high-bitrate channel (2.5, 10, or 40 Gb/s). Although a DCS could conceivably have an optical switch fabric, electrical switch fabrics are strongly favored here because of the low bitrates and high port counts.

Finally, the node in Fig. 2d is a packet switch. It makes routing decisions on a per-packet basis; it therefore has to switch very fast. It also retrieves all the traffic routing information from the data itself.

2.3 SOME SIMILARITIES AND DIFFERENCES BETWEEN THE VARIOUS SWITCHING NODES

These various switching nodes will all need some sort of a switch fabric. The switch fabric usually must be strictly nonblocking (see Hui90), although for the WSC, obviously, blocking would still occur in the wavelength domain. However, cross-connects usually do not have to switch very fast; these connections are set up for long periods of time. Switching times of the order of milliseconds are required for protection switching. A packet switch on the other hand will

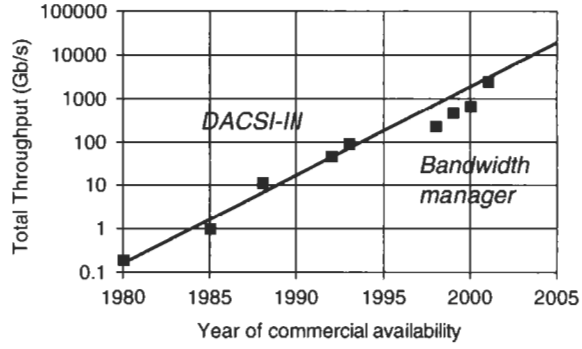


Fig. 3 Total throughput of Lucent's electrical cross-connects against year of commercial availability.

need a fabric that can be rearranged packet by packet. Since we don't want to lose much time to set up a connection through the packet switch, the switching time should be much less than the typical length of a packet.

In Fig. 2, the black-box functionality increases going from (a) to (d). If we keep the number of ports and port speed the same, a WIC can, of course, play the role of a WSC. Likewise, a DCS could be used as a WIC by just disabling its TDM demultiplexing function. Finally, a packet switch can be trained to behave just like a cross-connect by setting up fixed routing as in multiprotocol label switching (MPLS) (DR2000). This brings us to the obvious question: why would one bother to build any of the three left boxes if all the problems can be solved with a packet switch? The answer lies again in the cost and scalability issues of these different types of switches. For equal throughput a cross-connect costs approximately one tenth as much as a packet switch. Also, state-of-the-art cross-connects have about 10 times the throughput of state-of-the-art packet switches (see Fig. 3 and Fig. 17). This difference has obviously to do with the much richer functionality of a packet switch. By the same token, any new cross-connect technology will be successful only if it either dramatically lowers the per-port cost or has a much better scalability property than a corresponding packet switch. This point is sometimes lost when new optical switching technologies are proposed. We will now examine each of these nodes more closely to determine how optical switching technology can be used.

3. The Wavelength Selective Cross-Connect (WSC)

3.1 ARCHITECTURE OF WSC

Let us now take a closer look at the WSC. Assume we have M optical fibers coming into a central office, each fiber carrying N optical wavelength channels (Fig. 4). The WSC optically demultiplexes each wavelength channel. The

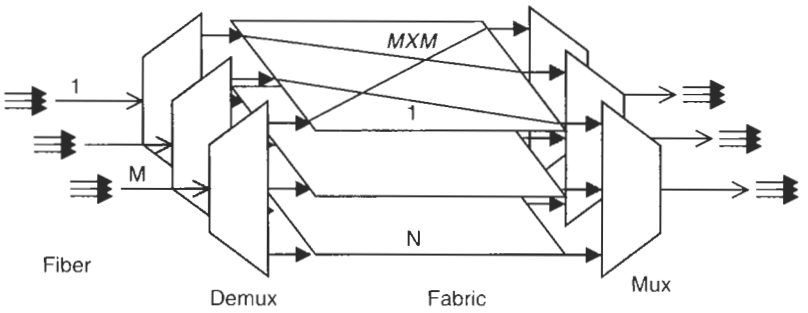


Fig. 4 A wavelength selective cross-connect interconnecting M incoming fibers to M outgoing fibers, each one carrying N wavelength channels.

switch fabric that cross-connects the channels thus consists of N $M \times M$ optical fabrics. A strictly nonblocking space fabric (Hui90) of size $M \times M$ has a complexity of M^2 (M log(M) type complexity can be obtained with certain multistage architectures). Thus, the switch fabric of our wavelength-selective cross-connect has in general a complexity of the order of NM^2. It further contains M pairs of $1 \times N$ optical demultiplexers–multiplexers. Since there is no OEO conversion, the fabric has to be optical.

3.2 ADVANTAGES AND DRAWBACKS OF WSCs

The main selling point of WSCs is the absence of optoelectronic regenerators. A regenerator consists of a back-to-back optical receiver and transmitter and usually performs a 3R functionality (retiming, reshaping and reamplification of the signal). In a fully loaded network, regenerators are the dominant cost, and therefore, a cross-connect that does not need them could have a major cost advantage over cross-connects that need them. Another advantage of WSCs is their format and bitrate transparency. Indeed, since the data is never electronically processed in the WSC, a wavelength channel could have any format or bitrate subject to the limitation of the total optical bandwidth available. This format transparency was a major motivation in the early push for WDM networks (WASG96). However, since transmission rates and formats are becoming increasingly uniform around the SONET standard or Ethernet, the importance of signal transparency has faded somewhat.

WSCs however have some major drawbacks that have so far prevented their use in real commercial applications. As already mentioned, WSCs are blocking in the wavelength domain. If channel x from input fiber 1 and channel y from input fiber 2 want to go to the same output fiber and if they happen to be on the same wavelength, they will block each other. Unblocking them would require one of the channels to change its wavelength, something that can only be done at the source, which might be hundreds of kilometers away. Thus,

wavelength channel routing would require coordination of the wavelengths at all transmitters in the entire optical network. Although this would be possible, it would certainly be very cumbersome and would add a lot of complexity to the network management. Nevertheless, it has been shown (Bayvel2000) that wavelength blocking would not be a fundamental problem if managed networkwide.

Another major issue with WSCs is their analog nature (TGNS98, Goldstein98, BCG95). Since the same photons coming from the input fibers go out into the output fibers, the transmission systems on both sides and all fibers connected to a WSC would have to be identical. For instance, they would be required to use the same wavelength set, similar power levels, and have similar signal-to-noise requirements. In short, their analog characteristics would have to be the same. Since there is no real analog standard for fiber transmission systems, all the systems connected to the WSC would have to come from the same vendor. Even worse, if one of the fibers needed to be upgraded to a denser wavelength set, that would not be possible. Thus, a WSC leads to an inherently closed-interface, single-vendor solution that is not upgradable. Needless to say, any service provider will be very hesitant to deploy such a system. A WSC also impairs the optical signal as it passes through. Impairments such as filter narrowing, loss, and polarization dependent loss all degrade the signal and limit the maximum reach of the optical system. Thus, when designing the total reach of the system, the impairments coming from the WSC would have to be accounted for. Finally, there is no easy way to monitor the quality of an optical signal without going through a full OEO conversion. Although optical power and optical signal-to-noise ratio could certainly be measured with relatively simple optical filters and low speed electronics, the ultimate health of the signal can only be determined by an actual bit-error rate measurement based on checking the data itself and this requires a receiver running at the full data bandwidth.

Despite all the shortcomings of WSCs, their potential for savings in network hardware are huge. For this reason, there is still a very intense research effort going on in this field that covers networking as well as hardware aspects (Saleh2000). We shall now review a few types of WSC without claim of presenting an exhaustive list.

3.3 SOME EXAMPLES OF WSCs

One of the first WSCs was built by the Multi Wavelength Optical Networking (MONET) consortium (WASG96). It fully demultiplexed each wavelength channel before cross-connecting them, just as shown in Fig. 4. Each fiber into the switch fabric carries only one channel. The problem with such an architecture is that it is hard to recover from optical losses that inevitably arise from the optical demultiplexing, switch fabric, and optical multiplexing. Although

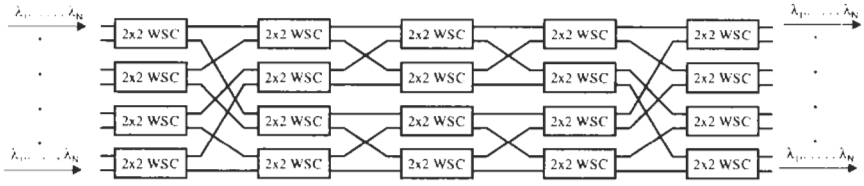


Fig. 5 Possible implementation of an 8 × 8 WSC based on a network of 2 × 2 WSC.

optical amplifiers could be used, this is an expensive proposition if one amplifier has to be used for each wavelength channel. If the total insertion loss becomes low enough, however, optical amplifiers could be used before or after the optical demultiplexers or multiplexers, in which case one amplifier would be shared over many wavelength channels. Low-loss fabrics based on microelectromechanical switch (MEMS) technology (see below) might change the picture for this type of WSC.

One can also build a $M \times M$ WSC by cascading 2×2 WSC in Benes-type architectures (Benes35, WMDZ99) as shown in Fig. 5. A 2×2 WSC is a four-port device that can selectively switch into a bar or cross state for each of the N wavelength channels. There are different ways to build 2×2 WSCs (Doerr98, SCBBJ96, RTBDC2000, NTS99). The advantage of such a WSC is that despite lossiness, the wavelength channels are never individually demultiplexed, and thus only a few optical amplifiers are needed to recover from the loss. As a disadvantage, we note that in such a multistage architecture issues like optical crosstalk and analog impairments of the signal are more severe.

4. The Optical Add/Drop: Special Case of WSC

4.1 A/D IN NETWORKS

Electronic Add/Drop (A/D) nodes are important elements in today's SONET networks. They are used in low-dimensional nodes with only one input/output port where a small fraction of the traffic is dropped and added. Imagine now a fiber carrying a hundred wavelength channels going through an A/D node. Instead of an electronic A/D, which would require a full OEO conversion on each wavelength channel, we now drop and add a few wavelengths but keep the channels that want to go through in the optical domain. Such optical A/D has a huge cost advantage over a solution that would require OEO on each channel (SGS96). Of course it is assumed that the traffic has been groomed onto the wavelength channels in such a way that none of the data on the through channels actually wants to drop.

4.2 OPTICAL A/D TECHNOLOGIES

The optical A/D can be viewed as a 2 × 2 WSC with an optical multiplexer at the add port and a demultiplexer at the drop port (Fig. 6). The most straightforward approach to building an optical A/D is to take a back-to-back demultiplexer–multiplexer pair and interconnect them with an array of 2 × 2 switches as shown in Fig. 7a. The drawback of this approach is that it leads to extensive filtering of the signal. Indeed, every time the signal passes through an optical multiplexer or demultiplexer, it is optically filtered, which cuts into the total optical bandwidth available to the signal and eventually destroys it (ALC98, CFFKS97). In order to be useful in networks, an optical A/D should be cascadable many times before full regeneration of the signal is necessary. Thus, a key figure of merit for optical A/Ds is their filter response. Ideally, we would like this filter response to be rectangular: totally flat in the middle of the channel and infinitely steep on the edges with no gaps between neighboring channels. Of course this is only an ideal, but there are ways to get close. One approach is to use coherent reconstruction of the output optical

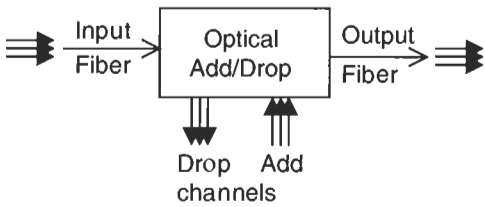


Fig. 6 Schematic view of an optical A/D.

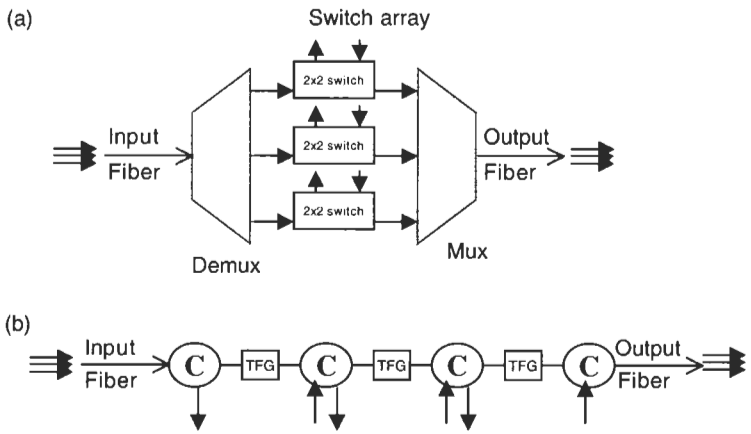


Fig. 7 Possible implantations of flexible A/D based on: a) optical demultiplexer interconnected to an optical multiplexer via 2 × 2 switches, b) tunable fiber gratings and circulators.

field (DSCLP1999). Such a device has, in principle, a totally flat spectrum if all channels are in the through state. A channel that drops, however, still causes some amount of filtering on its neighbor and puts an upper limit on how many times it can drop before destroying its neighbors. A four-node network with 1.6 Tb/s capacity has been shown using such a device (KCNDS2000). Optical A/D based on tunable fiber gratings (Fig. 7b) have also a good filter response (KLKCJ98, GM95); but that comes at the expense of dispersion and low wavelength ripple. It is fair to say that, as of the writing of this chapter, the ideal solution to optical A/D has not yet been found, although optical A/D has been on the wish list of every service and systems provider from very early days. The requirements of cascadability, flexibility, and manufacturability makes the optical A/D a tough problem to solve.

5. Summary on Transparent Routing

WSCs could potentially lead to major cost savings in optical networks because they eliminate expensive per-channel optoelectronic regeneration. However, their analog nature makes them difficult to use in an open multivendor environment, and their wavelength blocking characteristics complicate network management. One more immediate application of a WSC is the optical A/D. But even for this much simpler application, analog impairments of the through signals are still an issue for many network applications.

6. Wavelength Interchanging Cross-Connect

6.1 APPLICATIONS

Let us now take a look at a WIC (Fig. 8). Again we have M incoming fibers, each carrying N wavelengths. First, all the wavelength channels are demultiplexed. Thus, we have NM optical inputs into the cross-connect. The fabric itself is a $NM \times NM$ strictly nonblocking switch fabric. It therefore has a complexity of N^2M^2. Note that this is N times more complex than the switch fabric of a WSC of the same size. As far as the switch fabric is concerned, it is not really relevant whether the signals come from a WDM system; all that matters is that there are NM optical signals going into the fabric. For purposes of illustration in Fig. 8, we have an optical receiver at the input of the fabric and a transmitter at the output. We will see later that we actually could place the optical receiver at the output and not have any optoelectronic conversion at the input in certain cases.

Often the best technology choice for a cross-connect not only depends on the required throughput but also on where the boundary between cross-connect and transmission system lies. In Fig. 9, we display 3 possible boundaries. In the first case (Fig. 9a), the optoelectronic receiver and transmitter are part of the cross-connect; in this case the signals to be cross-connected are electrical

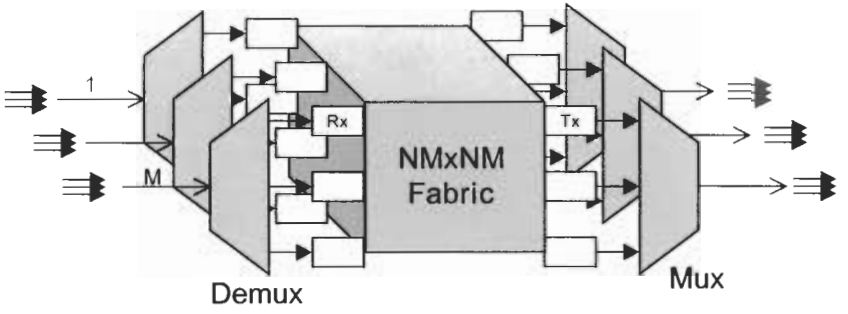


Fig. 8 A wavelength interchanging cross-connect interconnection M incoming fibers to M outgoing fibers each one carrying N wavelength channels.

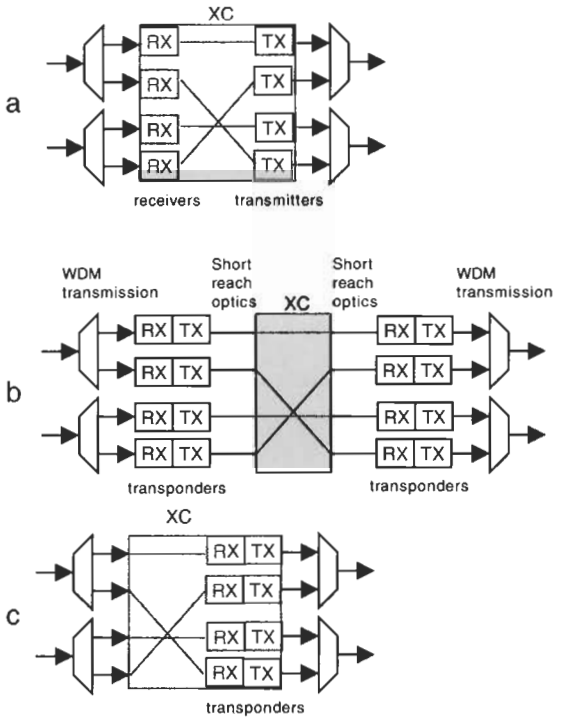


Fig. 9 System boundaries for WICs: a) receiver on ingress and transmitter on egress side are part of cross-connect, b) Transmission receiver/transmitters are outside cross-connect and c) no Receiver/transmitter on ingress side but on egress side.

in nature. In the second case (Fig. 9b), the signals to be cross-connected are optical, and therefore there are no receivers and transmitters on the cross-connect itself. The third case we consider (Fig. 9c) is a cross-connect that has no O-E conversion at the input but has OEO at the output. We should

emphasize that we only have defined the black-box functionality, not what is inside the actual cross-connect. Optical or electrical switch fabrics could be used for any of these three cases but it is clear that electrical switch fabrics lend themselves best to the case in Fig. 9a, whereas optical fabrics are best suited for the cases in Fig. 9b and c. A cross-connect whose receivers and transmitters are directly connected to a WDM transmission system is said to have compatible optics.

6.2 ELECTRICAL CROSS-CONNECTS

6.2.1 History and State of the Art

Figure 10 shows the schematics of an electrical cross-connect. Each optical signal is received in an optical receiver and converted to an electrical signal. The port card (PC) on the cross-connect performs a full 3R regeneration of the signal as well as electrical signal processing functions such as signal error check. The electrical signal is then transported to and from the switch fabric over an electrical backplane. Today's electrical backplanes support only up to 2.5 Gb/s (SMITHGALL2001). So if the cross-connect has a granularity of 10 Gb/s, then the 10 Gb/s datastream would be routed as 4 parallel 2.5 Gb/s channels within the switch.

Figure 3 shows the throughput of commercially available DCSs from Lucent plotted over the last two decades (Hunt99). As throughput we define the number of input/output ports times the port speed. Thus a 16 port cross-connect with throughout of 2.5 Gb/s per port would have a total throughput of 40 Gb/s. Of course, taking throughput as a figure of merit can sometimes be misleading, because it does not take into account the number of ports. For instance, a single-port cross-connect with 40 Gb/s would also have 40 Gb/s throughput; needless to say, such a cross-connect would be useless. So if we

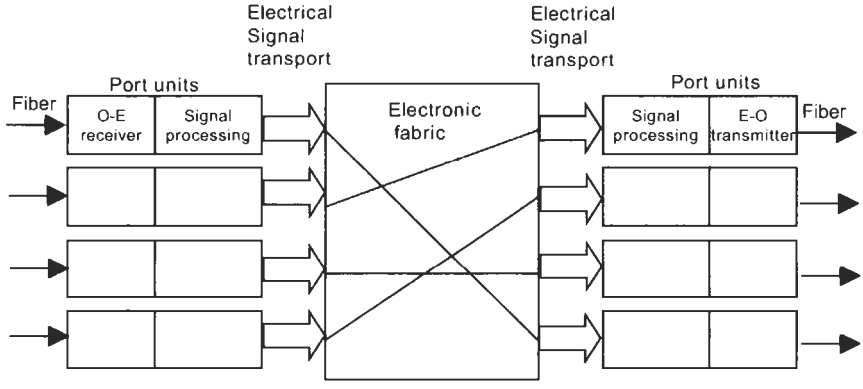


Fig. 10 Schematic view of an electrical cross-connect.

compare throughput of different cross-connects, we just have to keep in mind the number of ports and port speed in order not to compare apples to oranges. In the early 1980s, the first series of digital cross-connects had a total throughput of less than 1 Gb/s. The per-port speed of these systems was only 64 Kb/s (DS0 rate); they had several thousand ports. Later on, the per-port speed increased to 50 Mb/s, and the total throughput to several Gb/s. From this plot, we deduce the "Moore's law" for electrical cross-connects to be a doubling of the throughput every 18 months. The question is, will this Moore's law continue for electrical cross-connects?

In Fig. 11 we have listed some commercially available state-of-the-art cross-connects (RHK2001). Clearly for cross-connects with less than 1 Tb/s throughput, electrical-fabric-based solutions can do the job, although some of these cross-connects already use parallel optical interconnects, because they need multiple bays of equipment (see below). It should be noted that some of these electrical cross-connect support sub-rate grooming which means that they can cross-connect individually lower speed channel (typically STS1 or 50 Mb/s). We are now witnessing the appearance of the first cross-connects based on optical switch fabrics. Lucent's Lambdarouter (Lucent2001) is a 256 × 256 port 40 Gb/s per-port-speed cross-connect with a total throughput of 10 Tb/s; other major equipment suppliers have announced cross-connects with optical switch fabrics as well. Clearly there is an industry-wide move away from electrical fabrics toward optical fabrics for throughputs in excess of 1 Tb/s. Will this crossover move up or down over time? Let us take a look at the key pieces in an electrical cross-connect.

Product/company	*Fabric/size*	*Granularity/ throughput*	*Footprint*	*Interconnection*
Ciena CoreDirector	256 × 256 electrical	OC-48 640 Gb/s	Single bay	Electrical backplane
Sycamore SN16000	512 × 512 electrical	OC-48 1.2 Tbs	3 bays	Parallel optics
Tellium Aurora	512 × 512 electrical	OC-48 1.2 Tb/s	4 bays	Electrical cables
Alcatel Crosslight	512 × 512 Optical	OC-192 5 Tb/s	8 bays	Optical
Lucent Lambdarouter	256 × 256 Optical	OC-7868 10 Tb/s	3 bays	Optical

Fig. 11 Table of state-of-the-art commercially available cross-connects in the year 2001.

6.2.2 Scalability Issues of Electrical Cross-Connects

There are many ways to build electrical switch fabrics; the reader is referred to the excellent review by Keshav (Keshav97). There are now commercially available crossbar chips with 136 × 136 ports and 2.5 Gb/s per-port speed (CONEXANT2000, GLW2001). With 4 parallel chips, one can already build a fabric with a throughput in excess of 1 Tb/s. Using a multistage architecture, switch fabrics of many Tb/s are certainly feasible.

Obviously, the electronic switch fabric itself is not the problem to reach multi-Tb/s. Why, then, are people so much interested in optical switching technology? The answer is that transporting data from the port card to the switch fabric poses a problem for large cross-connects. The port card has nonzero size and also consumes quite a bit of power. For a 10 Gb/s receiver–transmitter pair, typical power consumption ranges from 20–50 W. The maximum power consumption per piece of equipment is typically 6 kW. The ports of the cross-connect cannot be spaced infinitely close, and the signal has to travel a certain distance between the port card and the switch fabric over the backplane.

Let us illustrate this point by studying a typical layout of a single-bay electrical cross-connect (Yang2000) as shown in Fig. 12. The cross-connect has circuit-pack slots containing either the port cards or switch. The communication between the two flows in electrical form over a backplane (Fig. 13), is a printed circuit board consisting of connectors and multiple levels of microstrip lines. Attenuation and signal degradation limit the distance the signal can travel at 2.5 Gb/s to approximately 0.5 m (CPR2000). The switch fabric resides in the center of the bay such as to minimize travel distance to and from the port cards.

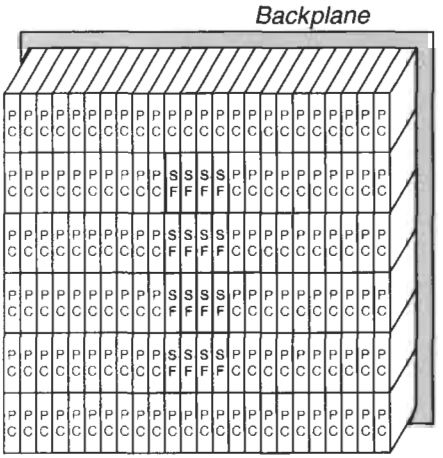


Fig. 12 Typical layout of a single bay electrical cross-connect with port-cards (PC) and switch fabric (SF) circuit packs interconnected via an electrical backplane.

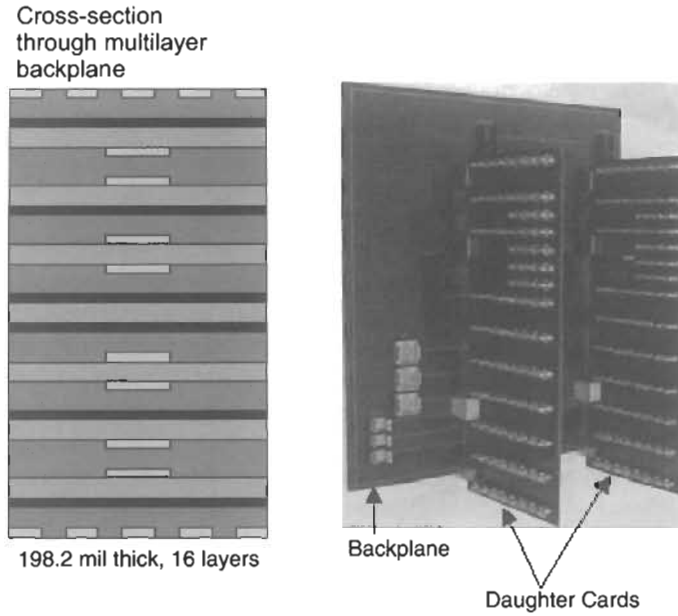


Fig. 13 Cross-section and photograph of a multilayer electrical backplane.

If a cross-connect can be built as shown in Fig. 12, it is usually a cost-effective solution for the type of application in Fig. 9a since the backplane and switch fabric chips are fairly low in cost. However, if the switch has too many port cards for a single-bay solution, we will have a major problem interconnecting multiple bays in a cost-effective and practical way. That describes the problem of the scalability for electrical switches. It is not the switch fabric itself, which can be scaled to a very large size; it is the transport between the port cards and the switch fabric. Thus we claim that, as long as an electrical switch can be built in a single bay with a single backplane, it is very hard to beat it economically by a solution based on an optical fabric, except in places where the signals are already in optical form as in Fig. 9b and 9c. The major impact of optical switch fabrics is for cross-connects that no longer fit into a single bay.

There are various interconnection technologies to solve the transport problem within multibay cross-connects (Fig. 14); these are currently used in commercially available cross-connects (Fig. 11). We can of course transport electrical signals between different bays of equipment, depending on the signal bandwidth, for example, coax, twisted-pair or LVDS (low voltage differential drive) cabling may be used. But remember there may be potentially thousands of signals travelling between these different bays of equipment, and this quickly leads to a cabling nightmare. One technology increasingly used to solve this problem is parallel optical interconnects. This technology is based

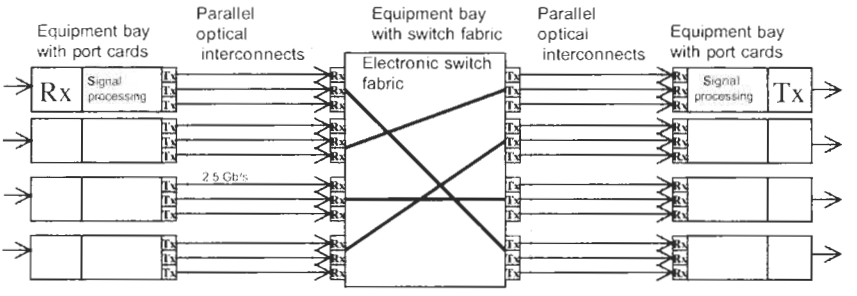


Fig. 14 Schematic view of an electrical cross-connect with optical interconnects between port cards and switch fabric.

on vertical cavity laser (VCSEL) arrays (PMFCH2000) connected to a multimode fiber ribbon of typically 12 fibers. State-of-the-art optical interconnects carry 12 2.5 Gb/s signals over a distance of 100–300 m; thus one ribbon has 30 Gb/s of throughput. The essence of this technology is low, small form factor connectors and low-power consumption arrays of transmitters/receivers. If the interconnection technology itself required a lot of space and power on the switch fabric side, then we would have solved much of the problem, since scalability and power consumption would again limit the size of the switch.

6.3 THE OPTICAL CROSS-CONNECT (OXC)

If the size of the cross-connect is such that optical signals have to be used within the switch anyway, then there is a strong case to be made to do the switching optically as well (Fig. 15). By eliminating all the electronics on the switch fabric side, we have solved the power consumption and size problem. Furthermore, optical fabrics have almost unlimited per-port bandwidth; thus we can cross-connect any rate (10 Gb/s, 40 Gb/s, 160 Gb/s); even several WDM channels or entire fiber capacities can be routed simultaneously, making the per-port speed multiple Tb/s. The total throughput is therefore basically unlimited (Ryf2001), although the number of ports is certainly not. Furthermore, the feature of sub-rate grooming is lost in an optical cross-connect.

6.3.1 Requirements for Optical Switch Fabrics

There are many potential technologies that might be used for optical cross-connects (OXC). They have to satisfy two main criteria: low optical loss and scalability. The low-loss requirement comes from the fact that in a fully loaded OXC (one where all the port units are installed) the total system cost is largely dominated by the cost of the port cards. The port card on an OXC has a transmitter that sends a signal into the switch fabric and a receiver that receives the signal coming out of the switch fabric. The difference between transmitter

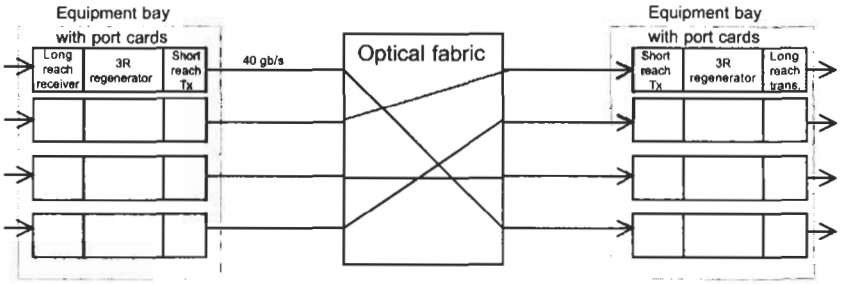


Fig. 15 Schematic view of optical cross-connect with 3R regeneration on both input and output ports.

output power and receiver sensitivity is the power budget that equals the maximum loss the overall optical fabric is allowed to have. The cost of the transmitter/receiver is very sensitive to the transmitter output and power and receiver sensitivity. Thus, a high-loss fabric will require high-power transmitters and high-sensitivity receivers. Since the cost of the fully loaded system is already largely dominated by the cost of the transmitters/receivers on the PCs, the cost of the overall OXC is mainly determined by the loss in the fabric, even if the fabric itself were free. A typical loss budget for a pair of short reach transmitters/receivers at 10 Gb/s is 7 dB.

The scalability requirement comes from the phenomenal growth in fiber transmission capacity and the fact that electrical cross-connects can cover applications for one and maybe several Tb/s throughputs. If we extrapolate the growth rate a few years into the future (Fig. 3), we would forecast that a 40 Tb/s throughput cross-connect (1000 × 1000 at 40 Gb/s line rate) will be needed by the year 2005. Thus, a technology that would not get us there is not very interesting. Two promising technologies for optical switching are microelectromechanical switching (MEMS), see other chapters in this book, and microbubbles.

6.3.2 Technologies for Optical Fabrics for Cross-Connects

MEMS technology, based on mirror arrays, has attracted much interest lately, and the first commercial optical cross-connects are now commercially available (Lucent2001); the promise of this technology is to provide very low-loss fabrics that are scalable. The low loss comes from the fact that the beam propagates in free space and is therefore not attenuated. The difficulty is the very tight control of the mirror positions. Another attractive technology is based on microbubbles (Fouquet2000). Here, N input ports are connected to N output ports through a matrix of N^2 switches that are digitally activated by heating a microbubble. The advantage of this switch is that it is easy to control and it is very cheap. The major drawback is that the insertion loss goes up linearly with

the number of ports. Thus, creation of low-loss large-size fabrics is a major challenge, even if multistage architectures are used.

6.4 SUMMARY ON CROSS-CONNECTS

Electrical cross-connects run into a scalability problem because of the signal transport between the port cards and the switch fabric. The electrical switch fabric itself is not the limiting factor and can be scaled to several Tb/s throughput. Key technologies for electrical cross-connects are (a) small form factor, (b) low-power consumption port card, (c) high-performance backplanes and (d) cheap, reliable small footprint and low-power-consumption parallel optical interconnects. These technologies are likely to significantly enhance the performance of cross-connects with electrical switch fabrics. Above 1 Tb/s total throughput, optical switch fabrics are useful today, because they avoid the interconnection problem by switching in the optical domain. Optical switch fabrics have almost unlimited bandwidth: critical issues are low optical loss and scalability.

7. Packet Switches

7.1 APPLICATION OF PACKET SWITCHES

Unlike a cross-connect, which acts more like a railroad switch, a packet switch has to be capable of routing individual packets, more like a mail office. The routing information is contained in the packet itself. From a high-level point of view, a packet switch has these three distinct elements: (a) a fast, strictly nonblocking switch fabric that connects ingress ports to egress ports, (b) buffers on ingress and egress ports, and (c) a scheduler that manages connectivity on a packet-by-packet basis through the switch fabric (Fig. 16). Thus, the switch fabric for a packet switch has to be able to set up connections in nanoseconds, in distinction to cross-connects, where switching times of the order of milliseconds are generally sufficient. Our main point is that if there is a crossover from electrical to optical fabrics for cross-connects above a certain throughput, then the same crossover has to occur for packet switches for the same throughput or at even lower throughput. Remember, the packet switch contains the cross-connect functionality, so if it were practical to build an electrical 5 Tb/s-capacity packet switch, we could just take this box and train it as a cross-connect. Admittedly, this argument is a little simplistic, because there are other factors such as cost, but in general the argument will hold true.

7.2 STATE-OF-THE-ART PACKET SWITCHES

The reason why optical fabrics for packet switches have not yet been very prominent is that the throughput of commercially available packet switches has

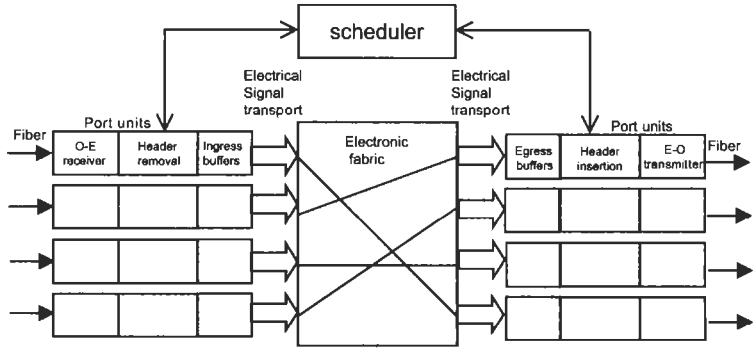


Fig. 16 Schematic view of an electrical packet switch.

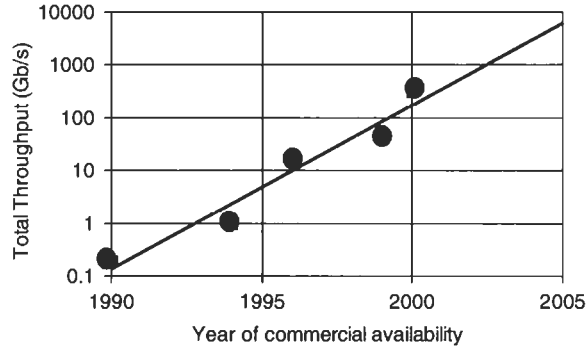


Fig. 17 Total throughput of packet switches versus year of commercial availability.

been an order of magnitude less than that of their cross-connect counterparts (Fig. 17) (McKeown2001). Indeed, the largest packet switches on the market today have a throughput of 160 Gb/s. But if one wishes to scale packet switches beyond 1 Tb/s, the same backplane limitations that we discussed above for the cross-connect will apply for the packet switch as well.

To solve this problem, large packet switches therefore need optical interconnection technology. Again, parallel optics can solve this problem to a certain extent. In Fig. 18, we show a typical architecture of a state-of-the-art packet switch that includes parallel optical interconnects between multiple bays of equipment (Stiliadis2000). Buffering is done on the switch fabric side and port card side. The flow of packets is managed through back-pressure: if a buffer is full, it tells all the other buffers in the previous stage to stop sending packets. Thus, the switch fabric side of such a packet switch is fairly complex and needs a lot of interconnections.

Since optics is already used within such a packet switch, it becomes very tempting to replace the electronic switch fabric by an optical fabric (Fig. 19). We eliminate all the optoelectronic conversion and buffering on the switch

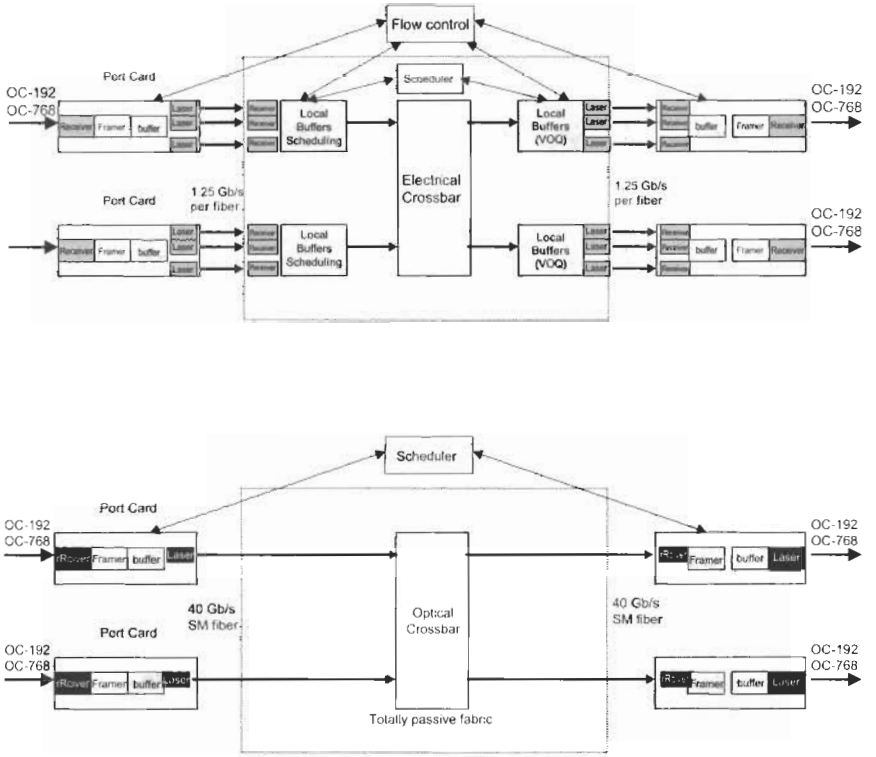


Fig. 19 Schematic view of a packet switch based on a passive optical switch fabric.

fabric side and run the optical fibers at the maximum optical data rate (for instance 10 Gb/s or 40 Gb/s). This allows us to cut down significantly on the number of fiber interconnection running between the port cards and the switch fabric and also solves the power consumption and size issue, just as it does for the cross-connect discussed above.

7.3 SOME OBSERVATIONS ON OPTICAL PACKET SWITCHING AND TRANSPARENT PACKET-SWITCHES

It is very important to note that a packet switch using an optical switch fabric looks to the network side just like any other packet switch. Recently, there have been attempts to build transparent packet switches (MSATT99, ZT98) that also include optical buffering and some optical signal processing capabilities. Our belief is that signal processing and buffering are very difficult to do in optics, and optical technologies attempting to do them are far from being mature; it is also not clear what problem they are solving.

The term optical packet switching is often used in a context where one sends real optical packets over the transmission network (GCHCK98, SCDJ99).

Conceptually this is similar to the architecture in Fig. 19, except that the distance between the port cards is hundreds of kilometers. The problem with such an approach is that it becomes difficult to use bandwidth efficiently when the transmission delay exceeds the packet duration; it also mixes transmission issues such as dispersion and noise accumulation with switching issues. In what we will describe below, we will assume that packet duration is always much longer than transmission delay; typical distances between port cards are 10–100 m, so that there are no signal impairments arising from fiber transmission.

7.4 REQUIREMENTS OF SWITCH FABRICS FOR PACKET SWITCHES

An optical switch fabric for a multi-Tb/s packet switch should have the following characteristics: it should be strictly nonblocking, it should switch in nanoseconds, the size should be of the order of 100×100 and the per-port speed should be 40 Gb/s. Such a fabric would allow us to get to a 4 Tb/s packet switch. There are several optical switching technologies that could potentially do the job. First of all, the technologies discussed above for the cross-connect are not candidates here. Both the MEMS and bubble switch have response times of the order of milliseconds, much too slow for packets at a 10–40 Gb/s rate. In general, any switching technology that relies on mechanical or thermooptical effects can be eliminated for this application. A typical packet has the order of 1000 bits (although packets vary wildly in length), which is 100 ns at 10 Gb/s or 25 ns at 40 Gb/s. A certain amount of packet aggregation (KLST2000) has been proposed for large packet switches, because scheduling large fabrics is a difficult computational problem regardless of the switching speed. The scheduler needs to have some finite computing time to figure out what the best connections are. Switching times that are significantly longer than a small fraction of the packet size might therefore be acceptable, but remember, packet aggregation comes at the expense of larger buffers and increased latency through the switch.

7.5 OPTIONS FOR FAST OPTICAL SWITCHING

There are several optical switching technologies that can indeed have switching times of nanoseconds or less. First to mind comes the $LiNbO_3$ Mach-Zehnder interferometer, a very popular data modulator that can switch in a few picoseconds. 16×16 switch fabrics have been built using $LiNbO_3$ based on multistage 2×2 switches (MSCCI2000). The main problem with these fabrics, though, is that the throughput loss increases very quickly with the number of stages. Another potentially fast switching technology is semiconductor optical amplifiers (SOA); prototype switches have been demonstrated using such devices (MSTH98). However, SOAs are not really a mature technology and are still

not commodity items, but they could potentially become very attractive for fast optical switching.

7.6 WAVELENGTH SWITCHING

7.6.1 Principle of Wavelength Switching

The switching technology we will focus on here is wavelength switching. Fast wavelength switching is not exactly a new idea (Kaminow91, SYHY97); what has changed over the last few years, however, is that the components needed for wavelength switching are increasingly maturing thanks to the strong demand for WDM systems and that the need for Tb/s packet switches is becoming urgent. As we will show below, there is a direct correspondence between wavelength switching and space switching. Let's start with an $N \times N$ optical multiplexer (Dragone91) as shown in Fig. 20. This component can be implemented with an arrayed waveguide grating (AWG). It has a prism-like functionality—the light that is injected into one of the ports is deflected to one of the output ports, the angle of deflection being determined by the wavelength. From each input port there is a wavelength that will connect this input port to one and only one output port. In general, we can write the wavelength connectivity in form of a matrix λ_{ij} $i = 1, \ldots, N$ and $j = 1, \ldots, N$. Under

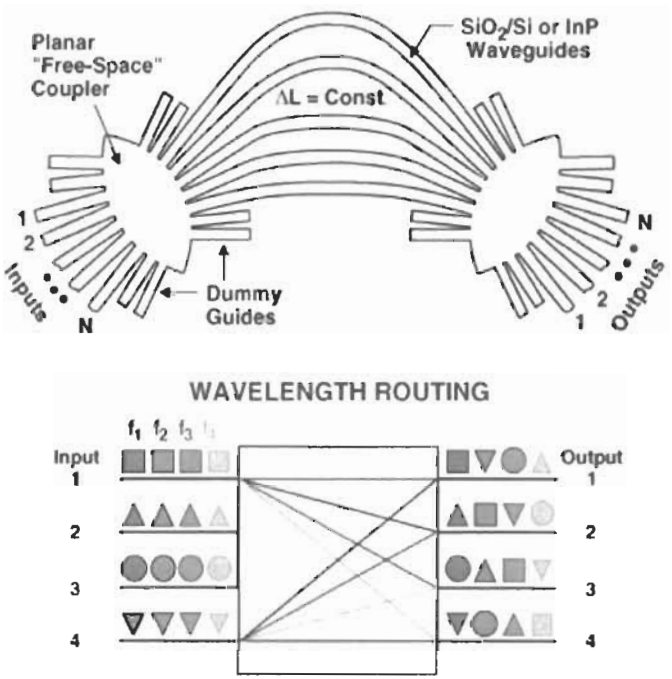


Fig. 20 $N \times N$ integrated arrayed waveguide multiplexer (AWG).

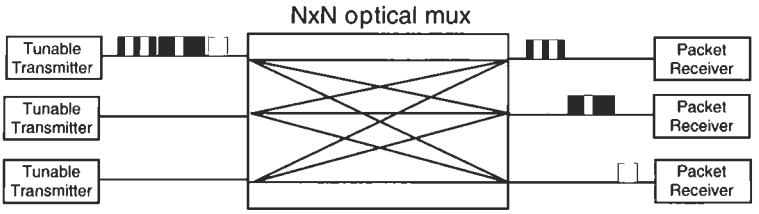


Fig. 21 Optical switch fabric based on tunable transmitters, $N \times N$ AWG and broadband receivers.

certain circumstances (BDDCL2000), the AWG represents a wraparound feature, which means that $\lambda_{11} = \lambda_{22} = \lambda_{33}$ etc. and $\lambda_{12} = \lambda_{23} = \lambda_{34}$; thus, the same set of wavelengths are used for each input port, and moving from one input port to another will just permute the set of wavelengths. It is important to note that wraparound is not a necessary condition; all that really matters is that the wavelengths of a particular set for any input port $x, \lambda_{xj}\ j = 1, \ldots, N$ are all distinct. To each input port of the AWG we connect a tunable laser followed by a data modulator. Each tunable laser can rapidly change (on a time scale of nanoseconds) its wavelength to any of the λ_{ij}; the light output from the laser is then modulated with data by a high-speed modulator. Thus, by choosing the appropriate wavelength, the laser selects the output port to which the packet information is sent. The tunable transmitter together with the $N \times N$ AWG become a strictly nonblocking switch fabric with a switching speed equal to the tuning speed of the lasers (Fig. 21). The data rate of the packet itself is limited only by the bandwidth of the data modulator and the optical bandwidth of the AWG channel. We have thus replaced the active switch fabric in Fig. 18 with a passive switch fabric that is basically a piece of glass. All the active electronics and intelligence is on the port card (Fig. 22). Of course, we now need a global scheduling algorithm that has to make sure that no two packets will arrive at the same time on the same output receiver.

7.6.2 Key Components for Wavelength Switching

7.6.2.1 The $N \times N$ Arrayed Waveguide Grating Multiplexer

We will now discuss in more detail the key components of the optical switch fabric. Let us start with the AWG. This device is based on integrated waveguide technology using a SiO_2-on-Si platform although other material systems can be used. Invented about 10 years ago (Dragone91, VS91), it is now in widespread use in WDM systems as an optical multiplexer and has become a commodity component with several vendors aggressively marketing it. This in turn has led to a steep price erosion. The per-port cost of such a device has become fairly low ($100–$1000). The three key performance parameters are insertion loss, crosstalk, and filter bandwidth.

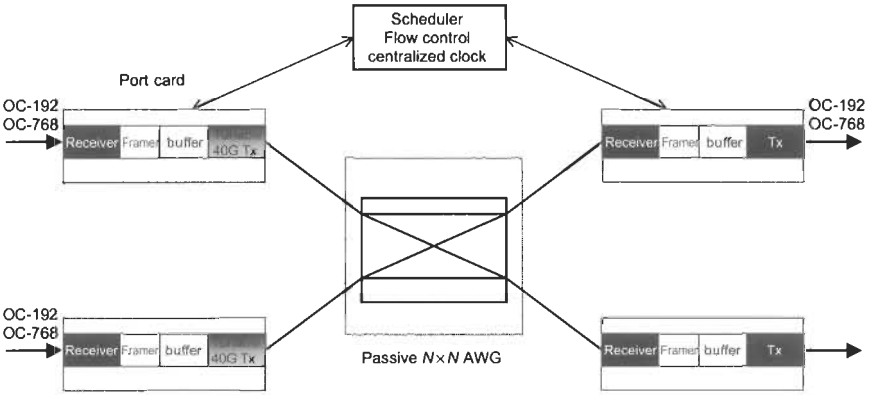


Fig. 22 Schematic view of a packet switch based on tunable lasers and passive $N \times N$ AWG.

Insertion losses arise from coupling into and out of the device, from waveguide absorption losses and from design losses. Typical losses guaranteed by vendors range around 5–10 dB; although hero prototypes can have losses of less than 2 dB (SKOIH2000).

Crosstalk is the amount of light that gets routed to the wrong output port. We usually distinguish between next-neighbor crosstalk and background crosstalk. Next-neighbor crosstalk is strongly dependent on the exact filter shape. The wider the filter shape, the more light will leak into the ports that are spectrally adjacent. Depending where exactly this crosstalk level is measured, suppression of 15–20 dB between adjacent channels is usually achievable. Background or nonadjacent crosstalk is caused by imperfection and random phase errors (Zirngibl98), crosstalk suppression of 35–45 dB is achievable (SKOIH2000).

There is a tradeoff among filter shape, bandwidth, and insertion loss. Since the modulation rate per channel is limited by the optical filter bandwidth, one wants to make the filter bandwidth as large as possible. This however, leads to next-neighbor crosstalk. In order to have both large filter bandwidth and low adjacent crosstalk, one can attempt to flatten the filter bandwidth that naturally tends to have a Gaussian shape. In order to do so, one has to sacrifice insertion loss (Dragone98). A rule of thumb is that the 3 dB filter bandwidth can be half the optical channel spacing while still offering acceptable adjacent crosstalk levels and low insertion loss.

7.6.2.2 Tunable Lasers

Tunable lasers have been a hot research topic ever since the advent of WDM transmission systems. Their main application is to serve as universal, wavelength-agile transmitters. With fixed transmitters, each wavelength

has to be inventoried separately. Fast tunability is not required for this application, since the wavelength dimension is not used for networking purposes. However, wavelength stability and general telecommunications-grade performance are required. A laser can be tuned in several ways, usually by changing the refractive index in the cavity. This can be accomplished by either changing the temperature or the current density in some part of the lasing cavity. Since temperature tuning is intrinsically slow, we can eliminate it for our fast packet-switch application. Index tuning remains. In principle, geometrically tuned lasers (Zah96, ZJ94) could be very fast too, but unfortunately these lasers are far from being commodity components today.

There are several manufacturers of tunable lasers that use a Distributed Bragg Reflector (DBR) approach to tuning the wavelength. In a DBR, there is a gain section and at least one separate filter section that can be tuned through changing the carrier density through current injection. To achieve a wide tuning range, multiple filter sections have to be used, and the wavelength is determined through use of coarse/fine filter pairs or by choosing filters with slightly different periodicities (Vernier effect) (AKBMY92). The goal of each tunable laser is to have a wide tuning range, continuous tunability, and modal stability. Commercially available tunable lasers have now guaranteed tuning ranges of 40 nm; hero research devices show tunability of up to 100 nm (TYIKT93) which is probably the upper limit because of the gain bandwidth of the semiconductor medium. Nanosecond tuning speed has been demonstrated (LRB2000).

7.6.2.3 Data Modulator

The data modulator is another component that has been driven down in price and up in performance by long-haul WDM systems. The prime candidate for the application here is the $LiNbO_3$ Mach-Zehnder modulator, because of its very large optical bandwidth. These are now commercially available for the 40 Gb/s modulation rate. Other high-speed modulators are based on the electroabsorption effect; their advantage includes smaller size and smaller drive voltage (LBRPD99). However, they only work over a relatively narrow wavelength range, making them less suitable for our application.

7.6.2.4 The Receiver

On the egress side of the fabric, we will need an optical receiver. Optical receivers work over very large wavelength ranges, so we do not have to worry about this. The difference between the receiver here and a receiver in a telecommunication system is that our receiver has to be able to recover phase and clock quickly. Remember that while the laser tunes its color, the receiver will see no light; also, packets from different ingress ports will have different phase positions. The clock problem can be solved by distributing a centralized clock

signal to all ingress and egress ports. It has been shown that phase recovery can be done within a few bits (PGELM99).

7.6.3 Demonstration of Fast Wavelength Switching

A switch fabric similar to the one described above has been demonstrated (Gripp2001). One major concern is optical crosstalk. The are two types of crosstalk, in-band and out-of-band. For in-band crosstalk, the mixing term between the electrical fields beats at a frequency that is within the data signal bandwidth; therefore this signal is not rejected by the electrical receiver. The beat signal for out-of-band crosstalk, on the other hand, has a mixing term outside the signal bandwidth. Its crosstalk is still felt because the detector sees it as an additional signal. The total current through the photodetector is given by $I_{photo} \sim E_s^2 + E_{xtalk}^2 + 2E_sE_{xtalk}\cos\{2\pi(f_s - f_{xtalk})t\}$. In order to have less than 1 dB power penalty, the total intensity of the out-of-band crosstalk should be less than 10 dB down from the signal intensity (HP85). For in-band crosstalk, the isolation has to be doubled on a dB scale, i.e., 20 dB. The worst case is clearly when all channels are on the same wavelength. If we allow the adjacent channel crosstalk and the nonadjacent channel crosstalk both to be 23 dB down from the main signal (so total crosstalk adds up to 20 dB) then the adjacent crosstalk isolation has to be 26 dB and the nonadjacent crosstalk suppression has to be (23 dB + log (# port-3)) or about −40 dB for a 64 channel fabric. While these numbers are within reach of the performance of current AWGs, they can be relaxed somewhat by using crosspolarized channels. It also should be noted that having all the channels on the same wavelength is a very unlikely scenario and could be eliminated through appropriate scheduling.

8. Conclusion

We believe that optics will have a big impact on switching in the future. It has already started happening for Tb/s cross-connects. In these switches, optics primarily solves the transport problem within the switch itself. The transport problem arises because electronic transmission lines are very limited in their capacity × distance product. For packet switches, the same problem will present itself once these packet switches reach Tb/s throughput. Whereas slow optical switching technologies such as micromechanical mirrors are fine for cross-connects, packet switches will require an optical fabric that can switch in a few nanoseconds. Wavelength switching appears so far the most promising approach due to the rapid maturing of various WDM components. The crossover from electrical to optical fabrics may move up in the future because of new optical interconnection technology and reduced size and power consumption of port cards. We believe that transparent routing and switching has a limited future. With the exception of optical A/Ds, transparent wavelength

selective cross-connects would make networks very complex to design and manage. Also, optics is not well suited to replace electrical functions such as header processing and buffering. Thus, switching nodes will likely stay opaque as seen from the transport network, although there will be a lot of photonic transport and switching within the node.

References

AKBMY92 – R. C. Alferness, U. Koren, L. L. Buhl, B. I. Miller, M. G. Young, T. L. Koch, G. Raybon, and C. A. Burrus, Applied Physics Letters, Vol. 60, p. 3209, 1992.

ALC98 – E. Almstrom, S. N. Larsson, and H. Carlden, "Cascadability of Optical Add/Drop Multiplexers," ECOC '98, Madrid, Spain, pp. 589–590, September 20–24, 1998.

BAYVEL2000 – P. Bayvel, "Future High-Capacity Optical Telecommunication Networks," Phil. Trans. R. Soc. London A., 358, pp. 303–329, 2000.

BCG95 - K. Bala, R. R. Cordell, and E. L. Goldstein, "The Case for Opaque Multiwavelength Optical Networks," Proceedings of the IEEE/LEOS Summer Topical Meeting on Technologies for a Global Information Infrastructure, Keystone, CO, Paper WA-4, August 1995.

BDDCL2000 – P. Bernasconi, C. R. Doerr, C. Dragone, M. Cappuzzo, E. Laskowski, and A. Paunescu, "Large N × N Waveguide Grating Routers," IEEE Journal of Lightwave Technology, Vol. 18, No. 7, pp. 985–991, July 2000.

BENES35 – V. E. Benes, "Mathematical Theory of Connecting Networks and Telephone Traffic," New York, Academic, 1935.

CFFKS97 – C. Caspar, H. M. Foisel, R. Freund, U. Kruger, and B. Strebel, "Cascadability of Arrayed-Waveguide Gratings (de)Multiplexers in Transparent Optical Networks," OFC 97, Dallas, Texas, Paper TuE2, February 16–21, 1997.

CONEXANT2000 – Conexant Systems, Inc., Spec Sheet for CX20487.

CPR2000 – T. Cohen, G. Patel, and K. Rothstein, "Design Considerations for Gigabit Backplane Systems," 2000 High-Performance System Design Conference, 2000.

DOERR98 – C. R. Doerr, "Proposed WDM Cross Connect using a Planar Arrangement of Waveguide Grating Routers and Phase Shifters," IEEE Photonic Technology Letters, Vol. 10, pp. 528–530, April 1998.

DNPBL2001 – B. Doshi, R. Nagarajan, G. N. Srinivasa Prasanna, N. Blackwood, Z. Lei, M. Sharma, T. Mueller, T. Hu, and N. Raman, "A Systematic Study of Next Generation Optical Core Networks: Are Flatter Networks Better?," NFOEC, 2001.

DR2000 – B. Davie and Y. Rekhter, "MPLS: Technology and Applications," The Morgan Kaufmann Series in Networking, David Clark, series editor, May 2000.

DRAGONE91 – "An N × N Optical Multiplexer using a Planar Arrangement of Two Star Couplers," IEEE Photonic Technology Letters, Vol. 3, pp. 812–815, 1991.

DRAGONE98 – C. Dragone, "Efficient Techniques for Widening the Passband of a Wavelength Router," IEEE Journal of Lightwave Technology, Vol. 16, No. 10, pp. 1895–1906, October 1998.

DSCLP99 – C. R. Doerr, L. W. Stulz, M. Cappuzzo, E. Laskowski, A. Paunescu, L. Gomez, J. V. Gates, S. Shunk, and A. E. White, "40-Wavelength Add-Drop Filter," IEEE Photonics Technology Letters, Vol. 11, No. 11, p. 1437, November 1999.

FOUQUET2000 – J. E. Fouquet, "Compact Optical Cross-Connect Switch Based on Total Internal Reflection in a Fluid-Containing Planar Lightwave Circuit," Paper TuM1-1, pp. 204–206, OFC 2000, Baltimore, Maryland, March 5–10, 2000.

Gripp2001 – J. Gripp et al., "Demonstration of a 1.2 Tb/s optical packet switch fabric (32 × 40 Gb/s) based on 40 G burst-mode clock-data-recovery fast tunable lasers and high performance N × N AWG," ECOC 2001 post-deadline paper.

GCHCK98 – C. Guillemot, F. Clerot, M. Henry, A. Le Corre, J. Kervarec, A. Dupas, and L. Billes, "Optical Packet Switching for WDM High Speed Backbones," ECOC '98, pp. 83–85, Madrid, Spain, September 20–24, 1998.

GLW2001 – E. L Goldstein, L.-Y. Lin, and J. A. Walker, "Lightwave Micromachines for Optical Networks, Vast Promise Amid Vaster Promises," Optics and Photonics News, March 2001.

GOLDSTEIN98 – E. L. Goldstein, "Transparency and Opacity in Multiwavelength Long-Haul Networks" (Invited), Proceedings of the European Conference on Networks and Optical Communications, IEE and IEEE Communications Society, Manchester, UK, June 1998.

GM95 – C. R. Giles and V. Mizrahi, "Low-Loss Add/Drop Multiplexer for WDM Lightwave Networks," Technical Digest 10th International Conference on Integrated Optics and Optical Fibre Communication, Vol. 3, Paper ThC2-1, 1995.

HP85 – A. M. Hill and D. B. Payne, "Linear Crosstalk in Wavelength-Division-Multiplexed Optical-Fiber Transmission Systems," IEEE Journal of Lightwave Technology, Vol. LT-13, p. 643, 1985.

HUI90 – J. Y. Hui, "Switching and Traffic Theory for Integrated Broadband Networks," Kluwer Academic Publishers, Norwell, MA, 1990.

HUNT99 – C. Hunt, Private Communication.

KAMINOW91 – I. P. Kaminow, "Photonic Local Networks," Chapter 26, Optical Fiber Telecommunications II, 1991.

KCNDS2000 – H. Y. Kim, S. Chandrasekhar, T. Nielsen, C. R. Doerr, L. W. Stulz, L. L. Buhl, R. Monnard, S. Radic, and M. Zirngibl, "1.6 Tb/s (40 × 40 Gb/s) Total Capacity 4-Node Optical Networking with Fully Programmable A/D Devices," Postdeadline Paper, OFC 2000, Baltimore, MD, March 5–10, 2000.

KESHAV97 – S. Keshav, "An Engineering Approach to Computer Networking: ATM Networks, the Internet and the Telephone Network," Addison-Wesley, 1997.

KLKCJ98 – S. Y. Kim, S. B. Lee, S. W. Kwon, S. S. Choi, and J. Jeong, "Channel-Switching Active Add/Drop Multiplexer with Tunable Gratings," Electronics Letters, Vol. 34, No. 1, pp. 104–105, January 8, 1998.

KLST2000 – K. Kar, T. V. Lakshman, D. Stiliadis, and L. Tassiulas, "Reduced Complexity Input Buffered Switches," Proceedings of Hot Interconnects VIII, August 2000.

LBRPD99 – O. Leclerc, P. Brindel, D. Rouvillain, E. Pincemin, B. Darry, E. Desurvire, C. Duchet, E. Boucherez, and S. Bouchoule, "40 Gbit/s Polarization-Insensitive and

Wavelength-Independent InP Mach-Zehnder Modulator for All-Optical Regeneration," Electronics Letters, Vol. 35, No. 9, April 29, 1999.

LRB2000 – O. A. Lavrova, G. Rossi, and D. J. Blumenthal, "Rapid Tunable Transmitter with Large Number of ITU Channels Accessible in Less than 5 ns," ECOC 2000, Munich, Germany, 2000.

LUCENT2001 – http://www.lucent.com/serviceprovider/h/feature_1112.html

MCKEOWN2001 – N. McKeown, "How Scalable is the Capacity of (Electronic) IP Routers?," OFC 2001, Anaheim, California, March 17–22, 2001.

MSATT99 – Y. Maeno, Y. Suemura, S. Araki, S. Takahashi, A. Tajima, H. Takahashi, and N. Henmi, "A Bandwidth-Effective Terabit/s Optoelectronic Packet-Switch with Skew-Insensitive Synchronisation Scheme," ECOC '99, p. I-118, Nice, France, September 26–30, 1999.

MSCCI2000 – T. Murphy, S.-Y. Suh, B. Comissiong, A. Chen, R. Irvin, R. Grencavich, and G. Richards, "A Strictly Non-Blocking 16 × 16 Electrooptic Photonic Switch Module," ECOC 2000, Munich, Germany, 2000.

MSTH98 – Y. Maeno, Y. Suemura, A. Tajima, and N. Henmi, "A 2.56-Tb/s Multi-wavelength and Scalable Switch-Fabric for Fast Packet-Switching Networks," IEEE Photonic Technology Letters, Vol. 10, No. 8, pp. 1180–1182, August 1998.

NTS99 – T. Nakazawa, S. Taniguchi, and M. Seino, "Ti : $LiNbO_3$ Acousto-Optic Tunable Filter (AOTF)," Fujitsu Science Technology Journal, Vol. 35, No. 1, pp. 107–112, July 1999.

PGELM99 – I. D. Phillips, P. Gunning, A. D. Ellis, J. K. Lucek, D. G. Moodie, A. E. Kelly, and D. Cotter, "10-Gb/s Asynchronous Digital Optical Regenerator," IEEE Photonic Technology Letters, Vol. 11, No. 7, pp. 892–894, July 1999.

PMFCH2000 – V. Plickert, L. Melchior, M. Franke, R. Carl, P. Hildebrandt, and N. Staudemeyer, "Progress in Packaging of Parallel Optical Interconnects," IEEE 2000 ECOC 2000.

RHK2001 – RHK Telecommunications Industry Analysis, "WDM and Optical Networks: Optical Core Switch Systems," Technology Report, January 2001.

RTBDC2000 – S. Rotolo, A. Tanzi, S. Brunazzi, D. DiMola, L. Cibinetto, M. Lenzi, G.-L. Bona, B. J. Offrein, F. Horst, R. Germann, H. W. M. Salemink, and P. H. Baechtold, "Integrated Optic Tunable Add-Drop Filters for WDM Ring Networks," IEEE Journal of Lightwave Technology, Vol. 18, No. 4, pp. 569–578, April 2000.

RYF2001 – R. Ryf, J. Kim, J. P. Hickey, A. Gnauck, D. Carr, F. Pardo, C. Bolle, R. Frahm, N. Basavanhally, C. Yoh, D. Ramsey, R. Boie, R. George, J. Kraus, C. Lichtenwalner, R. Papazian, J. Gates, H. R. Shea, A. Gasparyan, V. Muratov, J. E. Griffith, J. A. Prybyla, S. Goyal, C. D. White, M. T. Lin, R. Ruel, C. Nijander, S. Arney, D. T. Neilson, D. J. Bishop, P. Kolodner, S. Pau, C. Nuzman, A. Weis, B. Kumar, D. Lieuwen, V. Aksyuk, D. S. Greywall, T. C. Lee, H. T. Soh, W. M. Mansfield, S. Jin, W. Y. Lai, H. A. Huggins, D. L. Barr, R. A. Cirelli, G. R. Bogart, K. Teffeau, R. Vella, H. Mavoori, A. Ramirez, N. A. Ciampa, F. P. Klemens, M. D. Morris, T. Boone, J. Q. Liu, J. M. Rosamilia, and C. R. Giles, "1296-Port MEMS

Transparent Optical Crossconnect with 2.07 Petabit/s Switch Capacity," OFC 2001. Postdeadline, Anaheim, California, March 17–22, 2001.

SALEH2000 – A. A. M. Saleh, "Transparent Optical Networking in Backbone Networks," Invited Talk, OFC 2000, Baltimore, MD, March 5–10, 2000.

SCDJ99 – N. Le Sauze, D. Chiaroni, E. Dotaro and A. Jourdan, "Optical Packet Versus Circuit Switching in Metropolitan Networks," ECOC 99, pp. I14–I15. Nice, France. September 26–30, 1999.

SCBBJ96 – D. A. Smith, R. S. Chakravarthy, Z. Bao, J. E. Baran, J. L. Jackel, A. D'Alessandro, D. J. Fritz, S. H. Huang, X. Y. Zou, S. M. Hwang, A. E. Willner, and K. D. Li, "Evolution of the Acousto-Optic Wavelength Routing Switch," IEEE Journal of Lightwave Technology, Vol. 14, pp. 1005–1019, 1996.

SGS96 – J. M. Simmons, E. L. Goldstein, and A. M. Saleh, "Quantifying the Benefit of Wavelength Add-Drop in WDM Rings with Distance Independent and Dependent Traffic," IEEE Journal of Lightwave Technology, Vol. 17, pp. 48–57, January 1996.

SKOIH2000 – A. Sugita, A. Kaneko, K. Okamoto, M. Itho, A. Himeno, and Y. Ohmori, "Very Low Insertion Loss Arrayed-Waveguide Grating with Vertically Tapered Waveguides," IEEE Photonic Technology Letters, Vol. 12, No. 9, pp. 1180–1182, September 2000.

SMITHGALL2001 – G. Smith, private communication.

STILIADIS2000 – D. Stiliadis, private communication.

SYHY97 – K. Sasayama, Y. Yamada, K. Habara, and K.-I. Yukimatsu, "FRONTIER-NET: Frequency-Routing-Type Time-Division Interconnection Network," IEEE Journal of Lightwave Technology, Vol. 15, No. 3, pp. 417–429, March 1997.

TGNS98 – R. W. Tkach, E. L. Goldstein, J. A. Nagel, and J. L. Strand, "Fundamental Limits of Optical Transparency," OFC 98, San Jose, California, February 22–27, 1998

TYIKT93 – Y. Tohmori, Y. Yoshikuni, H. Ishii, F. Kano, T. Tamamura and Y. Kondo, "Over 100 nm Wavelength Tuning in Superstructure Grating (SSG) DBR Lasers," Electronics Letters, Vol. 29, No. 4, pp. 352–354, February 18, 1993.

VS91 – A. R. Vellekoop and M. K. Smit, "Four-Channel Integrated-Optic Wavelength Demultiplexer with Weak Polarization Dependence," IEEE Journal of Lightwave Technology, Vol. 9, No. 3, pp. 310–314, March 1991.

WASG96 – R. E. Wagner, R. C. Alferness, A. A. M. Saleh, and M. S. Goodman, "MONET: Multiwavelength Optical Networking," IEEE/OSA JLT/JSAC Special Issue on Multiwavelength Optical Technology and Networks, Vol. 14, No. 6, pp. 1349–1355, June 1996.

WMDZ99 – G. Wilfong, B. Mikkelsen, C. R. Doerr and M. Zirngibl, "WDM Cross-Connect Architectures with Reduced Complexity," IEEE Journal of Lightwave Technology, Vol. 17, No. 10, October 1999.

YANG2000 – W. Yang, private communication.

YYIKT93 – Y. Tohmori, Y. Yoshikuni, H. Ishii, F. Kano, T. Tamamura, and Y. Kondo, "Over 100 nm Wavelength Tuning in Superstructure Grating (SSG) DBR Lasers," Electronics Letters, Vol. 29, No. 4, pp. 352–354, February 18, 1993.

ZAH96 – C. E. Zah, et al., "Wavelength Accuracy and Output Power of Multiwavelength DBF Lasers with Integrated Star Couplers and Optical Amplifiers," IEEE Photonics Technology Letters, Vol. 8, pp. 864–866, 1996.

ZIRNGIBL98 – M. Zirngibl, "Expected Performance of Waveguide Grating Routers in the Presence of Random Phase and Amplitude Errors," Optics and Laser Technology, Vol. 30, No. 7, pp. 419–420, April 1998.

ZJ94 – M. Zirngibl and C. H. Joyner, "12 Frequency WDM Laser Based on a Transmissive Waveguide Grating Router," Electronics Letters, Vol. 30, No. 9, p. 701, April 28, 1994.

ZT98 – W. De Zong and R. S. Tucker, "Wavelength Routing-Based Photonic Packet Buffers and Their Applications in Photonic Packet Switching Systems," IEEE Journal of Lightwave Technology, Vol. 16, No. 10, pp. 1737–1745, October 1998.

Chapter 9 | Planar Lightwave Devices for WDM

Christopher R. Doerr

Bell Laboratories, Lucent Technologies, Holmdel, New Jersey

1 Why Waveguides?

With the proliferation of free-space optics devices such as thin-film filters, microelectromechanical (MEMs) cross-connects,[1] and liquid-crystal wavelength-selective switches,[2] is not the field of integrated optics, the field of devices consisting of planar arrangements of waveguides and active components on one substrate, dying? Why confine oneself to two dimensions when one can use all three in bulk optics? One can achieve very low insertion loss with bulk optics.

The field of integrated optics is like a giant glacier, a mass of ice obliterating obstacles one by one. It is often true that one can initially obtain superior performance with bulk optics. This is because bulk optics have almost no constraints. One can handpick the best components and spend as much time as one wishes optimizing putting them together. But the words "hand" and "time" in the former sentence explain why integrated optics is so important. In integrated optics, all of the components are already together, and they can be done tens, hundreds, or thousands at a time. Often all that is left to do is to attach the optical fibers and place in a package. Integrated optics is about saving cost.[3]

This chapter covers photonic integrated circuits that are or could be used in wavelength-division multiplexed (WDM) optical networks. It is organized in terms of the currently most successful material systems: silica, polymer, silicon, indium phosphide, and lithium niobate. The field is so immense that the author has chosen to concentrate mostly on devices that use arrays of waveguides. This chapter unfortunately does not cover all of integrated optics, leaving out microring resonators, photonic bandgap structures, devices not specifically usable in WDM systems, fixed-wavelength lasers, and others.

2 Silica

Silica (i.e., glass) waveguide circuits are ideal for fiberoptic systems. The refractive index matches that of fiber, the propagation loss is very low (<0.02 dB/cm), and the material is highly durable.

OPTICAL FIBER TELECOMMUNICATIONS,
VOLUME IVA

ISBN: 0-12-395172-0

2.1 WAVEGUIDE STRUCTURE

The integrated optics designer must choose a layer structure for the waveguides. For silica, there are conventionally two main choices: to have a slab or not, as shown in Fig. 1. To make waveguides in silica, first a layer of silica, which will form the lower cladding, is deposited onto a substrate, usually a 5-inch-diameter silicon wafer or sometimes a glass wafer. Then the core layer, which has a slightly higher index, usually by doping the silica with phosphorous, but sometimes germanium, is deposited. This silica must have a lower melting temperature than the lower cladding. Then the wafer is patterned with photoresist and etched using a reactive ion etcher.

Having a slab simply means that the core is not etched all the way through. This requires more process control. The slab allows the light to still be guided vertically when no pattern is present. This allows certain device features such as a refractive lens. However, it also increases the amount of stray light in the outputs, possibly increasing the crosstalk of filters. Most silica devices do not employ a slab, although most other material systems do (see indium phosphide and silicon-on-insulator).

Finally another layer of silica is deposited over everything. This silica must have an even lower melting temperature than the previous two layers.

To determine the desired waveguide thickness, one can use the effective index method. The idea behind the effective index method is shown in Fig. 2. First, one finds the propagation constant of the field if the core was present infinitely in both lateral directions, i.e., a slab waveguide. The propagation constant is the number of radians per unit distance that the field advances in the waveguide. It is given by

$$\beta = \frac{2\pi n_{\text{eff}}}{\lambda_0} \tag{1}$$

where λ_0 is the wavelength as it would be in vacuum, and n_{eff} is the effective index. Also note that β is just another name for k_z. k is also used for propagation constants, and z is usually the direction of travel along the waveguide. For asymmetric three-layer slab, one can find β by solving the following transcendental equation for TE-polarized light:

$$\sqrt{\beta^2 - k_{\text{clad}}^2} = \sqrt{k_{\text{core}}^2 - \beta^2}\tan\left(\sqrt{k_{\text{core}}^2 - \beta^2}\frac{h}{2}\right) \tag{2}$$

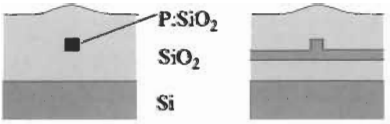


Fig. 1 Silica waveguide structure.

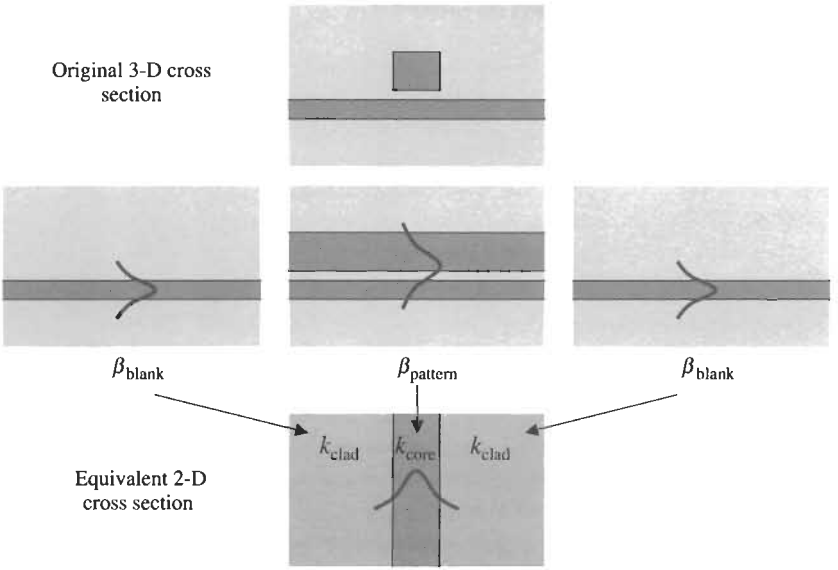


Fig. 2 The effective index method for determining waveguide modes.

and for TM-polarized light:

$$\sqrt{\beta^2 - k_{clad}^2} = \frac{k_{clad}^2}{k_{core}^2}\sqrt{k_{core}^2 - \beta^2}\tan\left(\sqrt{k_{core}^2 - \beta^2}\frac{h}{2}\right) \tag{3}$$

where h is the height of the slab (see [4] for a good tutorial). If more than one solution (β real) exists, then the waveguide is multimoded vertically, when the waveguide is very wide. This is usually not a problem. The designer will want to use only the zeroth order mode, so the highest β is used in the following. This β will be $\beta_{pattern}$. The effective index method is a scalar approach, so it works best when $k_{clad} \approx k_{core}$. Then it does not matter much whether the TE or the TM equation is used. Then one likewise finds β for the field where no core is present (although a slab may be). This β will be β_{blank}. One then can use these two effective β ($\beta_{pattern}$ and β_{blank}) and assume now that the cores are infinitely tall to find the actual β, using Eq. 2 or Eq. 3 with $k_{clad} = \beta_{blank}$, $k_{core} = \beta_{pattern}$, and $h = w =$ width. The normalized effective index step is given by

$$\left(\frac{\Delta n}{n}\right)_{eff} = \frac{2\left(\beta_{pattern} - \beta_{blank}\right)}{\beta_{pattern} + \beta_{blank}} \tag{4}$$

Note that there is also the index step between the core and cladding:

$$\frac{\Delta n}{n} = \frac{2\left(k_{core} - k_{clad}\right)}{k_{core} + k_{clad}} \tag{5}$$

Waveguide fabricators usually mention the index step, while waveguide designers usually mention the effective index step (a smaller number). A typical index step of a silica waveguide is 0.8%. This typically means an effective index step of ~0.57%. A convenient non-dimensional parameter for the modal behavior of a waveguides is the V-number,

$$V = \frac{2\pi n w}{2\lambda_0}\sqrt{2\left(\frac{\Delta n}{n}\right)_{\text{eff}}} \tag{6}$$

The waveguide is monomode provided $V < \pi/2$.

The main criteria for choosing the layer thickness and the index step are the bending radius, the matching of the waveguide mode to optical fiber, and the tolerance and loss of couplers. The bending loss in dB per radian is given approximately by the following equation for waveguides that are very wide, i.e., whispering gallery bends:[5]

$$\frac{20}{\ln(10)\sqrt{2(\Delta n/n)_{\text{eff}}}} \exp\left\{-\frac{4}{3}\left[Z^2 - 2.338 + \frac{1}{Z}\left(1 + \frac{0.65}{Z^2}\right)\right]^{1.5}\right\} \tag{7}$$

where

$$Z = \left[2\left(\frac{\Delta n}{n}\right)_{\text{eff}}\right]^{1/2}\left(\frac{2\pi n R}{2\lambda_0}\right)^{1/3} \tag{8}$$

To match the waveguide mode well to fiber, the waveguide is usually multimoded both horizontally and vertically. When higher-order modes can be guided, the designer must include spatial mode filters at strategic points to strip them off. The higher the effective index step, the closer waveguides must come to have a directional coupler of a given length, making fabrication more demanding. Also, a higher index step tends to have more scattering loss.

The effective index method works well only for structures that have a small index step. For structures with high index steps, such as air-clad semiconductor waveguides, one has to use a more complicated computer program, called a mode solver, that solves for the guided modes using finite element methods, Fourier transform methods or the method of lines[6–8] or launch an arbitrary field distribution down a long waveguide of the desired cross-section using a beam propagation method (BPM) program and wait until the nonguided mode components have radiated away.[9]

Note that a waveguide that has different dimensions in the x- and y-directions often has birefringence, i.e., a different refractive index for quasi-TE and quasi-TM modes (for a mode guided in two dimensions, there is no such thing as a pure TE- or a pure TM-guided mode). However, this shape birefringence is generally much weaker than the strain birefringence in the

glass for silica waveguides on a silicon substrate. When a silicon substrate is used, upon cooling after silica deposition, the silica is under enormous compressive strain because of the larger thermal coefficient of expansion of silicon than pure silica. Nearly all of the birefringence comes from the upper cladding silica pushing on the waveguide cores. The typical birefringence of silica waveguides is 0.3 nm/1550 nm $= 2 \times 10^{-4}$. Several methods to mitigate this birefringence have been used: use dopants in the upper cladding silica to make its thermal expansion coefficient match that of silicon;[10] use a silica substrate; use stress-relieving grooves;[11] or vary the width of the core,[12] which in some cases can change the birefringence, making it possible to construct interferometers that have the same effective path-length difference for TE- and TM-polarized lights.

2.2 HOW TO DECIDE THE WAVEGUIDE TOPOLOGY

Next, the designer must determine the topological pattern. Generally, the designer wants to keep the waveguides as wide as possible in order to minimize the effects of pattern width on the complex transmissivity of the light. Change in propagation constant per change in waveguide width for the fundamental mode decreases with increasing waveguide width. The drawback to a wide waveguide is that the zeroth order mode stays in phase with higher-order modes longer, increasing the chance of coupling between them.

Even simple waveguide structures, such as a directional coupler, are too complicated for analytical solutions (even with many approximations), and so a computer program, a BPM program that can simulate the light propagating in the waveguides, is indispensable. The two most popular BPMs are the Fourier transform [Fast-Fourier transform (FFT)] and finite difference (FD) BPMs. They are explained well in [13]. Roughly, the wave equation is a partial differential equation; FFT-BPM solves it using Fourier transforms, while FD-BPM solves it by approximating it as a difference equation. FFT BPM can handle large propagation angles but not large index steps. FD BPM cannot handle large propagation angles (the light must propagate within a few degrees of the direction of the calculation; there are modified FD-BPMs that handle larger angles at the expense of more computation time)[14] but can handle large index steps. FD-BPM generally executes much faster than FFT BPM,[15] and nearly all commercial BPM programs use FD-BPM. We will give a simple derivation of the FFT method here; it is instructive and easy to implement.

In FFT BPM, the actions of field distortion in the spatial domain due to refractive index variation and plane wave propagation in the angular domain are treated separately for each small step in the propagation direction, z. For each step Δz, the complex optical field $u(x, y)$, where x and y are the dimensions transverse to the plane of the circuit, is multiplied by

$$\exp[j\Delta z k(x, y)] \tag{9}$$

Then $u(x,y)$ is simply Fourier transformed to get the angular spectrum $U(k_x, k_y)$. $U(k_x, k_y)$ is multiplied by

$$\exp(jk_z \Delta z) \tag{10}$$

where

$$k_x^2 + k_y^2 + k_z^2 = k^2 \tag{11}$$

and is then inverse Fourier transformed, z is advanced by Δz, and the cycle repeats until the end of the propagation region.

Both FFT- and FD-BPM have the issue of spatial quantization—the structure must be artificially snapped to the simulation grid. A very small Δx grid must usually be chosen. Also, FFT- and FD-BPM can handle arbitrary index profiles, but in integrated circuits the waveguides are usually step-index, i.e., digitally constrained.

The elimination of this quantization is the idea behind a new type of BPM, sinc-BPM.[16] Sinc-BPM can be hundreds of times faster than FFT- and FD-BPM for many structures of interest. We will also give a derivation of sinc-BPM.

The inputs to sinc-BPM are the waveguide location, d, the width w, and center-to-center spacing a between a periodic array (sinc-BPM assumes periodic arrays of waveguides).

Because we are dealing with step index waveguides, we can modify Eq. 9 to be

$$\{f(x,y)\{\exp[j\Delta z(k_{\text{core}} - k_{\text{clad}})] - 1\} + 1\} \exp(j\Delta z k_{\text{clad}}) \tag{12}$$

where $f(x,y)$ is a function that is either 0 or 1.

We can express any waveguide distribution as a sum of waveguide arrays. If there is only one waveguide, then a can be a distance at which we are sure the coupling between waveguides is negligible. If we take the Fourier transform of Eq. 12, then for each set of waveguides spaced periodically in two dimensions, we get

$$\delta(k_x, k_y) + [\exp(jk\Delta z \Delta n/n) - 1] \frac{w_x}{a_x} \frac{\sin(k_x w_x/2)}{k_x w_x/2} \exp\left(j2\pi \frac{d_x}{a_x}\right)$$
$$\times \frac{w_y}{a_y} \frac{\sin(k_y w_y/2)}{k_y w_y/2} \exp\left(j2\pi \frac{d_y}{a_y}\right) \sum_{p,q} \delta\left(k_x - \frac{2\pi p}{a_x}, k_y - \frac{2\pi q}{a_y}\right) \tag{13}$$

No approximations have been made from FFT-BPM. Now we can convolve $U(k_x, k_y)$ with Eq. 13 instead of Fourier transforming u and back. To further speed the computation, we abruptly truncate the sinc function after $ka/25$ terms, empirically chosen, from the center. Because of the truncation, the total

power conservation is imperfect. C-language code for the two-dimensional case, i.e., when $a_y = w_y$, in a format that can be called from MATLAB is given below:

```
#include <math.h>
#include "mex.h"
#define PI 3.14159265358979
void mexFunction(int nlhs, mxArray *plhs[], int nrhs,
                 const mxArray *prhs[])

{
    double Dlocd,Dkx,Dz,k,ampr,ampi,Dnn,lockx,
           mult1,mult2,tempd,xlimit,Dx,kx,temp1,temp2;
    double *ufr,*ufi,*ufnewr,*ufnewi,*tfr,*tfi,*w,*aa,*d;
    short Dloc,loc,h,temp,N,imp,QN1,QN3,lim,cnt;
    int numsteps;

    if (nrhs != 8) {
        mexErrMsgTxt("Eight input arguments required.");
    } else if (nlhs > 2) {
        mexErrMsgTxt("Only two output arguments allowed.");
    }

    /* get input parameters */
    ufr = mxGetPr(prhs[0]);
    ufi = mxGetPi(prhs[0]);
    N = mxGetN(prhs[0]);
    xlimit = mxGetScalar(prhs[1]);
    Dz = mxGetScalar(prhs[2]);
    k = mxGetScalar(prhs[3]);
    w = mxGetPr(prhs[4]);
    numsteps = mxGetN(prhs[4]);
    aa = mxGetPr(prhs[5]);
    Dnn = mxGetScalar(prhs[6]);
    d = mxGetPr(prhs[7]);

    /* define output parameters */
    plhs[0] = mxCreateDoubleMatrix(1,N,mxCOMPLEX);
    ufnewr = mxGetPr(plhs[0]);
    ufnewi = mxGetPi(plhs[0]);
    plhs[1] = mxCreateDoubleMatrix(1,N,mxCOMPLEX);
    tfr = mxGetPr(plhs[1]);
    tfi = mxGetPi(plhs[1]);

    /* calculate parameters */
    QN1 = N/4;
    QN3 = 3*N/4;
    Dx = 2*xlimit/N;
```

```
Dkx = 2*PI/Dx/N;
mult1 = sin(k*Dz*Dnn);
mult2 = (cos(k*Dz*Dnn) - 1);
for (h = 0;h < N;h++){
    kx = -PI/Dx + Dkx*h;
    tempd = k*k - kx*kx;
    if (tempd > 0){
        tfr[h] = cos(sqrt(tempd)*Dz);
        tfi[h] = sin(sqrt(tempd)*Dz);
    } else {
        tfr[h] = exp(-sqrt(-tempd)*Dz);
        tfi[h] = 0;
    }
}

/* initialize ufnewr */
for (h=0;h < N;h++){
    ufnewr[h] = ufr[h];
    ufnewi[h] = ufi[h];
}

/* beam propagation */
for (cnt = 0;cnt<numsteps;cnt++){
    for (h=0;h < N;h++){
        ufr[h] = ufnewr[h]*tfr[h] - ufnewi[h]*tfi[h];
        ufi[h] = ufnewi[h]*tfr[h] + ufnewr[h]*tfi[h];
        ufnewr[h] = ufr[h];
        ufnewi[h] = ufi[h];
    }
    lim = (short)(0.04*k*aa[cnt]);
    Dlocd = 2*PI/aa[cnt]/Dkx;
    for (imp=-lim;imp < lim+1;imp++){
        loc = (short) Dlocd*imp;
        if (imp != 0){
            ampr = sin(w[cnt]*Dkx*loc/2)/(PI*imp);
        } else {
            ampr = w[cnt]/aa[cnt];
        }
        temp1 = cos(d[cnt]*Dkx*loc);
        temp2 = sin(d[cnt]*Dkx*loc);
        ampi = ampr*(mult1*temp1 + mult2*temp2);
        ampr = ampr*(mult2*temp1 - mult1*temp2);
        for (h=QN1;h < QN3;h++){
            temp = h - loc;
            if (temp < 0) temp = temp + N;
            if (temp > N-1) temp = temp - N;
            ufnewr[h] = ufnewr[h] + ufr[temp]*ampr
                        - ufi[temp]*ampi;
```

```
                ufnewi[h] = ufnewi[h] + ufr[temp]*ampi
                                + ufi[temp]*ampr;
            }
        }
    }
    return;
}
```

2.3 *PROPAGATION ACROSS FREE-SPACE REGIONS*

One would generally not want to simulate an entire complex circuit using BPM. It is too time-consuming and can lead to errors due to the accumulation of numerical round-offs. However, many circuits contain regions with waveguides well separated from each other so there is no mutual coupling, and many contain free-space regions, regions where the pattern is so wide that the field is not affected by the pattern boundaries (i.e., slab waveguides). For the uncoupled waveguides, each of length L, one can calculate the guided mode shapes and propagation constants β and simply phase shift each mode by βL. For the free-space regions, the field is guided vertically but has free propagation horizontally. If we have two waveguides terminating at locations A and B on opposite sides of the free-space boundary, then we can calculate the transmissivity (i.e., the similarity) between them. If the known fields are as shown in Fig. 3, then the transmissivity amplitude is given by

$$\left\{\int_{-\infty}^{\infty} u_B^*(x)\exp\left[jkx\left(\theta_B - \angle AB + \frac{\pi}{2}\right)\right]\right.$$
$$\left.\times F^{-1}\left\{F\left\{u_A(x)\exp\left[-jkx\left(\theta_A - \angle AB + \frac{\pi}{2}\right)\right]\right\}\exp\left(j\sqrt{k^2 - k_x^2}|AB|\right)\right\}dx\right\}$$
$$\Big/\left\{\sqrt{\int_{-\infty}^{\infty}|u_A(x)|^2dx\int_{-\infty}^{\infty}|u_B(x)|^2dx}\right\} \tag{14}$$

F is the Fourier transform

$$F\{g(x)\} = \int_{-\infty}^{\infty} g(x)e^{jxk_x}dx \tag{15}$$

Equation 14 is derived from breaking up the field u_A into a sum of plane waves (taking the Fourier transform), propagating those plane waves, converting back to the spatial domain, and calculating the overlap integral with u_B. An easy way to find AB given that A and B are known in terms of distance and angle is to use complex algebra (i.e., write A as $re^{j\phi}$ and subtract B from A).[17] There is a free-space region in the middle of the star coupler, which is discussed next.

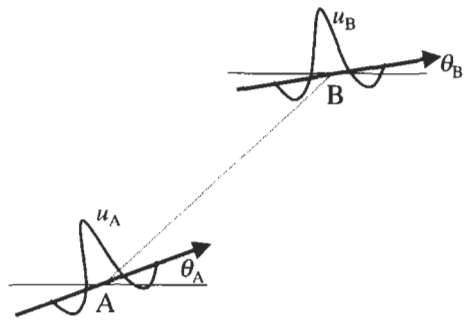


Fig. 3 Field distributions for transmissivity calculation between them.

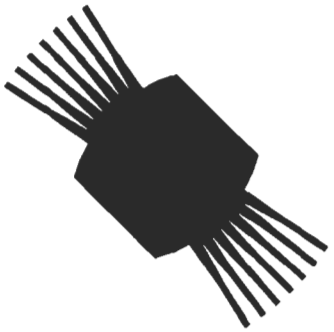


Fig. 4 The star coupler.

2.4 THE STAR COUPLER

A key element for redistributing light among waveguides is the star coupler.[18] The basic star coupler is a planar arrangement of waveguide cores that converge to a point, but before reaching the point, they are terminated on an arc in a free-space slab, this arc on which is the convergence point for a waveguide array diverging in the opposite direction, as shown in Fig. 4. The remarkable feature of the star coupler is that it can theoretically have zero excess loss and even power distribution among the output ports when one input port is illuminated, and one can nearly achieve this perfection with practical design parameters.[19] For such a conventional star coupler,

$$\mathrm{AB} = R\left(e^{-j\theta_1} + e^{j\theta_2} - 1\right) \tag{16}$$

which can be substituted into Eq. 14 to find the transmissivities. u_A and u_B are found by calculating the zero-order mode of a waveguide in the array when the waveguides are uncoupled and propagating this mode to the edge of the free-space region via BPM $u_A = u_1$, and $u_B = u_2^*$. θ_1 and θ_2 are the angular positions of the ports of interest as measured from the star coupler axis of symmetry. $\theta_A = -\theta_1 - \pi/2$ and $\theta_B = \theta_2 - \pi/2$. If θ_1 and $\theta_2 \ll 1$, then $|\mathrm{AB}| \approx R - R\theta_1\theta_2$ and $\angle\mathrm{AB} \approx \theta_2 - \theta_1$.

Mutual coupling in the converging waveguide arrays causes aberrations (a star coupler acts like two lenses separated by a focal length).[20] The mutual coupling is inevitable in a low-loss star coupler because in such a case the gaps between the waveguides are as small as possible at the edges of the free-space region (see Sec. 2.5). The aberrations will cause the response to become rectangular, filling the central Brillouin zone (central diffraction order) given by the angle λ/a, where a is the center-to-center spacing between waveguides at the free-space region boundary. The phase aberrations have approximately the following nature:

$$\phi(m) = \alpha_1 \cos\left[\frac{2\pi}{kaN}\left(m - \frac{M+1}{2}\right)\right] + \alpha_2 \cos\left[\frac{2\pi}{N}\left(m - \frac{M+1}{2}\right)\right] \quad (17)$$

where $\phi(m)$ is the phase of the field in waveguide m on the receiving side of the star coupler, N is the number of waveguides in one Brillouin zone, and M is the number of waveguides present.

The easiest way to understand what happens in an array of mutually coupled waveguides is to think in terms of the local normal eigenmodes of the array. In other words, if one looks at a cross-section perpendicular to the direction of propagation through the array, one has a periodic, sawtooth shape of index vs. space. The eigenmodes of a periodic index distribution are "Bloch modes,"[19] of which there is an infinite number. They are equivalent to plane waves in free space in that each travels in a different direction (i.e., has a straight phase front at a different angle). However, unlike in free space, Bloch modes are unique only in a travel direction interval of $\pm\lambda/2a$, where a is the center-to-center spacing between waveguides. A Bloch mode traveling at angle $P\lambda/a + d$ is the same as one traveling at d, where P is an integer.

The more strongly the waveguides are coupled, the larger the propagation constant difference between Bloch modes. A star coupler is most efficient when it adiabatically transforms each Bloch mode into just one plane wave in the free-space region. For example, suppose we start with energy in the fundamental mode of just one waveguide in the array when the waveguides are extremely weakly coupled. This means that we have equal energy in all the Bloch modes and they are all in phase. In the weakly coupled case, the propagation constants of all the Bloch modes are the same, and so the mode propagates undistorted and stays in one waveguide. As it approaches the free-space region, the waveguides get closer together, and so the Bloch modes start to have diverging propagation constants. The Bloch modes become out of phase, and so the energy spreads to other waveguides. Finally, at the free-space region boundary, we have an abrupt transition from Bloch modes to plane waves. The less abrupt the transition, the more each Bloch mode couples to only one plane wave traveling in its direction. Otherwise, energy is scattered to plane waves traveling in directions outside $\pm\lambda/2a$ and is lost in conventional star coupler designs. Also, the phase of each transmissivity, through the entire star coupler, is partially periodic with the Brillouin zone because the Bloch

modes start in phase and then at the free-space boundary each Bloch mode couples to a set of plane waves spaced in angle by a Brillouin zone width. The phase has a periodic component (period of λ/a, where a is the spacing at the end of the waveguide array) and a shifted phase center component (out-of-focus component where the propagations primarily start diverging which is before reaching the free-space region), which are the second and first parts of Eq. 17, respectively.

α_1 and α_2 are constants that can be found using BPM: one starts with the waveguide fundamental mode in one of the waveguides on the transmitting side of the star coupler and propagates until reaching the boundary, using BPM. Then one does the same thing for the receiving side, also finding the field at the star coupler boundary as excited by a single waveguide. Since the waveguide arrays are generally uniform, one can assume that the field at the star coupler boundary is the same regardless of which waveguide in the array was excited. Then one can find the transmissivity from one waveguide of the transmitting array to all of the receiving waveguides using Eq. 14. Then, taking the angle of this [i.e., $\tan^{-1}(\text{Im/Re})$], one can fit Eq. 17 to it to find α_1 and α_2. Finally, one can adjust the path lengths in the waveguide arrays attached to the star coupler and/or shift the aiming points radially to correct for these aberrations. It is convenient to use sinc-BPM for this calculation, because this calculation is very sensitive to spatial quantization and FFT- and FD-BPM often give inaccurate results unless the grid size perpendicular to the direction of propagation is chosen to be extremely small, giving very long computation times. For example, consider a 50×50 star coupler with $\Delta n/n = 0.5\%$, $n = 1.45$, and waveguide widths that start at 6 and taper to $7\,\mu\text{m}$ at the free-space boundary with a parabolic taper length of $180\,\mu\text{m}$. If one launches into one waveguide of the star coupler where the waveguides are well uncoupled and propagates to the free-space boundary, and if one launches into a neighboring waveguide and propagates to the free-space boundary, then since the initial uncoupled modes are orthogonal, the final two modes should remain orthogonal. Figure 5 shows such a final computed orthogonality between the propagated mode and its shifted self for the star coupler described above for propagations by FFT-, FD-, and sinc-BPM for various Δx and N. The x-window size of $900\,\mu\text{m}$ is kept constant. The normalized orthogonality of the final modes falls below 1% for sinc-BPM with a $\Delta x = 1.8\,\mu\text{m}$, whereas Δx must be decreased below $0.007\,\mu\text{m}$ before it falls below 1% for both FFT- and FD-BPM. Thus, for simulating this star coupler with reasonable accuracy, using the number of complex multiples per step Δz as a benchmark, sinc-BPM is ~2000 and ~660 times faster than plain FFT- and FD-BPM, respectively.

2.5 VERTICAL TAPERS AND SEGMENTATION

A plain $N \times N$ star coupler can have high excess loss, and the excess loss for the central inlets is higher the smaller the N. A plain 2×2 star coupler in typical silica waveguides has ~1.2 dB excess loss. Also, for a large star coupler,

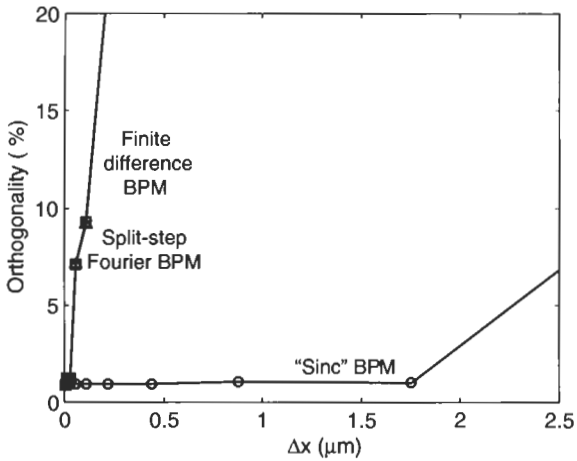


Fig. 5 Final orthogonality between modes that are initially orthogonal after propagation through a mutually coupled waveguide array in a star coupler as computed by FFT-, FD-, and sinc-BPM. From Ref. 16.

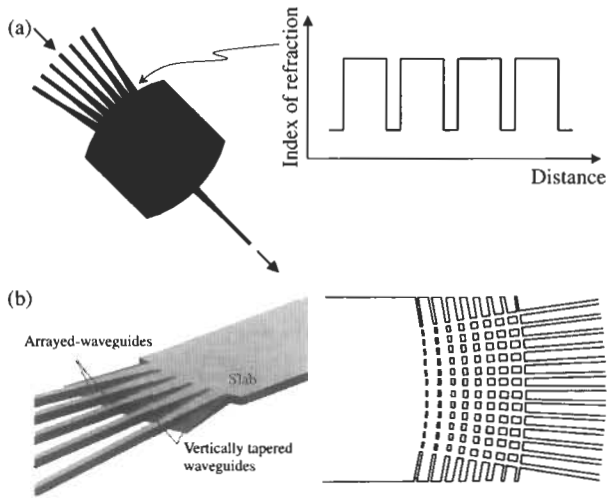


Fig. 6 Reduction of loss in the star coupler: (a) conventional case and (b) left vertical tapering and right segmentation. From Refs. 22 and 23.

the edge inlets can have excess losses approaching 3 dB. The transmissivity magnitude is essentially given by the overlap of the local normal mode of the index distribution right at the boundary of the star coupler with a slab mode (a two-dimensional plane wave) (see Fig. 6). The gaps between waveguides at the free-space boundary are finite because of lithography limitations (about 2.5 μm in silica). Even worse, for relatively small N, the convergence of the waveguides is fast, so the gaps are effectively even larger.

However, if one can raise the effective index in the gaps between the waveguides at the star coupler boundary, then one can have a better mode match to free space, reducing the loss. This lowering must be done adiabatically. One way is to gradually reduce the amount of etching between the waveguides as the free-space boundary is approached. This has been demonstrated using a double-etch technique[21] and with vertical tapers[22] (see Fig. 6). Both were successful, the vertical tapers resulting in 32-channel packaged waveguide grating routers with only 1.1 dB fiber-to-fiber insertion loss.

However, both of these techniques require additional critical processing steps. It has been shown that the insertion loss can be reduced by placing waveguides perpendicular to the waveguides in the array,[23] as shown in Fig. 6. This is called segmentation. The segmentation raises the effective index in the gaps between the waveguides, making the local normal mode more like a plane wave, thus increasing the transmissivity. The main advantage of the segmentation is that no extra critical processing steps or special masks are required, unlike with vertical tapering. To reduce the loss even further, one can reduce the effective gap, which is larger than the final gap because of the converging nature of the waveguides, by widening the waveguides to achieve the minimum gap as allowed by lithography and maintaining the gap until reaching the free space region.

2.6 STAR COUPLER SIZE

Where are the boundaries of the star coupler? The answer is not right after the free-space region boundary, since the waveguides are coupled there. We will do the following calculation approximating the coupled waveguides as vertical parallel slabs of higher refractive index material, infinite in extent (i.e., the effective index method). We will also assume that there is an infinite number of waveguides; in the actual design, this is simulated by using dummy waveguides. The amount of mutual coupling among the waveguides depends on the difference in propagation constants between the eigenmodes of the structure. We will compute the propagation constant difference between the slowest (largest propagation constant) and fastest guided local normal modes (Bloch modes in this case).[24] Moving away from the star coupler center, when the propagation constants are nearly identical, the net mode propagates undistorted, and the mutual coupling is over.

By matching waveguide boundary conditions for the TE mode [quasi-transverse magnetic (TM) mode in the structure], we find the dispersion relations for the slow and fast Bloch modes:

$$\kappa_{\text{slow}} \tan(\kappa_{\text{slow}} w/2) = \gamma \frac{1 - \exp[\gamma_{\text{slow}}(w-a)]}{1 + \exp[\gamma_{\text{slow}}(w-a)]} \tag{18}$$

$$\kappa_{\text{fast}} \tan(\kappa_{\text{fast}} w/2) = \gamma \frac{1 + \exp[\gamma_{\text{fast}}(w-a)]}{1 - \exp[\gamma_{\text{fast}}(w-a)]} \tag{19}$$

respectively, where $\kappa = \sqrt{n_{\text{core}}^2 k^2 - \beta^2}$ and, $\gamma = \sqrt{\beta^2 - n_{\text{clad}}^2 k^2}$, w is the waveguide width, and a is the center-to-center waveguide spacing.

Let $\beta_{\text{slow}} = \beta_{\text{fast}} + \Delta\beta$. Constraining ourselves to the case of waveguide separation large enough so that $\exp[\gamma(w-a)] \ll 1$, we find

$$\Delta\beta \approx \frac{4\kappa^2\gamma^2 \exp(\gamma w)}{n_{\text{core}} k[\kappa^2 + \gamma^2 + w(\kappa^2\gamma + \gamma^3)/2]} \exp(-\gamma a) \equiv C \exp(-\gamma a) \tag{20}$$

which is the speed difference between the fastest and slowest Bloch modes.

Suppose we start with light in every other waveguide of the array. This is a sum of the slow and fast modes. When the phase accumulation between the propagation constants

$$\phi = \int_{z_0}^{\infty} \Delta\beta dz \tag{21}$$

reaches π, the light will have shifted to the other set of every other waveguide. We want to know the distance z_0 at which the coupling thereafter is <-25 dB in power of the complete transfer. The transfer is sinusoidal so it happens when $\phi < 0.11$. Substituting Eq. 20 into Eq. 21, we find

$$z_0 < \ln\left(\frac{C}{0.11\gamma\alpha}\right)\frac{1}{\gamma\alpha} \tag{22}$$

where α is the angle between neighboring waveguides.

From Eq. 22, we learn that the mutual coupling between waveguides in an array diverging at an angle is finite. Thus, once the waveguides are extended to a separation of $a > \alpha z_0$ the amount of further mutual coupling is negligible and the waveguides can be bent, phase-shifted, or continued with negligible effect on the star coupler. We thus call the entire region of length $L = 2z_0 - R$ the star coupler, where

$$R = \frac{a_{s_1} a_{s_2} M}{\lambda} \tag{23}$$

is the free-space radius, a_{si} is the center-to-center inlet spacing at the edges of the free-space region on side i and M is the number of inlets on each side. This is derived by having each side fill the zone, λ/a. Taking the symmetric case of $a_{s1} = a_{s2} \equiv a_s$,

$$L = \frac{a_s M}{\lambda}\left[\frac{2}{\gamma}\ln\left(\frac{c a_s M}{0.11\gamma\lambda}\right) - a_s\right] \tag{24}$$

Shown in Fig. 7 are plots of αz_0 and L vs. a_s for a typical case in silica with an effective index step of 0.004 and of 0.006, $n = 1.45$, $w = 6\,\mu\text{m}$, and $M = 50$ and 100. As one can see, decreasing a_s allows one to decrease the star coupler physical size, and thus that of the entire device. It is important to realize that z_0 is relatively insensitive to the width of the waveguides w. Instead, it is the

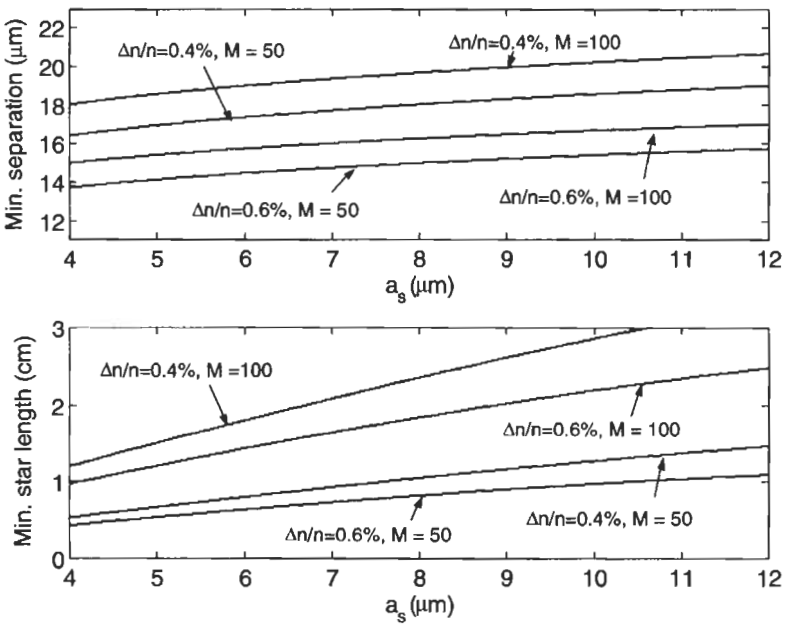


Fig. 7 Total star coupler length vs. center-to-center spacing of waveguides at the free-space region boundary.

waveguide center-to-center spacing a and the index step to which the mutual coupling is substantially more sensitive. However, the danger of having a_s too small is a very strong mutual coupling, making the filter have stronger aberrations and be more sensitive to fabrication.

2.7 WAVEGUIDE GRATING ROUTER

2.7.1 Multiplexer/Demultiplexer

By connecting two star couplers by an array of waveguides of linearly progressively increasing path length, one can make a waveguide grating router (WGR), which can serve as a spectral multiplexer or demultiplexer (see Fig. 8). The first transmissive waveguide device demonstrating optical angular dispersion was published by Smit in 1988 and consisted of a parallel array of curving waveguides in Al_2O_3 ridge waveguides.[25] Also in 1988, as discussed previously, Dragone published the star coupler.[18] Then, Smit and Vellekoop proposed and demonstrated converging the waveguide array to points on both sides, terminating the waveguides on an arc before reaching the points with output waveguides on the other side, similar to Dragone's star coupler.[26] Two months later, Takahashi submitted an experiment using an array of waveguides similar

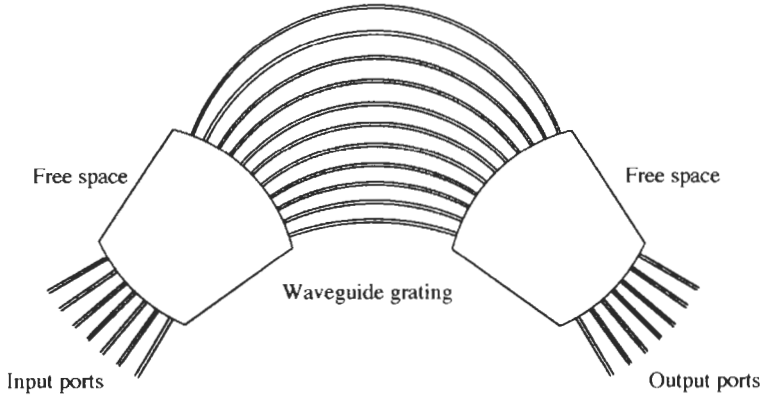


Fig. 8 The waveguide grating router.

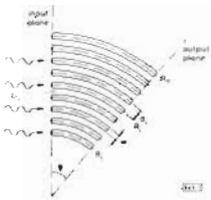
to Smit's parallel array of waveguides and used external bulk lenses to couple to fibers, also similar to a star coupler but in free-space optics.[27] One day after Takahashi's submission date, Dragone submitted the proposal for the WGR as it is known today, i.e., two star couplers connected by a waveguide array of linearly progressively increasing path length, including the $N_1 \times N_2$ case (see next section). Dragone experimentally demonstrated the WGR in 1991.[20,28] As a result of this complicated history, each group has given its own name to the device: Dragone's group calls it the WGR, Smit's group calls it the PHASAR, and Takahashi's group calls it the arrayed waveguide grating (AWG). A figure detailing this history is shown in Fig. 9.

One can use a modification to the usual grating equation to describe the operation of a conventional WGR:

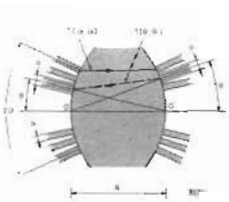
$$\beta_{\text{waveguide}} \Delta L + \beta_{\text{slab}} \left(a_1 \sin\theta_1 + a_2 \sin\theta_2\right) = 2\pi A \tag{25}$$

where a_1 and a_2 are center-to-center spacings of the grating arms at the free-space boundary in star couplers 1 and 2, θ_1 and θ_2 are the angular locations of the ports in star couplers 1 and 2 as measured from the star coupler center line, and A is the grating order (which must be an integer). The main modification is the addition of ΔL, which is the path-length difference between adjacent connecting waveguides (grating arms). For a bulk diffraction grating, A is usually 1 or -1, but for a WGR, A is usually $>\sim 20$ when operating at the wavelength of interest.

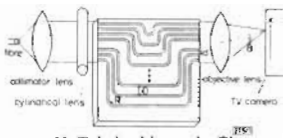
From Eq. 25 one can see that if the ports are equally spaced in angle and the chromatic dispersion of β is small, then in the small-angle approximation the passband peaks are equally spaced in wavelength from port to port. Within a given port, there is one passband for each grating order, and from Eq. 25 one can see that these are evenly spaced in inverse wavelength, i.e., in frequency. Note that the spacing is also inversely proportional to the group refractive



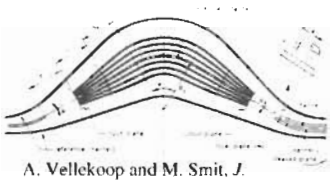
M. Smit, *Electron. Lett.*, p. 385, 1988.



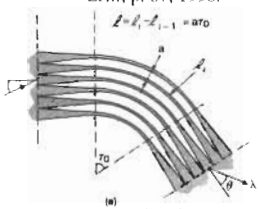
C. Dragone, *Electron. Lett.*, p. 942, 1988.



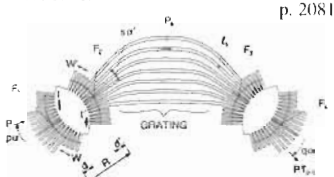
H. Takahashi, et. al., *Electron. Lett.*, p. 87, 1990.



A. Vellekoop and M. Smit, *J. Lightwave Technol.*, p. 310, 1991.



C. Dragone, *J. Opt. Soc. Am. A*, p. 2081, 1990.



C. Dragone, *IEEE Photon. Technol. Lett.*, p. 812, 1991.

Fig. 9 History of the waveguide grating router.

index ($n - \lambda dn/d\lambda$), rather than just the index, which can be significantly different, even in silica (typical silica waveguides have an index of 1.45 and group index of 1.48 at 1.55 μm wavelength).

A conventional WGR has negligible chromatic dispersion.[29] This is because different wavelengths do not take different recombining path lengths in the WGR.

2.7.2 $N_1 \times N_2$ Waveguide Grating Router

The $N_1 \times N_2$ WGR is a WGR with multiple inputs and multiple outputs.[20] Each port on one side is demultiplexed on the other side, each demultiplexed spectrum shifted in wavelength with respect to the others. It allows one to make a strict-sense nonblocking cross-connect using tunable lasers (see Chapter 8). In such a case, each laser is connected to a WGR input and can reach one of the WGR outputs by choosing an appropriate wavelength.

One might think one could make an $N \times N$ WGR with a free spectral range of N channels, and then the total number of wavelengths required in the cross-connect would be only N. However, this does not always work.[30] Since the outputs are evenly spaced in angle with even spacing in wavelength, one would need to choose to have the channels evenly spaced in wavelength (whereas the

channels are evenly spaced in frequency in WDM systems). Secondly, such an $N \times N$ WGR employs 3 grating orders over all the possibilities of operation, and thus since the channel spacing changes slightly with grating order (see Eq. 25), it is impossible to make a perfect $N \times N$ WGR that needs only N wavelengths to make a cross-connect using a plain WGR. However, if one keeps

$$N \ll A \tag{26}$$

then the deviation from ideality is small. Thus one can make a 40×40 WGR with a channel spacing of 0.4 nm at a center wavelength of 1550 nm that needs only 40 wavelengths to make a cross connect with negligible impairments.[30]

One solution to make a $N \times N$ WGR that needs only N wavelengths with N not $\ll A$ is to make a WGR with greater than $2N$ ports. Then place N ports on one side and $2N$ ports on the other side. Then on the $2N$ side, combine the 1st port with the $N + 1$st port in a 50/50 coupler, the 2nd port with the $N + 2$nd port, etc.[31] The penalty is a 3 dB excess loss, a large device size, and if the 50/50 couplers are integrated with the WGR, then in general waveguide crossings are required.

2.8 MAKING FABRICATION-ROBUST WAVEGUIDE ARRAYS

To minimize aberrations in waveguide arrays, one would like the array length distribution to change at most by a proportionality constant and an additive constant in the presence of fabrication uncertainties. This way the shape of the distribution is unchanged with polarization and fabrication errors. For a grating (waveguides of linearly increasing path length) there will be only a wavelength shift, and for a lens (waveguides all of the same path length), nothing will happen. This can be done by using the same radius for all bends and making sure that the sum of the lengths of the straights of each waveguide obey the desired distribution independently.

To accomplish this robustness, a conventional waveguide grating arm should be composed of three straights and two bends,[20] as shown in Fig. 10. The only constraint is that α_1 and α_3 have linear distributions.

A conventional waveguide lens (an array of waveguides nominally all of the same length) arm should be composed of five straights and four bends (it is not possible to make it with only three straights and two bends if both sides are connected to star couplers).[32] The only constraint is

$$\alpha_2 - \alpha_4 = \frac{\alpha_1 - \alpha_5}{2} + \text{constant} \tag{27}$$

where the constant is the same for all the waveguides. This makes the sum of the bend lengths for each waveguide equal. One (more constraining) way to achieve this is to make α_2 the average of α_1 and α_3 plus a constant and α_4 the average of α_3 and α_5 plus a constant.

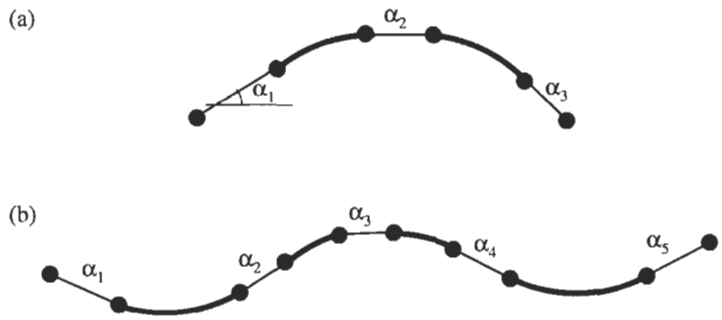


Fig. 10 Waveguide patterns that are robust to fabrication changes. The thin and thick lines are straights and bends, respectively.

2.9 SPECTRAL SAMPLING

We will substitute Eq. 16 into Eq. 14 and make some approximations in order to better understand the star-coupler transmissivity. Remember, though, that if one wants to calculate the star coupler very accurately, one cannot make these approximations, especially in the case of a low-loss star coupler with significant mutual coupling among waveguides. Assuming $k_x \ll k, \theta_1 \ll 1$, and $\theta_2 \ll 1$, the transmissivity from a port at θ_A to a port at θ_B becomes approximately

$$\exp[jkR(1-\theta_1\theta_2]\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u_2(x)\exp(jkx\theta_1)\,u_1(x')\exp\left(jkx'\theta_2+jk_xx'\right)\right.$$
$$\left.\times\exp\left(-j\frac{k_x^2}{2k}R-jk_xx\right)dx'dxdk_x\right)$$
$$\Big/2\pi\sqrt{\int_{-\infty}^{\infty}|u_1(x)|^2\,dx\int_{-\infty}^{\infty}|u_2(x)|^2\,dx} \tag{28}$$

If the Fourier transforms of u_A and u_B vary slowly with k_x then the transmissivity is approximately

$$\exp(-jkR\theta_1\theta_2)\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}u_2(x)\exp(jkx\theta_1)\,u_1(x')\exp\left(jkx'\theta_2\right)dx'dx}{\sqrt{(\pi R/2k)\int_{-\infty}^{\infty}|u_1(x)|^2\,dx\int_{-\infty}^{\infty}|u_2(x)|^2\,dx}} \tag{29}$$

dropping constant phase shift.

Finally for a $M \times M$ star coupler, if each waveguide m_1 is at an angle $[m_1-(M+1)/2]/(a_2nM)$, and likewise with side 2, then one can write the star coupler transmissivity as

$$t_{\text{star}}(m_1,m_2)=\exp\left(-\frac{j2\pi m_1m_2}{M}\right)f_1(m_1)f_2(m_2) \tag{30}$$

where f_1 and f_2 are some functions dependent only on the star coupler geometry. The important point is that f_1 is a function of only m_1 and f_2 of only m_2. If the input to each port on one side of the star coupler has complex amplitude $s_1(m_1)$, then the output from each port $s_2(m_2)$ on the other side is given by

$$\frac{s_2(m_2)}{f_2(m_2)} = \sum_{m_1=1}^{M_1} \exp\left(-\frac{j2\pi m_1 m_2}{M}\right) f_1(m_1) s_1(m_1) \tag{31}$$

Thus $s_2(m_2)/f_2(m_2)$ is the discrete Fourier transform of $s_1(m_1)f(m_1)$. Employing the Nyquist sampling theorem, if $s_1(m_1)f(m_1)$ is zero outside the interval $m_1 = 1 \ldots M$, which is the minimum sampling bandwidth of $s_2(m_2)/f_2(m_2)$, then all of the information in $s_1(m_1)f(m_1)$ can be in $s_2(m_2)/f_2(m_2)$. $s_2(m_2)/f_2(m_2)$ is periodic with period M, so we need only M consecutive samples of $s_2(m_2)/f_2(m_2)$. Thus if there are less than or equal to M values of s_1, then all of their information is contained in M samples of s_2. So we can exactly reconstruct s_1 in a second star coupler that is cascaded with the first. The contraint on s_1 is equivalent to requiring the ports to occupy less than or equal to the arc subtended by the angle λ/a_2, which is the Brillouin zone as determined by the s_2 ports.

Suppose we have two WGRs connected by equal-length waveguides. To transport a signal through the entire structure without distortion, the two central star couplers must have their grating arms occupy equal to or less than the Brillouin zone determined by the inlet spacing of the connecting waveguides.[33] In other words, the transmissivity will be perfectly flat if the number of grating arms in the WGRs is equal to or less than the number of connecting waveguides that can fit in the star-coupler central Brillouin zones (this number is equivalent in many cases to the number of channels in the WGR free-spectral range (FSR)).

In a conventional WGR, the number of grating arms is ~2.5 times the number of channels in the WGR FSR, and the spectrum is undersampled. Thus a back-to-back configuration of such WGRs, useful, for instance, as a programmable filter (see Sec. 2.14), would exhibit a series of peaks and dips. If we cut out the number of grating arms to be equal to the number of channels in the WGR FSR (i.e., remove the outermost grating arms without changing anything else), then we have perfect sampling and the spectrum is perfectly flat. We can even go past perfect sampling and oversample. In such a case, the spectrum is also perfectly flat but the entire structure is larger.

Generally, as the sampling increases, the amount of mutual coupling also increases. Thus as discussed in Section 2.4 one must use a BPM program to calculate the aberrations and then adjust the grating arm lengths (and lens arm lengths, if appropriate), and possibly their aiming point, to compensate for them. The losses can be significant in the case of perfect sampling because the limited grating arms truncates the field. However, this loss goes to zero as the waveguide array is more adiabatically transitioned into free space. As the loss is reduced to zero, and if the WGRs are symmetric (i.e., have dummy

waveguides on their other sides) then in the case of perfect sampling, the individual passbands, looking at the transmissivity through just one WGR, cross exactly at their 3 dB points. In the case of some loss or in the case of an undersampled spectrum on the input/output sides of the WGRs, then the passbands cross at a point higher than the 3 dB point, with additional contributions from next-to-nearest neighboring passbands. One can employ segmentation in the output inlets in the case of high sampling to reduce the losses.

Simulations of the transmissivity through two WGRs (N channels in the free-spectral range, M grating arms) connected by equal-length waveguides for the cases of under ($M > N$) and perfect ($M = N$) sampling are shown in Fig. 11.

One may note that the Brillouin zone width is wavelength-dependent, being equal to $\lambda_0/(na)$, where n is the index of refraction and a is the inlet spacing. In the case of two back-to-back WGRs, if the lens inlets are evenly spaced in wavelength or frequency, perfect sampling can occur at only one wavelength. To avoid undersampling, the rest of the controls must be designed to oversample, resulting in excess loss and/or a larger structure. However, there is an inlet spacing distribution that maintains perfect sampling for all the inlets:

To keep the Brillouin zone width (λ/a) constant, the following must hold:[34]

$$\frac{d\alpha(p)}{dp} \propto \frac{\lambda(p)}{n[\lambda(p)]} \tag{32}$$

where $\alpha(p)$ and $\lambda(p)$ are the angular position and wavelength of the lens inlet p, respectively. Equation 32 has the added benefit of equalizing the phase distortion caused by mutual coupling for all of the lens inlets. Solving Eq. 32, we find

$$\lambda(p) = \lambda_0 \exp\left[\frac{n}{n_g MAs}\left(p - \frac{P+1}{2}\right)\right] \tag{33}$$

where λ_0 is the center wavelength, n and n_g are the index and group index in the free-space region, M is the number of grating arms, A is the grating order, s is the sampling coefficient ($s = 1$ for perfect sampling, >1 for oversampling), and P is the number of lens inlets. Thus the optimum channel spacing

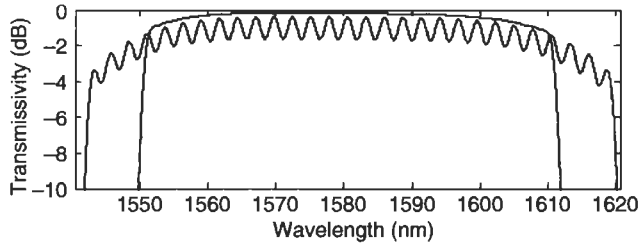


Fig. 11 Calculated transmissivity through back-to-back WGRs for the cases of under ($s = 0.8$; wiggly curve) and perfect ($s = 1.0$; smooth curve) sampling.

for two back-to-back WGRs with perfect spectral sampling is exponential in wavelength. Interestingly, for a given n and sampling coefficient, the possible wavelength sets are discrete, since M and A are integers.

To summarize, one can transmit through a star coupler without information loss (not the same as power loss) by ensuring that all the input ports are within the central Brillouin zone of the output ports. Note that one can use ports outside the Brillouin zone for loss reduction and/or frequency deviation reduction purposes (see Sec. 2.15.1.3), but these ports cannot contain new information, if one does not want to lose any information. Also, if the star-coupler transitions are perfectly adiabatic, light in any waveguides outside the central Brillouin zone will not be coupled through the star coupler anyway. Instead, that light, even though it propagates in the same direction as the waveguides, will not be found in the fundamental mode of any of the waveguides in the array once they are decoupled.

2.10 FOUR-PORT COUPLERS

We have discussed the star coupler which is a $2N$-port device. It essentially takes the discrete Fourier transform of the set of lightwave amplitudes in the incoming waveguides and puts the result in the amplitudes of the outgoing waveguides. However, when N is small, e.g., 2, the excess loss of the star coupler is significant (>1 dB). The advantage of the star coupler is that the splitting magnitude and phase are both very accurate despite wavelength and fabrication changes.

But many applications need $N = 2$ and require low loss and only a precise splitting magnitude or phase (or can posttrim the coupler). The lowest-loss coupler (<0.1 dB) is the directional coupler, consisting of two waveguides brought into close proximity. It has a precise splitting phase when entering from either side (always 90°), but its magnitude is quite sensitive to wavelength and fabrication changes. One can minimize the sensitivity by choosing the waveguide widths in the coupling region equal to the waveguide width that gives the minimum mode size (in a waveguide, the fundamental mode gets smaller as the waveguide width gets smaller up to a point and then rapidly grows larger).

One can trade off between splitting magnitude and phase sensitivity by cascading two or more directional couplers with a delay in one path between the couplers.[35,36] For example, making a 180° coupler (i.e., cross-state coupler) followed by a 120° length delay in one path followed by a 90° coupler (i.e., a 50/50 coupler) makes a 50/50 coupler with an extremely insensitive splitting magnitude of 50/50 but an uncertain relative phase between the two outputs.

Probably the most elegant four-port coupler is the adiabatic coupler,[37] which has an extremely robust splitting magnitude and a very uncertain splitting phase when entering the symmetric side. It involves adiabatically transforming two waveguides of very different widths to two waveguides of the same width (adiabatic means that the energy in the nth eigenmode is conserved,

and it is true for all the eigenmodes). The eigenmodes of two waveguides with very different widths are essentially the fundamental mode of each waveguide alone. The eigenmodes of two waveguides of the same width are a mode with equal energy in each waveguide with the same phase and equal energy with opposite phase. It thus acts as a perfect 50/50 coupler. At first, this may seem surprising. If one starts with two waveguides of different widths a meter apart and changes both waveguides to have the same width, one does not generally expect to have a perfect 50/50 coupler. However, in order for the change in widths to be adiabatic for this structure, they must change extremely slowly, because the propagation constants for the two eigenmodes are nearly identical. Thus, for an adiabatic coupler to be reasonably short, the two waveguides must be brought close together, increasing the propagation constant difference between the eigenmodes. Even so, the adiabatic coupler is quite long, typically greater than 1 cm.

There is also a coupler called a multimode interference (MMI) coupler.[38] It uses the self-imaging property of a multimode waveguide. It works ideally for very high-index step waveguides. However, with silica waveguides the index step is generally low, so there are approximations which can result in significant excess loss. Also, MMI couplers have a splitting magnitude that is quite sensitive to their width. In silica they are best suited for applications like WGR passband flattening (see Sec. 2.16.2).

If the designer needs access to only 3 ports of the coupler with a 50/50 splitting ratio, the best choice is a y-branch coupler. The y-branch coupler has a slightly higher loss than the directional coupler (typically 0.2 dB) but has an extremely robust splitting magnitude and phase and is compact. All of these couplers are shown in Fig. 12.

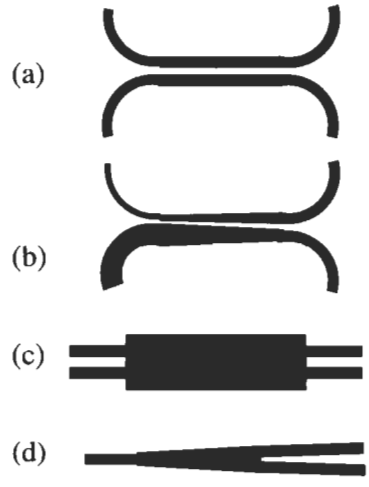


Fig. 12 Four-port couplers: (a) directional coupler, (b) adiabatic coupler, (c) multimode interference coupler, and (d) y-branch coupler (only three useful ports).

2.11 CASCADED MACH-ZEHNDER INTERFEROMETERS

One can construct an arbitrary periodic finite-impulse response (FIR) filter by connecting together 2 × 2 couplers and delay paths.[39] This is best viewed in the time domain. Suppose an impulse enters the cascade. The impulse is constantly split and sent with different delays. The final result is a train of impulses, the impulse response of the filter. If all of the delays are integer multiples of a constant, then the final impulses are spaced at equal intervals, and the filter is periodic. There is a procedure described in [40] for a 2 × 2 filter that allows one to determine the coupling percentages and delay lengths knowing the desired FIR. For a given FIR, there are four possible choices of sets of coupler angles (i.e., splitting ratios) and path lengths found from this deterministic procedure. These are simply taking the mirror image of the filter (i.e., using it backwards and forwards, which gives the same filter response magnitude), negating all of the coupler angles, and both. Making the filter out of directional couplers allows for very low insertion loss. This is because the filter never couples to more than two modes at once. Figure 13 shows a band splitter in silica that has less than 1 dB fiber-to-fiber insertion loss and less than −50 dB crosstalk.[39]

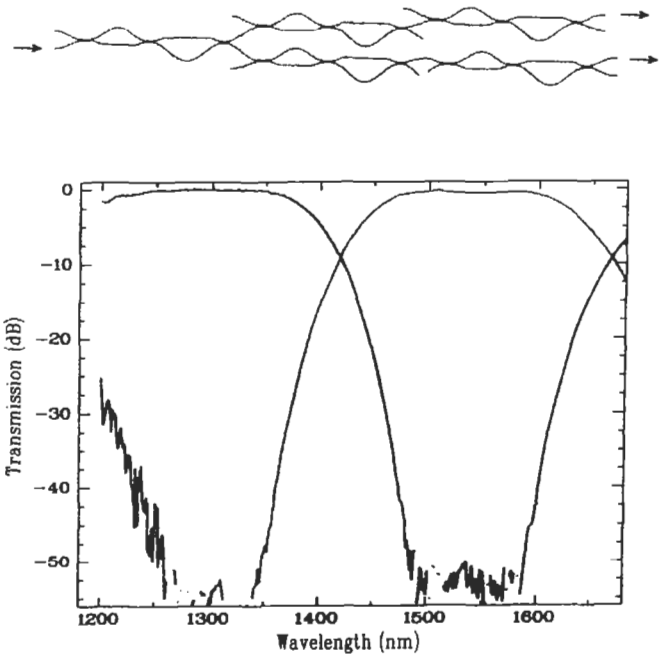


Fig. 13 1.3-μm/1.5-μm band filter constructed using cascaded Mach-Zehnder interferometers. From Ref. 39.

Unlike the conventional WGR, casacaded MZ filters do not in general have a linear phase vs. frequency response and thus exhibit chromatic dispersion. This can be understood by noting that in each stage different spectral portions predominantly travel different path lengths. However, since the frequency response between an input/output port of a 2 × 2 casacaded MZ filter is the complex conjugate of the response between the other input/output port combination, one can use a series of two identical 2 × 2 filters as shown in Fig. 14 to achieve a perfectly linear phase. Such a concept was used to make a dispersion-free interleaver in silica, seen in Fig. 15.[41] The cascading of the two filters also makes the contrast of the filter very high.

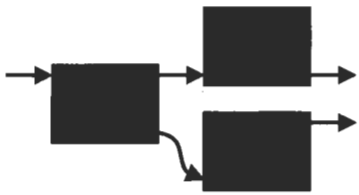


Fig. 14 Configuration for eliminating chromatic dispersion in four-port filters.

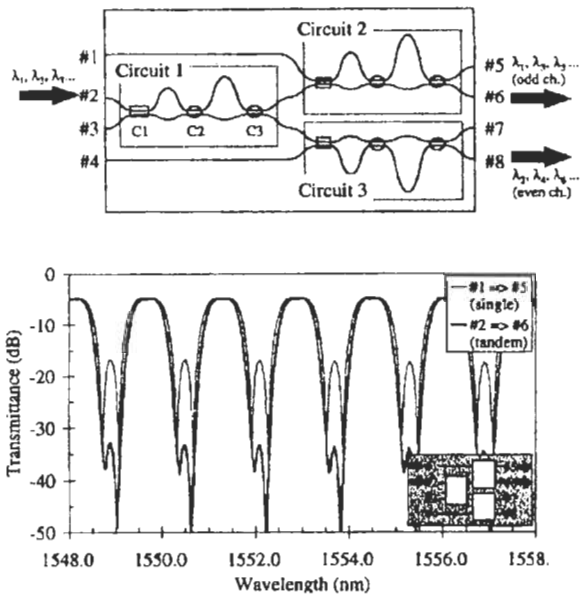


Fig. 15 Narrow-band interleaving filter using cascaded Mach-Zehnder interferometers. From Ref. 41.

Fig. 16 Thermooptic phase shifter in silica.

2.12 PHASE SHIFTER

Silica waveguides are amorphous and centrosymmetric, so no Pockels-based phase shifters with reasonable voltages have been attained in silica. Thus the usual way to make a phase shifter in silica is to make a thermooptic phase shifter.[42] It consists of simply a strip heater over the waveguide (see Fig. 16), usually chrome, that changes the refractive index via raising the temperature of the core. The advantages of the thermooptic phase shifter is that it has neglible insertion loss and the fabrication is simple and robust; the disadvantages are that it consumes significant electrical power, typically 1 W/(2π), is relatively slow, typically 2 ms (which is fast enough for circuit switching), exhibits some thermal crosstalk, and exhibits some polarization dependence.

The refractive index of silica changes about 10^{-5}/°C. The length change with temperature plays no significant role. Because the glass can expand vertically but not horizontally, TM-polarized light shifts more than TE by about 5%. The silicon substrate serves as an excellent heat sink, so the temperature gradient is roughly linear from the heater to the substrate, and the heat spreads laterally by about the same distance. Since the typical total silica thickness is 40 μm, the nearest waveguide must be more than 40 μm away to avoid significant thermal crosstalk. This is why adiabatic ("digital") switches,[43] which need a controllable phase shift between two closely spaced waveguides, require too much power to be practical in silica. A typical thermooptic phase shifter is 4 mm long. Thus, to achieve π of phase shift, the core temperature must be raised by ~20°C. The temperature at the heater is about twice as high, 40°C above the substrate, and is thus about 65°C at π.

One can achieve a permanent refractive index change in silica waveguides by driving a thermooptic phase shifter with a very high drive power (5–8 W).[44] This hyperheating can result in a permanent increase in refractive index up to 10^{-6}. The mechanism is not known, but it is our belief that the glass is relaxed into a more compressed state by the local heat. This is because the waveguide cores are under extreme compressive stress by the upper cladding glass, caused by the cooling of the silicon wafer after upper cladding deposition. The local high heat from the phase shifter may allow the core to permanently compress more. This also explains why heating up the entire wafer to high temperatures (it is annealed at over 900°C after upper cladding deposition) does not have the same effect as the trimming since in that case the entire silicon wafer expands, reducing the stress. It is a nearly polarization-independent effect. It

is estimated that the waveguide core reaches temperatures over 350°C during the hyperheating. One needs to cover the heaters with an oxide or similar cap to prevent them from burning in the air during hyperheating.

2.13 MACH-ZEHNDER INTERFEROMETER SWITCH

A Mach-Zehnder interferometer consists of two three- or four-port couplers joined by two waveguides. The couplers are typically 50/50 couplers. When the two waveguides are equal in length to within a few wavelengths, and a phase shifter is on one or both arms (see Fig. 17), the interferometer acts as a switch. For example, if the couplers are 50/50 and the connecting waveguide lengths are the same, then light entering the top left port exits the bottom right port and vice versa. Thus, the switch is in a cross state. If one of the arms is then phase shifted by 180°, the switch will switch to a bar state ("bar" and "cross" comes from the physical appearance of the connections in a 2 × 2 switch.

If the two couplers are identical, the cross state extinction is much higher than the bar state extinction ratio in the face of fabrication variations. Also, because of the typical polarization dependence of thermooptic phase shifters, the best extinction ratio for both polarizations simultaneously is limited to about 23 dB. However, to avoid penalties due to in-band crosstalk, an extinction ratio greater than 35 dB is usually required. The best solution is to use dilation—two switches in series. To do a full 2 × 2 dilated switch, 4 Mach-Zehnder switches are required. A tutorial for Mach-Zehnder thermooptic switches is given in [42]. The power consumption and polarization dependence of Mach-Zehnder switches can be reduced by using one phase shifter on each arm of the interferometer in a push–pull fashion.[45] The interferometer must be biased in an intermediate state when no power is applied.

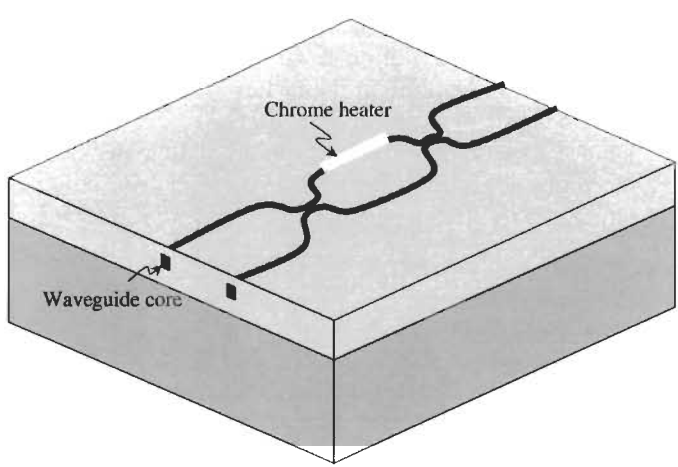


Fig. 17 Mach-Zehnder interferometer switch.

2.14 DYNAMIC GAIN EQUALIZATION FILTER

Dynamic gain equalization filters (DGEFs) are devices that can control channel powers in a WDM transmission line by having a chromatically variable transmissivity. DGEFs are especially needed in ultralong-haul cases in which the signals pass through so many optical amplifiers that small imperfections in the gain spectral flatness, plus Raman effects, result in large channel-power divergences without DGEFs.

One of the earliest DGEFs was a Mach-Zehnder interferometer in silica with one arm longer than the other so as to create a sinusoidal filter response.[46] The spectral position and depth of the sinusoid could be controlled thermooptically. Many applications require a much higher spectral resolution now.

One can create a DGEF by having a demultiplexer and multiplexer connected by an array of attenuators. However, the loss of a demultiplexer–multiplexer pair $1/T$ is often too high for many networks. One can reduce the insertion loss, giving up dynamic range, by placing the demux–mux pair inside one arm of a Mach-Zehnder interferometer and replacing the attenuators with phase shifters,[47] as shown in Fig. 18. If the couplers have splitting ratio $R/(1-R)$, then the transmissivity through the device is

$$\left\{R + (1-R)\sqrt{T}\cos[\phi(p)]\right\}^2 \tag{34}$$

where $\phi(p)$ is the phase of the phase shifter for the pth control band. Thus the dynamic range is

$$\left[\frac{R + (1-R)\sqrt{T}}{R - (1-R)\sqrt{T}}\right]^2 \tag{35}$$

For example, if $R = 0.5$, and $T = 0.1$ (i.e., 10-dB loss for the demux-mux pair plus phase shifters), then the total device loss is only 3.6 dB at minimum attenuation, and the dynamic range is 5.7 dB. Interestingly, even when R is chosen to make the attenuation range infinite, $R = \sqrt{T}/(1+\sqrt{T})$, e.g., a

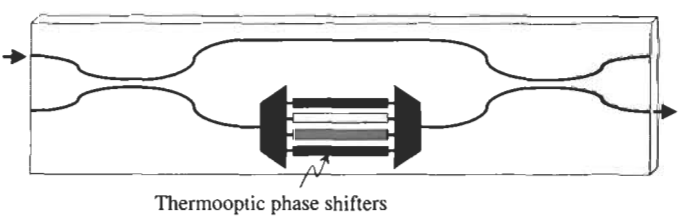


Fig. 18 Single-filtered arm Mach-Zehnder interferometer. This can be used as a dynamic gain equalization filter or a wavelength selective cross-connect.

wavelength-selective switch, the maximum transmissivity is still higher than T, approaching a value of 6 dB better for small T.

Note that if the couplers in the Mach-Zehnder interferometer are directional couplers and if $R = 0.5$, then it is best for fabrication robustness to use opposite input–output ports; if the upper port is the input on the left, then the output port should be the upper port on the right. This is the same reason as for the Mach-Zehnder switch (see Sec. 2.13).

The waveguide layout in silica for such a DGEF is shown in Fig. 19. It consists of a two waveguide gratings connected by an array of equal-length waveguides.[33] We would like the spectral response to be completely smooth, so the lens inlets perfectly sample the spectra from the waveguide gratings. Also, we would like the loss to be as small as possible, so all of the star couplers make their transitions as adiabatic as possible by using segmentation. The phase shifters are thermooptic. The nonfiltered path length is equal to the average length through each waveguide grating plus the length of the connecting waveguides, so that the resulting device has negligible chromatic dispersion (nearly same path length for all spectral portions).

A DGEF must have extremely low polarization-dependent loss (PDL) and polarization-mode dispersion (PMD). However, DGEFs of the single-filtered-arm design generally exhibit high PDL because of the significant physical difference between the interfering paths. Small polarization conversion in the waveguides, especially in the bends in which the lightwave interacts asymmetrically with the waveguide walls, foils attempts to use a waveplate inserted in the middle in such a case.

One way to solve these polarization problems is to employ polarization diversity (see Fig. 20).[48] In such a case, the input to the DGEF enters an optical circulator and then a polarization splitter. One polarization enters the circuit from one side, and the other from the side, both oriented so that they enter the same eigenpolarization of the circuit. The returning lightwaves then recombine in the polarization splitter, reenter the circulator, and then are sent to the DGEF output. In this way, the PDL of the device is determined only by the circulator, provided that the polarization beam splitter inside the polarization splitter has a high extinction ratio.

The circulator and polarization splitter can be combined in one bulk optic device to reduce loss. Shown in Fig. 21 are results from a DGEF with such a combined device. It achieves less than 4.5 dB fiber-to-fiber insertion loss over the C-band.[34] Shown in Fig. 22 are results from a DGEF with 100-GHz band

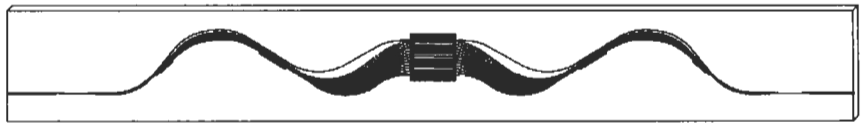


Fig. 19 Waveguide layout of a single-filtered-arm interferometer DGEF.

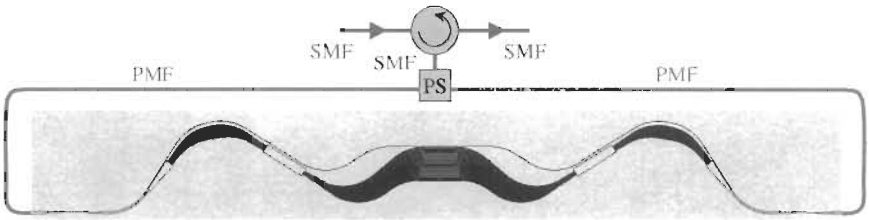


Fig. 20 Polarization diversity scheme for achieving polarization independence in a two-port integrated device, such as the DGEF. From Ref. 34.

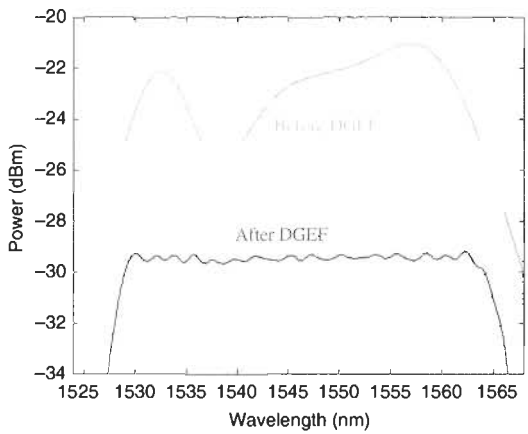


Fig. 21 Measured transmissivity through a single-filtered-arm interferometer DGEF.

spacing and 44 bands. Such a DGEF can somewhat control individual channel powers and yet have a smooth spectrum to minimize signal distortion.[48]

Because of fabrication imperfections, the quiescent state is not always flat. One can use hyperheating trimming with the DGEF to preflatten it, saving power consumption, maximum required drive power, and achieving a nice power-off state. The transmissivity through a DGEF before and after trimming is shown in Fig. 23.

2.15 WAVELENGTH-SELECTIVE CROSS-CONNECT

Recently, there has been significant progress in strict-sense nonblocking optical cross-connects using mechanically moving micro electromechanical systems (MEMS) mirrors. These are fabrics in which each of the N ports on one side can connect to any of the N ports on the other side. There are thus $N!$ possible switch settings. Fabrics with N up to 1000 have been demonstrated with MEMS.[1] However, for WDM networks, such fabrics still require optical demultiplexers and multiplexers. Also, such cross-connects are not naturally

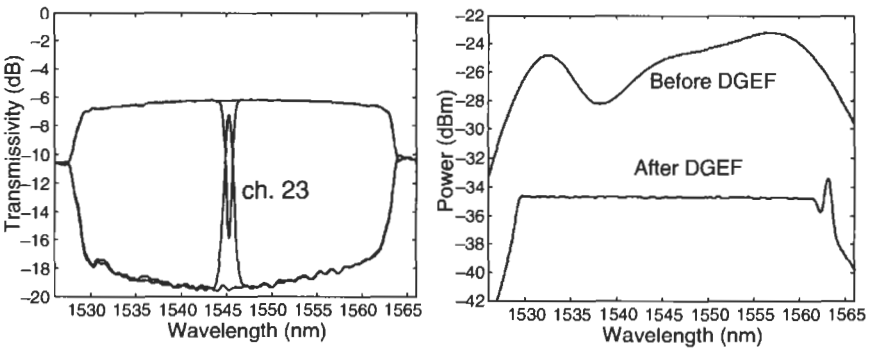


Fig. 22 Measured performance of a 100-GHz-band-spacing single-filtered-arm interferometer DGEF. From Ref. 48.

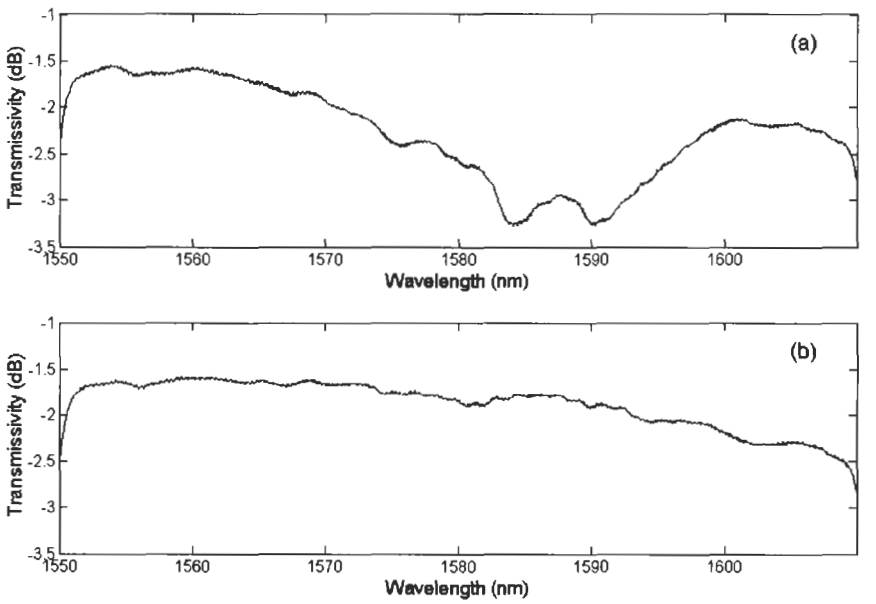


Fig. 23 A single-filtered-arm interferometer DGEF with no power applied before and after trimming.

modular, and thus the user must purchase initially the largest fabric they think they will ever need, resulting in a high startup cost.

Another approach is to use wavelength-selective cross-connects (WSCs).[49] A wavelength-selective cross-connect (WSC) is a wavelength-selective switch between fiberoptic lines. For instance, a 2 × 2 WSC can put each pair of wavelength channels on two fiberoptic lines in a bar or cross state. Certain

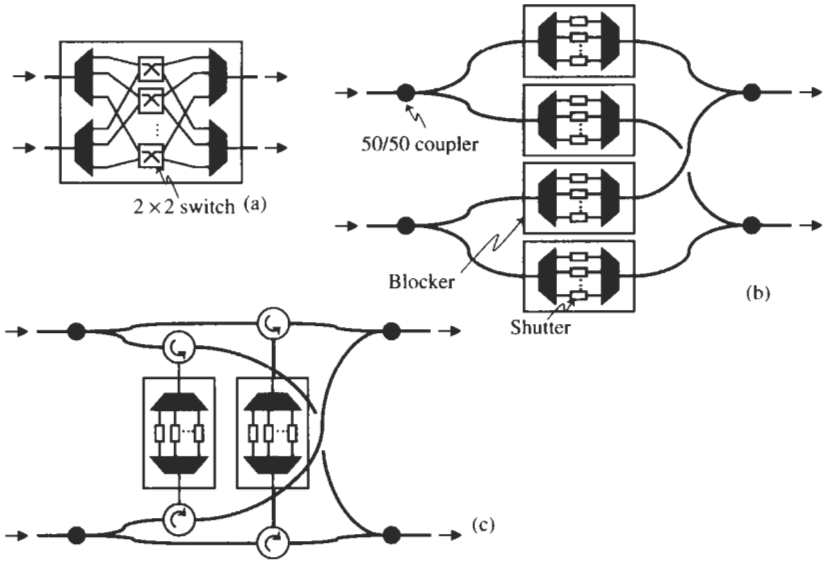


Fig. 24 Some possible wavelength-selective cross-connects.

architectures with WSCs are growable, allowing the user to start with only a small portion of the WSC and to add to it as network capacity grows.

There are two main approaches to WSCs. One is to integrate the entire switch, as shown in Fig. 24a for an $N \times N$ case. This is the most integrated way, and ultimately should be the lowest in cost. However, if such a device needs to be serviced, replaced, or upgraded, all N lines passing through the WSC must be interrupted. Carriers will likely not tolerate such a significant failure point. A more practical design is to use a broadcast-and-select architecture employing wavelength blockers,[49] as shown in Fig. 24b. A wavelength blocker is a 1×1 WSC and either passes or blocks each channel. A blocker is a much simpler device than a full WSC since it requires shutters instead of switches and needs no waveguide crossings. However, a broadcast-and-select architecture requires N^2 blockers, which is expensive. Fortunately, for the $N = 2$ case, one can use the design shown in Fig. 24c, reducing the number of blockers to two.[50]

2.15.1 Full Wavelength-Selective Cross-Connect

2.15.1.1 Employing 2 × 2 Switches

The first integrated WSC in silica is shown in Fig. 25.[51] It has 16 channels with Gaussian-shaped passbands, 100-GHz spaced. Note that the spectrum is undersampled, which is why there are dips between channels. The two demultiplexers and two multiplexers are all waveguide grating routers. The

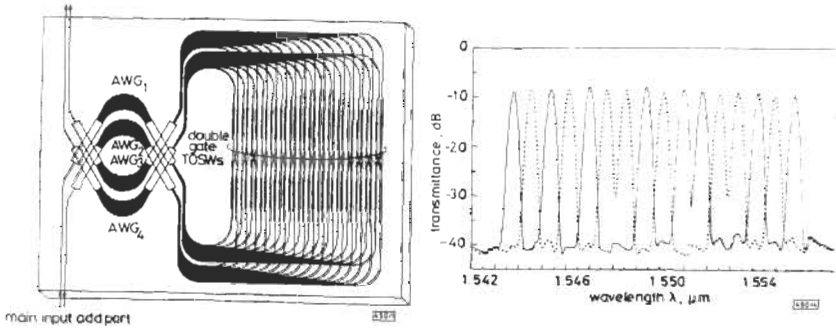


Fig. 25 16-channel WSC in silica that has waveguide crossings and the measured transmissivities. From Ref. 51.

switches are dilated 2 × 2 thermooptic Mach-Zehnder interferometer switches (see Sec. 2.13). As one can see, it has many waveguide crossings. Waveguide crossings have very low loss and crosstalk provided they cross at an angle greater than ~30°(<0.1 dB/crossing at 30° and becomes lower the wider the waveguides).[52] However, significant real estate is required to make the waveguides cross at such an angle. Thus, it is difficult to expand this design beyond 16 channels. By the way, one may wonder how well the waveguide grating routers match in terms of wavelength. On silica, the match can be very good (within 0.1 nm). For other material systems, such as InP, this is usually not the case, because of material nonuniformity.

There are many other possible arrangements of the WGRs and switches, and these are analyzed in [53]. However, nearly all require waveguide crossings.

2.15.1.2 Between-Channels Design

A 2 × 2 WSC design that does avoid waveguide crossings is shown in Fig. 26.[54] It consists of two WGRs and 2 × 2 switches. It eliminates waveguide crossings by having channels switch to adjacent ports. However, this design works if only every other channel is present, i.e., if there are large gaps between channels. Another drawback is that the switching crosstalk is limited to the adjacent-channel crosstalk of only one WGR. This is usually where the highest crosstalk is in a WGR.

2.15.1.3 Interleave Chirped Design

Another 2 × 2 WSC that avoids waveguide crossings yet has no gaps and theoretically has good crosstalk uses interleave chirped WGRs (see Fig. 27).[55]

The design works as follows: it consists of two interleave-chirped WGRs connected by equal-length waveguides (i.e., a waveguide lens). A conventional WGR creates one image per wavelength in each Brillouin zone (i.e., spatial

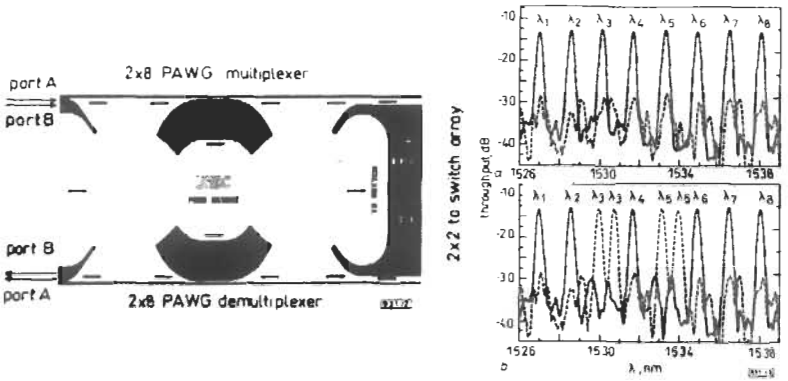


Fig. 26 8-channel WSC in silica that has no waveguides crossings but cannot use every other channel and the measured transmissivities. From Ref. 54.

diffraction order) of the star coupler on the opposite side. An interleave-chirped WGR produces P images per Brillouin zone. To minimize the loss and improve the switching performance, there are typically $P + 1$ connecting waveguides per channel. We will first cover the case of $P = 2$.

Multiplexed wavelength channels enter the device through one of the two lines on the left. Each channel is split by the left-hand WGR into multiple copies, two per star-coupler diffraction order. This is because every other arm in the WGR is shortened by $\lambda_c/4$, the interleave chirp. The connecting waveguides then collect the three copies with the highest optical powers from each channel and carry them to the right-hand WGR. This WGR then sends each channel to one of the two lines, the amount depending on the effective path lengths of the connecting waveguides, controlled by thermooptic phase shifters. Approximately, the bar state is when all connecting paths are equal, and the cross state is when the path length changes by $\pm\pi$ from waveguide to waveguide for the connecting waveguides of the respective channel.

As mentioned, the $P = 2$ interleave chirp involves increasing (or decreasing) the length of every other arm in the WGR by $\lambda_c/4$. The interleave chirps for $P = 2$ to 8 are shown in Table 1. For $P > 2$, the WSC is blocking and is best used as a $1 \times P$ or a $P \times 1$. However, one can use $2P$ $1 \times P$s to create a nonblocking $P \times P$ WSC.

The interleave cross-connect has several desirable attributes. First, the design is relatively compact and has no waveguide crossings. Second, the dominant loss in WGRs is due to radiation outside of Ω_0, where Ω represents a Brillouin zone, Ω_0 being the central one. The loss is especially high near the edges of Ω_0; this portion of Ω_0 is often even discarded. In the interleave cross-connect, the addition of connections in Ω_{-1} and Ω_1 significantly reduces and evens out the losses for all the channels. Third, as mentioned

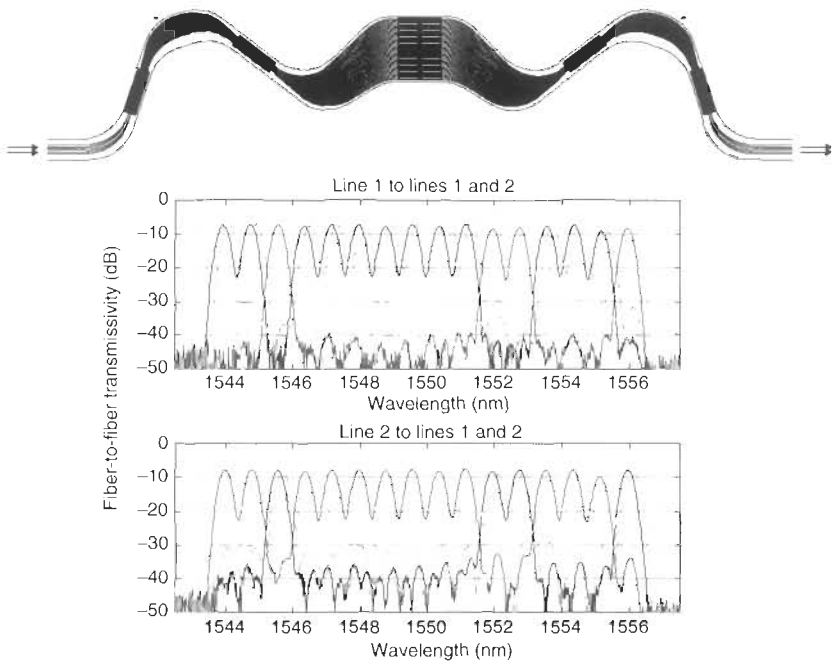


Fig. 27 16-channel WSC in silica that has no waveguide crossings and the measured transmissivities for a certain switched setting. Dark and light lines show TE- and TM-polarized responses, respectively. From Ref. 56.

Table 1 **Interleave-Chirp Sequences for Different Values of *P***

P	*Arm Length Change Series* [λ_c]
2	$\frac{1}{4}, 0$
3	$\frac{1}{3}, 0, 0$
4	$\frac{1}{2}, 0, 0, 0$
5	$\frac{1}{5}, \frac{3}{5}, \frac{1}{5}, 0, 0$
6	$\frac{3}{4}, \frac{1}{3}, \frac{3}{4}, 0, \frac{1}{12}, 0$
7	$\frac{1}{7}, \frac{3}{7}, \frac{6}{7}, \frac{3}{7}, \frac{1}{7}, 0, 0$
8	$\frac{1}{4}, \frac{1}{2}, 0, \frac{1}{2}, \frac{1}{4}, 0, 0, 0$

*From Ref. 56.

previously, a limitation of a single two-arm Mach-Zehnder switch is sensitivity to the power-splitting ratios of the two couplers. Often, dilation is required. However, by having three or more arms in the interferometer, as in the interleave WSC, one can always adjust the phases so as to have exactly zero power

in the one port provided that the sum of the powers in the arms with the lower powers is greater than the power in the third. In other words, one can always form a triangle from three segments provided that the sum of the lengths of the shorter two segments is greater than the third. A similar argument holds for the case of more than three arms. Thus, the phase shifters can make up for fabrication imperfections.

The interleave cross-connect has some drawbacks. First, the design as it is does not work when the channel number is greater than the grating order (this is the same as the limitation on the $N \times N$ WGR), i.e., when $N > A$. For instance, a 16-channel, 100-GHz-spaced WSC with $P = 2$ in the 1550 nm band works fine ($N = 16$ and $A = 54$). But a 40-channel version with otherwise the same parameters does not and needs a special modification. Second, it can be sensitive to drifts in the phase-shifter settings. Third, waveguides for each channel that are used to interfere with each other for the switching action are spread far apart, making it sensitive to phase-shifter crosstalk and mechanical strain.

Figure 27 shows the waveguide layout of a 16-wavelength-channel 2×2 interleave-chirped WSC with $P = 2$ and the measured fiber-fiber transmissivity fully packaged.[56] The phase shifters are thermooptic. A 6-channel version was previously demonstrated in InP.[57]

One can make a $1 \times K$ WSC by applying interleave chirps with longer periods. Minimally $P = K$. However, one can use a $P > K$ to make a smaller WSC that has more tolerance to phase-shifter drifts. For example, one could make a 2×2 WSC that is more robust to phase-shifter drift by having $P = 3$ and not using the third ports on either side and having 4 phase shifters per channel.

2.15.2 Wavelength Blocker

As shown earlier in Fig. 24b one can construct a WSC using wavelength blockers. Using blockers trades off insertion loss for device simplicity and flexibility. As a blocker, one can use an interleave-chirped WSC with $K = 1$. A partial demonstration of a 16-wavelength, 8×8 WSC using blockers is shown in Fig. 28.[58] Such a cross-connect is nonblocking in space but not in wavelength. However, such a cross-connect is growable by adding components and serviceable, since a device can be removed without interrupting any traffic.

Another design using wavelength blockers for 2×2 WSCs is shown in Fig. 24c.[50] The architecture works as follows: wavelength channels on two different fiberoptic lines enter the architecture from the left. Each is split into two by a 50/50 coupler and sent to two wavelength blockers surrounded by optical circulators. The left and right blockers control the bar and cross states, respectively. The outputs of the blockers are combined in two 50/50 couplers. For example, suppose we want the two λ_1 channels to stay in their rings but have the two λ_2 channels cross between the rings. Then the left blocker passes λ_1 and blocks λ_2 and vice versa for the right blocker.

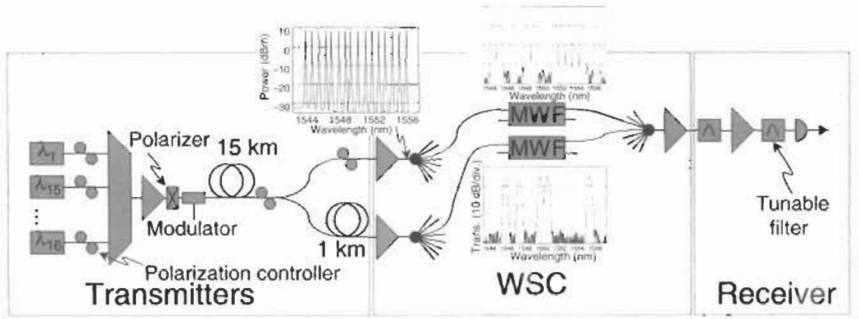


Fig. 28 Experimental setup for partial demonstration of a 16-wavelength 8 × 8 wavelength-selective cross-connect. MWF = multiwavelength filter, another name for a blocker. From Ref. 58.

A main advantage of this architecture, like all broadcast-and-select architectures, is that the optical devices can be serviced without interrupting traffic. For instance, suppose the right-hand blocker needs to be replaced. One can set up the network so that all of the channels are in the bar state. Then the right-hand blocker can be removed without interrupting any traffic. Another main advantage is that a device that requires only shutters, the blocker, is significantly easier to make than one requiring switches, the conventional WSC. For example, a blocker needs no waveguide crossings. Other advantages include the fact that the filter spectral response can be the same for both the bar and cross states, aiding the cascadability; and one can multicast for the purpose of, for example, 1 + 1 protection.

An additional device simplification is to make the blocker have a periodic filter response, such that channels are blocked in sets, trading off flexibility for device size. The waveguide layout of such a periodic blocker design in silica waveguides is shown in Fig. 29. It consists of two WGRs connected by an array of equal-length waveguides containing thermooptic Mach-Zehnder interferometer shutters. It can handle 128 channels, spaced by 50 GHz, in 16 sets of 8 channels each. The total number of channels is limited in a periodic device because of the following: each output from an arrayed waveguide has a slightly different FSR (see Eq. 25). This is related to the same reason why the $N \times N$ WGR generally needs more than N wavelengths to make an $N \times N$ cross-connect (see Sec. 2.7.2). Thus the channel spacing changes slightly from grating order to grating order. The worst-case frequency offset is given by

$$\text{worst offset} = \frac{\text{FSR}N\Delta f}{2\text{FSR} + 4f} \tag{36}$$

where N is the total number of channels, Δf is the channel spacing, and f is the center optical frequency of the span.

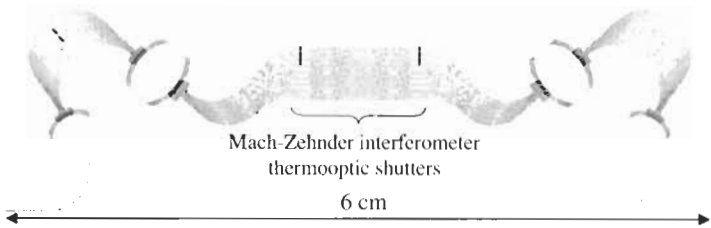


Fig. 29 Waveguide layout of 16-channel periodic wavelength blocker. From Ref. 50.

Consider a design for the case of $N = 128$, $\Delta f = 50\,\text{GHz}$, FSR $= 16 \times 50\,\text{GHz}$, and a center wavelength of 1550 nm, the frequency offset of the first and last channels from the centers of their passbands is 6.6 GHz. This frequency offset is made even smaller by connecting extra waveguides between the WGRs in the next-higher-order grating spatial diffraction orders. The secondary image of each wavelength has the opposite frequency offset, somewhat correcting the frequency error. Equally important, collecting the extra copies eliminates the usual 3 dB increase in loss for the channels at the edges of the FSR. This particular blocker has a slight oversampling of the spectra from the WGRs, and thus the transmissivity through the blocker is fully flat when adjacent shutters are open. It has two connecting waveguides per channel so as to have a flattop passband per 50-Ghz-spaced channel and collect extra images for seven of the FSR edge channels. Thus it has 31 grating arms and 46 connecting waveguides. Thus there are 46 shutters, but since the shutters are in pairs, each pair can share one thermooptic phase shifter, so there are only 23 active thermooptic phase shifters. Segmentation can be used on both sides of all the star couplers, except the input and output, to minimize the insertion loss.

An experiment was done with two packaged versions of these periodic blockers. The fibers and chips facets are angled at eight degrees because of the bidirectional usage of the blockers. Back reflections are below -55 dB in such a case. The fiber-to-fiber insertion loss is 5–7 dB, and the PDL is 0.5 to 1.0 dB. Each shutter pair requires $\sim$700 mW of electrical power to switch, making a worst case power consumption of $\sim$16W per device.

The shutters give the typical 20 to 25 dB extinction ratio expected in silica waveguide Mach-Zehnder switches. One could use dilation (two shutters in a row) to double this extinction ratio in the device. The entire setup of Fig. 24c was constructed using the two blockers, commercial circulators and 50/50 couplers. Figure 30 shows the measured transmissivities when all the channels are in the bar state, and Fig. 31 shows when five channel sets are in the cross state and one channel set is being multicasted. Note that, because there are no gaps between channels, this WSC can switch with a variable bandwidth, e.g., it can switch some 40 Gb/s channels spaced by 100 GHz and some 10 Gb/s channels spaced by 50 GHz. Figure 32 shows closeups of Fig. 31 from the span

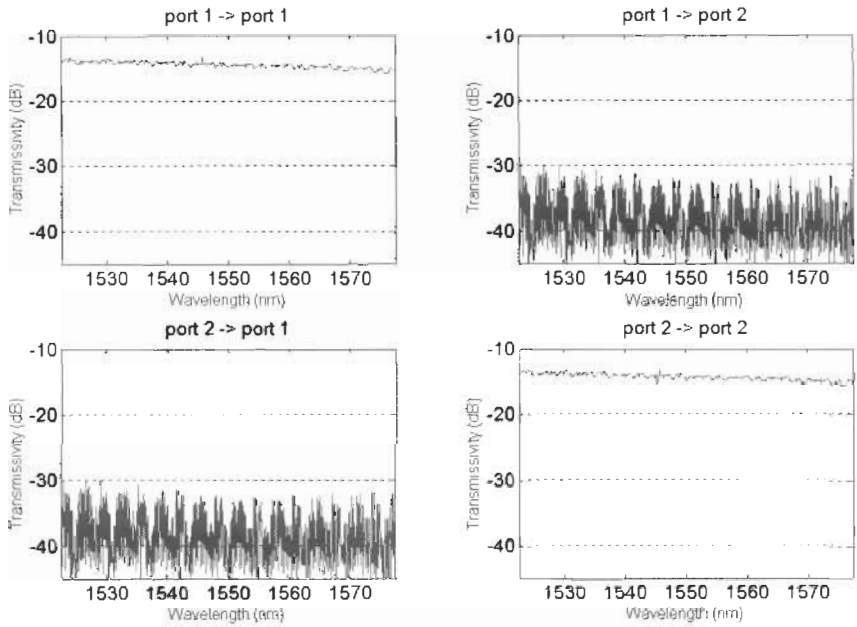


Fig. 30 Measured transmissivities of all four input–output combinations of the periodic broadcast-and-select wavelength-selective cross-connect for all channel sets in the bar state. From Ref. 50.

ends overlaid with a precise 50-GHz grid. As one can see, the passbands line up well to the grid across the entire span, although the stopband walls need to be steeper.

2.16 WAVELENGTH ADD-DROP

A wavelength add-drop (WAD) is a device that can remove from and add to wavelength channels on a transmission line. A WSC could be considered a subclass of a wavelength add-drop (WAD). One possible way to distinguish the more generic WAD from a WSC is that a WAD often has a different filter characteristic for the drop and add channels, while a WSC typically has the same filter characteristic between all its input and output ports. Also, some distinguish a WAD as a $2N+2$ port device, N being the number of wavelengths, and a WSC as a $2K$ (or sometimes $1+K$) port device, K being the number of fiberoptic lines.

2.16.1 Small Channel Count

Small-channel-count WADs are devices that can drop only a few wavelength channels. Most designs can pass only about the same number of channels

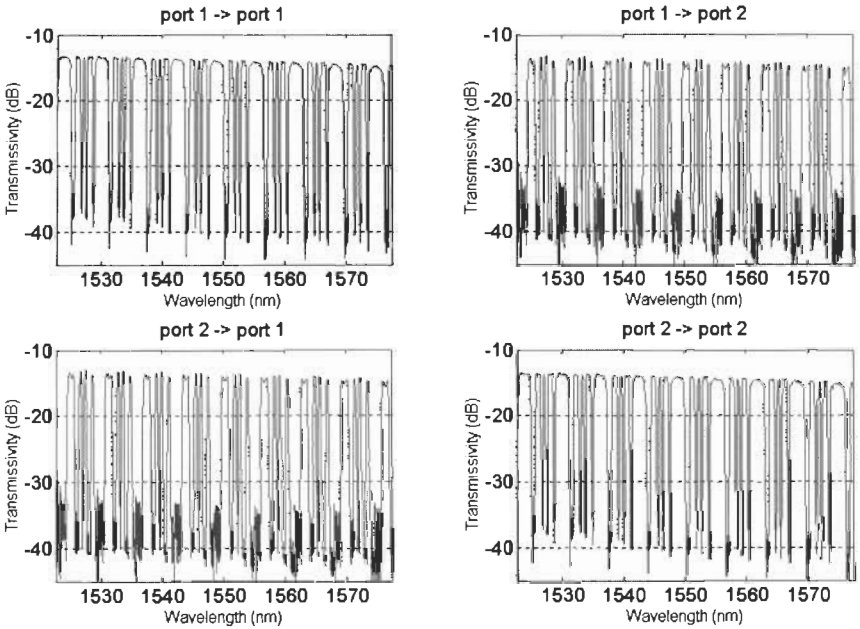


Fig. 31 For five channel sets in the cross state and one channel set being multicasted. From Ref. 50.

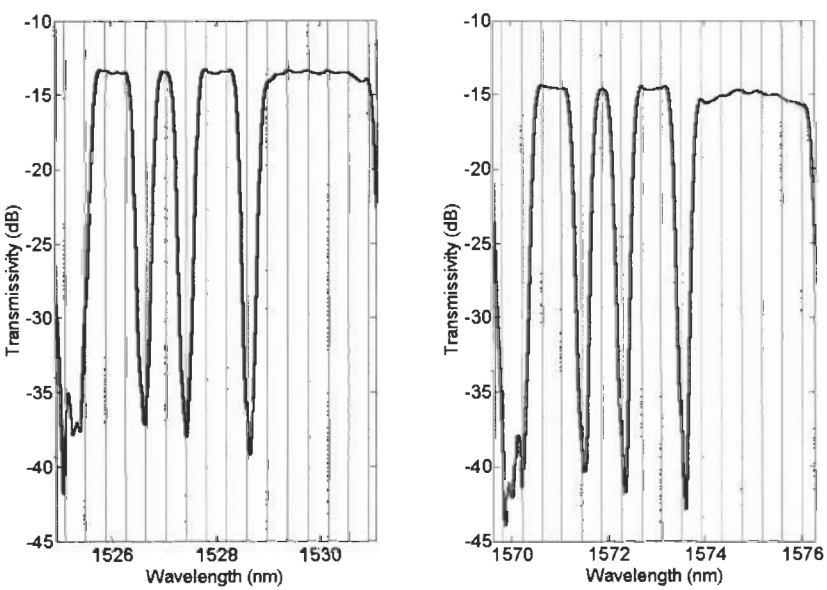


Fig. 32 Close-ups from previous figure. From Ref. 50.

as can be dropped, although there are some, such as those employing Bragg gratings (see Sec. 3.3 for an example), that can pass many more channels than they can drop.

One design in silica uses cascaded Mach-Zehnder interferometers, and an example is shown in Fig. 33. It consists of a tree structure of interleavers (see Sec. 2.11; the ones in Fig. 33 have nonlinear chromatic dispersion, though) that are tunable via thermooptic phase shifters, and is thus reconfigurable. The waveguides are actually in silicon-oxynitride (SiON), which is similar to silica except it can have a high index step (3.3%).[59]

Probably the very first integrated WAD is shown in Fig. 34. This 16-channel device consists of three WGRs and an output array connected by an array of Mach-Zehnder interferometer switches.[60]

2.16.2 Large Channel Count

As the number of WDM channels increases, the cascaded Mach-Zehnder approach soon requires too many stages to fit without needing 180-degree bends to fit on a wafer, and the approach of Fig. 34 quickly requires too many waveguide crossings and too much real estate to avoid crossing at angles less than 30° (to mitigate loss and crosstalk). One possible means to reach very large channel counts in an integrated device is to use waveguide arrays in a reflective arrangement with a striped mirror and a circulator.[61] The concept is shown in Fig. 35. Other advantages include automatic wavelength alignment of the demultiplexer and multiplexer, dilation with only one switch, and the ease with which all the possible paths a wavelength can take (via crosstalk)

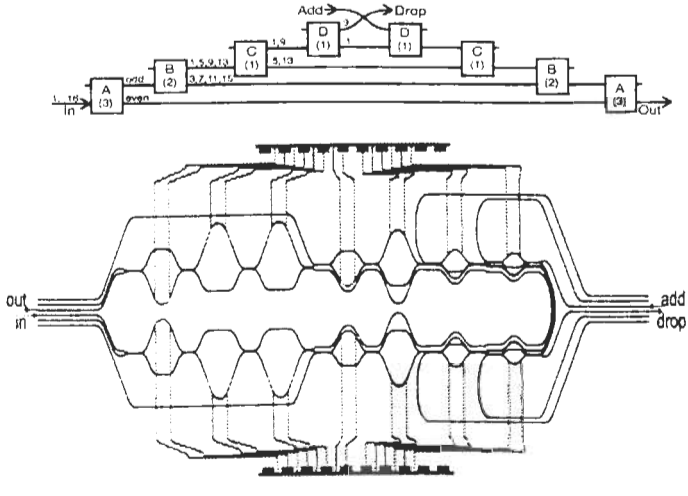


Fig. 33 Optical add-drop using cascaded Mach-Zehnder interferometers. From Ref. 59.

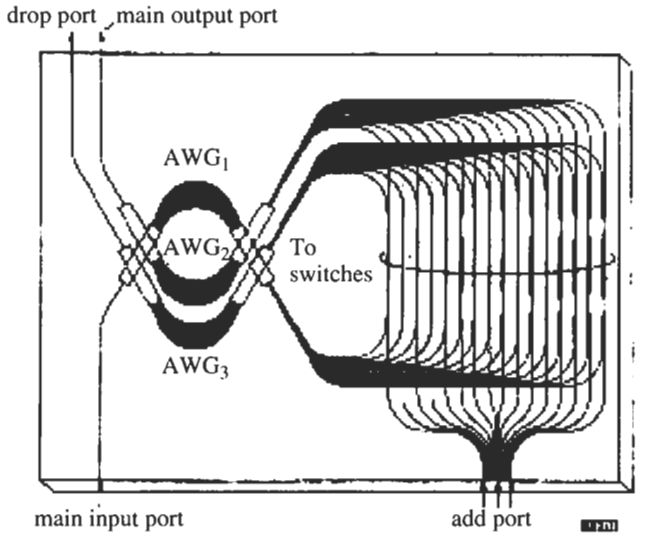


Fig. 34 Optical add-drop using waveguide grating routers and Mach-Zehnder switches with waveguide crossings. From Ref. 60.

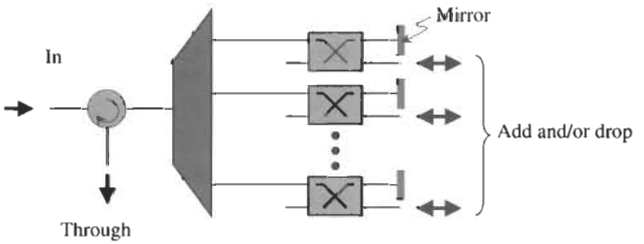


Fig. 35 Concept of reflective optical add-drop for large channel counts.

can be made to have the same length to within a few wavelengths, eliminating concerns of multipath interference.

The simplest type is to employ Gaussian-shaped passbands, which is the passband shape of the conventional WGR. For example, the add-drop of Fig. 34 has Gaussian passbands. However, Gaussian passbands exhibit strong passband narrowing when cascaded. One solution is to use an interleaver (see Sec. 2.11) and two WADs, one for the even-numbered channels and one for the odd-numbered, each with twice the channel spacing.[62] Such a net passband has a reasonable cascadability. An experiment doing this is shown in Fig. 36. Two 20-channel, 200-GHz-spacing Gaussian-passband reflective WADs are placed on one chip, one shifted by 100-GHz with respect to the other. A separate 100/200 GHz interleaver is used to send the channels to each WAD. It is thus a relatively compact 40-channel WAD.

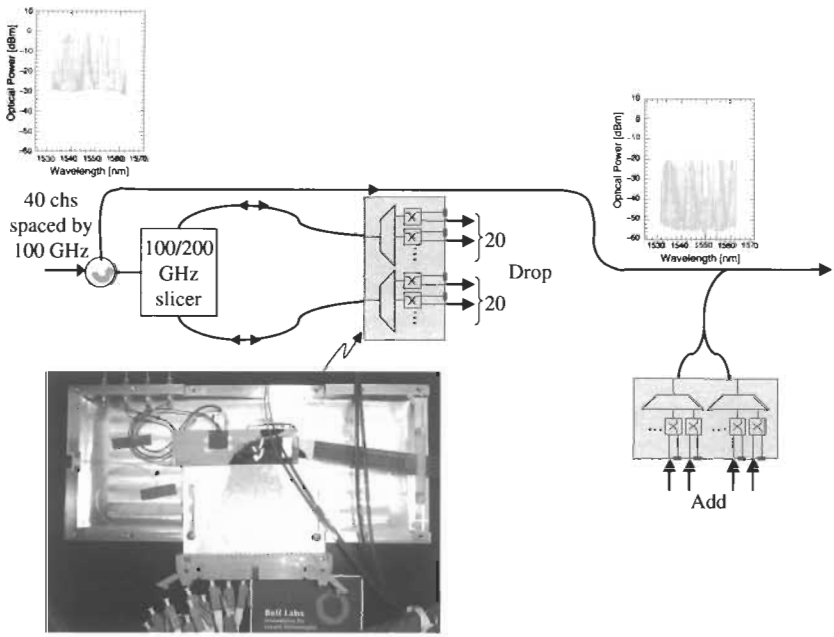


Fig. 36 Experimental demonstration of 40-channel add-drop using an interleaver and two integrated 20-channel Gaussian reflective add-drops. From Ref. 62.

However, many transmission lines need an even more box-like passband. One way to do this is to use excess-loss passband flattening of the WGR, such as with a y-branch,[63] multimode interference coupler,[64] or short parabolic horn[65] at the input. However, this adds significant excess loss, typically 3 dB. These methods could be termed "single-link passband flattening," since there is only one waveguide ultimately leading out from each side of each WGR per channel.

Another way to achieve a rectangular passband is to use multiple reflection waveguides per channel (with multiple switches). If each set then has perfect sampling, the transmission spectrum from the input to the through port has a rectangular spectrum. This could be called "multilink passband flattening," since there are more than one waveguide ultimately leading out from at least one side of each WGR per channel as done in the blocker of Fig. 29. For example, a good choice is to have two lens waveguides per channel. Then the in-to-through spectrum consists of two overlapping passbands making up a flat passband. For the add and drop channels there are two choices. There is generally no room to have a multiplexer for the two channel halves; thus the two passbands must be added in a power combiner. If they are added in phase, then the output is a single centered Gaussian-type passband. If they are added with $\pm 90°$ phase, the output is two rectangular passbands, each with 3 dB excess loss, which can be used as an add port and a drop port.

One difficulty with multilink passband flattening is that the lens waveguides must be in phase. In a reflective geometry, this is difficult, because for the lens waveguides to be in phase, the polishing angle of the reflective facet must be extremely precise and there can be no dimples in the polishing near the waveguides. However, one solution is to perform hyperheating trimming (see Sec. 2.12). In such a case metal heaters are run along the lens waveguides and heated to such a high temperature that a permanent (irreversible) phase shift is obtained. Thus by monitoring with an optical spectrum analyzer, the phase can be adjusted until the passbands are flat. However, this is a time-consuming process that currently must be done with probe needles moved from trimmer to trimmer.

For the in-to-through channels, the gap between each pair can be equal to the gap inside the pair or greater (see Fig. 37). In the former case, the in-to-through spectrum can be completely flat over a series of undropped channels, giving essentially infinite cascadability.[61] However, for this full-flat design all of the lens arms must be in phase (up to 180° total of trimming is required per lens waveguide), the dropped stopbands somewhat narrow, and the adjacent channel crosstalk marginal. In the latter "gap" design there are dips between the channels.[66] The advantages are that only each pair of lens waveguides must be in phase (at most 90° total of trimming is required per lens waveguide), and the stop bands and adjacent-channel drop crosstalk are better.

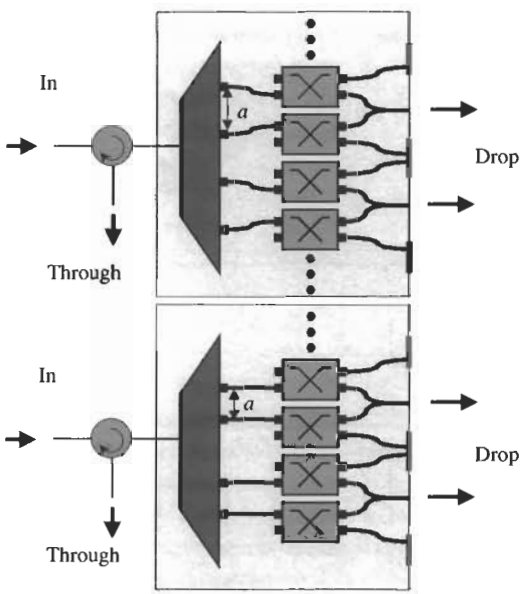


Fig. 37 Concept of multilink passband flattening for optical add-drop for the cases of (a) a full-flat spectrum and (b) a spectrum with dips between channels.

Both types of WAD were made—with and without gaps between channels. The waveguide layouts for both chips are shown in Fig. 38 and Fig. 39. Striped dielectric mirrors were deposited on the right-hand polished facets. For the full-flat device, all of the reflection waveguides were trimmed to be in phase with each other. For the gap device, only every pair was trimmed to be in phase. In both devices, the input fiber was connected directly to the free-space region of the star coupler. This allowed one to cut back the input free-space region length until the mutual coupling induced aberrations in the lens waveguides were minimized, and it allowed for precise adjustment of the center wavelength via lateral fiber movement before gluing. The measured transmissivity spectra of the two devices for in-to-though and in-to-drop are shown in Fig. 38 and Fig. 39. A systems experiment using 40 × 40 Gb/s channels was done using four full-flat WADs, showing their high cascadibility.[67]

PDL is a main difficulty with the multilink passband flattening approach to the WAD. For example, the average in-to-through PDL of the gap device is 1.3 dB for the worst channel. The lens waveguides are long, and adjacent ones bend in opposite directions at some points. Thus because of nonvertical sidewalls, small polarization conversion in bends occurs, and thus the relative phase between adjacent waveguides for TE and TM is quite different. A more robust design for multilink passband flattening is demonstrated in [45].

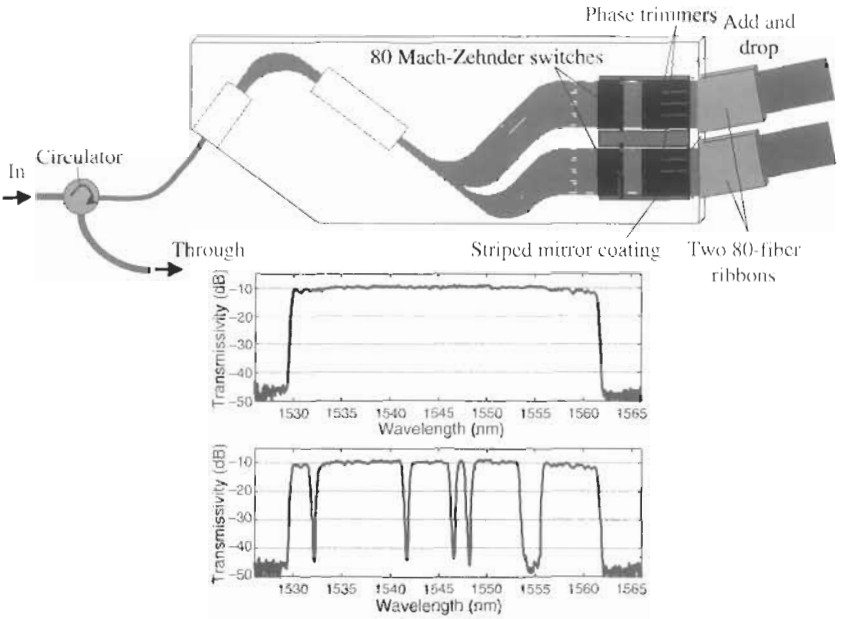


Fig. 38 Waveguide layout of 40-channel add-drop employing full-flat multilink passband flattening and the measured transmissivities. From Ref. 61.

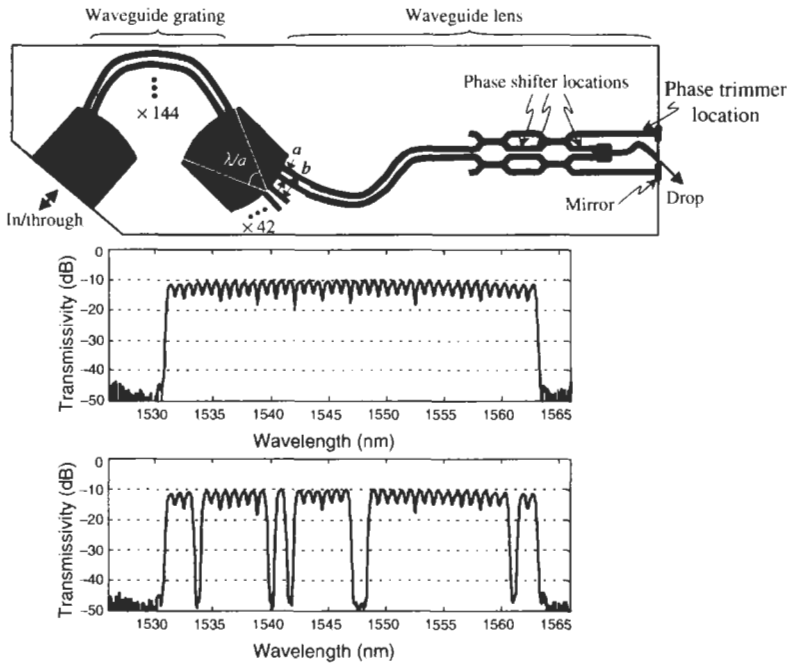


Fig. 39 Waveguide layout of 40-channel add-drop employing multilink passband flattening with dips between channels and the measured transmissivities. From Ref. 66.

2.17 DYNAMIC PASSBAND SHAPE COMPENSATOR

As the spectral efficiency of WDM networks increases, the passband widths of the filters that combine and separate the channels approach the signal bandwidth, making the requirements on the passband shapes more stringent. Transmission lines with many optical add-drop nodes are especially problematic, since small deviations from the ideal passband shape can accumulate to give significant signal distortion. Also, with the use of optical cross-connects, the filters that each channel passes through can change. Thus, it would be useful to have a filter with an adjustable passband shape for each channel that can be used in the multiplexing or demultiplexing stage to correct the signal distortions. An early proposal for such a filter consisted of two WGRs in series, one with a Gaussian passband and the other with a double-peaked passband. This could control the curvature of a single passband in a slow manner.[68]

Figure 40 shows a silica filter that can control the curvature, tilt, and attenuation of 40 passbands independently as well as act as a multiplexer.[69] It could also be used as a demultiplexer, but the advantages of using it as a multiplexer are that polarization dependence does not matter if one uses polarization-maintaining fiber connections to the sources, crosstalk does not matter, and the attenuation capability can be used to equalize the launched channel powers.

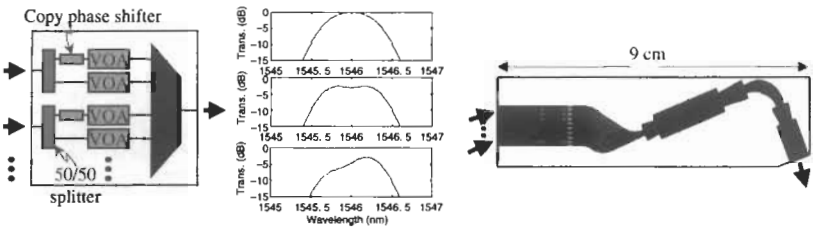


Fig. 40 Concept and waveguide layout of a 40-channel dynamic passband shape compensator. From Ref. 69.

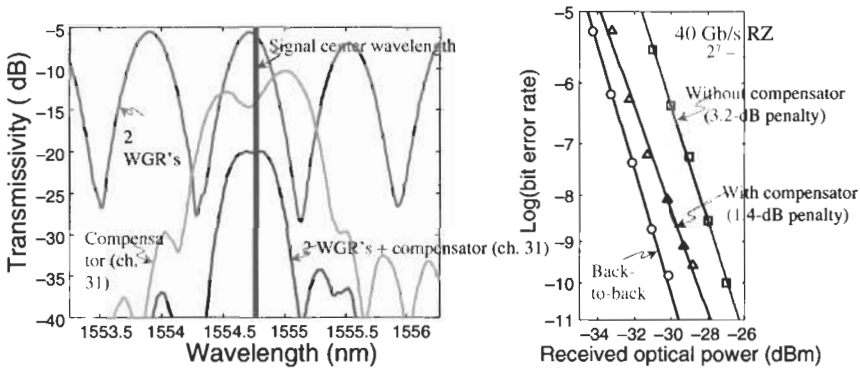


Fig. 41 Experimental results of dynamic passband shape compensating a 40-Gb/s return-to-zero signal. From Ref. 69.

Figure 41 shows the experimental results of using the compensator to improve the transmission performance of a 40-Gb/s return-to-zero signal through two Gaussian passbands offset from the channel with a net 3 dB passband width of 0.29 nm. After inserting the compensator the passband width broadened to 0.40 nm and became centered, and the bit-error rate penalty at 10^{-9} reduced from 3.2 dB to 1.4 dB.

2.18 DYNAMIC DISPERSION COMPENSATOR

A ring resonator is a waveguide that closes on itself, making a loop. Coupling a ring resonator to a waveguide creates a filter that ideally has no amplitude change with wavelength, just phase change, i.e., an "all-pass" filter for the signal passing through the waveguide.[70] If the complex amplitude transmissivity of any filter is $t(\lambda)$, then the group delay is

$$\text{delay} = \frac{-\lambda^2}{2\pi c_0} \frac{d\angle t(\lambda)}{d\lambda} \tag{37}$$

The group delay vs. wavelength through a straight waveguide coupled to a ring looks like a periodic pulse train. The chromatic dispersion is the derivative of group delay with respect to wavelength and usually has the units of ps/nm. If there are three or more rings in series with slightly different sizes and finesses, the net group delay vs. wavelength can have a linear region, the resulting curve looking like a sawtooth, where the period of the sawtooth is equal to the nominal free-spectral ranges of the rings.[70] By making the free-spectral range equal to or a subinteger divisor of the channel spacing and centering the channels on the linear regions, one has a device that applies second-order chromatic dispersion to all the channels at once. Furthermore, it can be tunable by placing phase shifters that control the ring sizes and couplings to the waveguide.[71] System tests have shown up to 4000 ps/nm tuning range.[72]

Two main difficulties, however, are getting the rings to be small enough to have their free-spectral ranges match the channel spacing and polarization dependence. The ring size has been reduced to the point such that the FSR is 50 GHz by using SiON waveguides with an index step of 3.3%.[73] One difficulty that arises with a high index step is that the fundamental mode size is reduced, no longer matching that of standard optical fiber. Thus either spot-size converters, high numerical aperture fibers, or a lensing system must be used to avoid high coupling losses to fiber. For polarization dependence, one solution is to use a substrate and upper cladding that have the same thermal expansion coefficient. Other solutions are polarization diversity, or, if the device is used at the transmitter and is attached via polarization-maintaining fiber, the polarization dependence may not matter.

2.19 PLANAR BRAGG GRATINGS

A Bragg grating is a grating that reflects directly back on itself. It thus turns a grating from being a normally finite-impulse response filter to being an infinite-impulse response filter. It consists of a corrugation in the refractive index, and the amplitude reflectivity Γ can be calculated by using the following recursive equation, applying the equation once per period of the grating:

$$\Gamma_{new} = r - \frac{\Gamma_{old}(1 - r^2)e^{j\beta 2L}}{1 - re^{j\beta 2L}} \tag{38}$$

where r is the amplitude reflectivity of each corrugation, and L is the period length.

Bragg gratings in planar silica waveguides are generally made by hydrogen loading the waveguides and then exposing to a diffraction pattern of ultraviolet light.[74,75] Such gratings have had poor performance to date, though, compared to their fiber relatives, exhibiting large polarization dependence and significant out-of-band ripple on the short-wavelength side. The polarization dependence arises from the strain in the glass to due to thermal expansion mismatch between the upper cladding glass and the substrate.

2.20 ADVANTAGES OF SILICA

The three main advantages of silica waveguides are the ease of index and mode matching to standard optical fiber, the very low propagation loss, and durability, leading to very low insertion loss reliable integrated devices. As ultralong-haul transmission systems become more prevalent, the need for extremely low insertion loss filters will increase. This is because such transmission systems are often limited by signal-to-noise ratio.

3 Polymer

Polymer waveguides are essentially plastic lightwave circuits. The refractive indices are similar to that of silica, and so the circuits have roughly the same dimensions and are buried rectangular-core cross-sections, as in silica. The insertion loss tends to be significantly higher, though, with many absorption peaks due to the rich molecular structures in polymer. The best have a propagation loss of ~0.2 dB/cm in the 1530–1580-nm band.[76] Also, whereas silica is impervious to nearly all chemicals and can withstand temperatures over 800°C, polymer must in general be protected from the environment.

3.1 WAVEGUIDE STRUCTURE

Polymer waveguides are typically buried rectangular cores, as in silica. Some can be processed by directly exposing the polymer with ultraviolet light. The layers, including the upper cladding, are typically spun on in a liquid form and then baked. This can result in very planar layers, even for the upper cladding. Eight-channel WGRs have been successfully made in polymer waveguides with losses of 5.8–7.5 dB.[77]

3.2 PHASE SHIFTER

Probably the main reason to use polymer is the phase-shifter performance. For a thermooptic phase shifter, many polymers have 20 to 30 times the change in index vs. temperature as silica (-2 to -3×10^{-4}/°C, note the opposite sign), greatly reducing the thermooptic power consumption.[76] More importantly, unlike in silica, which, even with poling, has a very weak electrooptic effect,[78] polymers can be made with a large electrooptic coefficient, reportedly even larger than $LiNbO_3$.[79] This allows for an ultrafast, zero-static-power-dissipation phase shifter, suitable even for a data modulator. Unfortunately, electrooptic polymers currently have a high propagation loss.

3.3 ADD-DROP

Another advantage of polymer waveguides is that one can write a Bragg grating via exposure to ultraviolet light in both the core and the cladding,

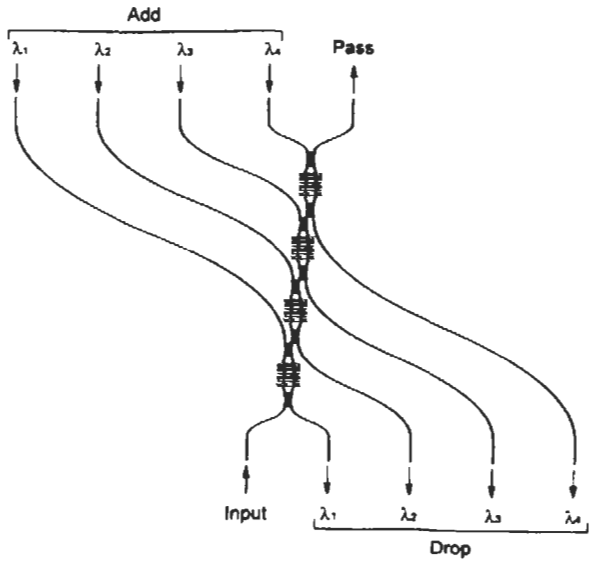


Fig. 42 Layout of polymer 4-channel add-drop using planar Bragg gratings. From Ref. 81.

unlike in fiber or silica waveguides. Having the grating exist both in the core and cladding helps to avoid the coupling of modes from the core to the cladding that gives ripple in the transmission spectra of ultraviolet-written Bragg gratings in fiber and silica waveguides on the short wavelength side. A single-channel add-drop using a planar, widely temperature-tunable Bragg grating and two separate circulators[80] and a four-channel add-drop using Mach-Zehnder interferometers with Bragg gratings in both arms[81] (see Fig. 42) have been demonstrated. The large index change with temperature allowed the grating in the single-channel device to be tuned over 20 nm.

3.4 ADVANTAGES OF POLYMER

The main advantages of polymer are its ease of processing with simple equipment, its potential capability for an efficient electrooptic phase shifter, its reduced-power thermooptic phase shifters, and its ripple-free Bragg gratings.

4 Silicon on Insulator

Silicon waveguides were started with the belief that they could take advantage of computer chip industry developments. One can easily integrate CMOS transistors with the waveguides and possibly use low-cost equipment to make the waveguides. The waveguides are made in crystalline silicon, and so can

have significant optoelectronic effects. However, because silicon has an indirect bandgap, optical control devices in silicon that use carrier injection are slow because of the long carrier lifetime; and because silicon is centrosymmetric there is no linear electrooptic effect. Silicon has a high index of refraction (3.5 at 1.55 μm) and cleaves poorly, so polishing and an anti-reflection coating is needed to couple the waveguides to fiber.

4.1 WAVEGUIDE STRUCTURE

Waveguides in silicon on insulator (SOI) are usually made by first taking a silicon wafer and depositing a thin layer of silica on it. This is then wafer-bonded (by having two perfectly flat and clean surfaces and obtaining adhesion by stiction) to another silicon wafer. One of the silicon wafers is then polished down to a typical thickness of 5 μm. This silicon provides the vertical optical confinement. The silicon is a crystal with a bandgap of ~1.1 μm and thus has relatively low loss (0.1 dB/cm reported)[82] at 1.55 μm. The horizontal confinement is achieved by etching a ridge, typically 2 μm high (see Fig. 43).

Because the index of silicon is so high, one might expect that in order to make a single-mode waveguide, the silicon layer must be extremely thin, but this is not so.[83] As long as the rib height is less than half of the total silicon thickness, then the waveguide can be single-moded even with a silicon thickness that is highly vertically multimoded when the rib is not present. Higher-order slab modes are simply not guided by the rib (although they are not easy to get rid of).

WGRs have been made in SOI using such an air-clad rib waveguide.[84]

4.2 PHASE SHIFTER/ABSORBER

Like silica, silicon is centrosymmetric and so has no Pockels effect. One can inject carriers into silicon to change its index and/or increase the absorption; this is sometimes called the plasma dispersion effect.[85,86] This effect is slow due to the long lifetime of an indirect bandgap material such as silicon. The carriers are usually injected by forming a p-n junction (see Fig. 44). To keep down the absorption losses from the doped silicon, the p-n junction is usually horizontal, i.e., the p-doped region is on one side of the waveguide, and the n-doped region is on the other.

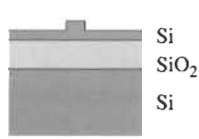


Fig. 43 Silicon-on-insulator ridge waveguide.

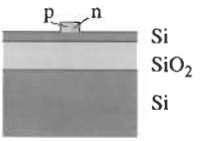


Fig. 44 Silicon-on-insulator plasma-based phase shifter.

One can also make a thermooptic phase shifter in silicon waveguides.[87] The index change of silicon with temperature is large, and so at first one might expect the required heater power to be much lower than that of silica. However, silicon is also a good thermal conductor, so the heat is quickly spread, making the power consumption high.

4.3 ADVANTAGES OF SILICON ON INSULATOR

The main advantages of SOI waveguides are a potential for low-cost processing, a smaller device size because of the high refractive index, the ability to integrate with silicon electronics and the opportunity for carrier-injection-induced absorption.

5 Indium Phosphide

Indium phosphide is a crystal and is the composition of the substrate and cladding. The core usually has the addition of gallium and arsenic. Like silicon, InP is a semiconductor, but unlike silicon, indium phosphide has a *direct* bandgap. Thus indium phosphide can give efficient stimulated emission, efficient detection, and efficient electrooptic and carrier-induced effects. Also, unlike silicon, indium phosphide cleaves well, eliminating the need for polishing. Finally, InP is not centrosymmetric and so has a linear electrooptic effect. Unfortunately, indium phosphide is delicate and expensive to grow. Nevertheless, essentially every existing telecommunications laser uses indium phosphide.

5.1 WAVEGUIDE STRUCTURE

The two most popular passive waveguide structures are shown in Fig. 45. The buried rib-loaded slab generally has a larger bend radius but lower propagation loss and can be wet-etched, rather than dry-etched like silica or silicon. Air-clad rib indium phosphide waveguides must be dry-etched. Dry etching requires a reactive ion etcher, which can be difficult to obtain good results with. Typical propagation losses are 0.2 dB/cm and 2 dB/cm for wet- and dry-etched waveguides, respectively.

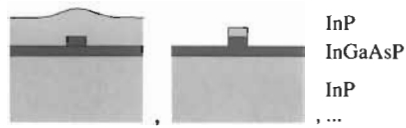


Fig. 45 Indium phosphide waveguides.

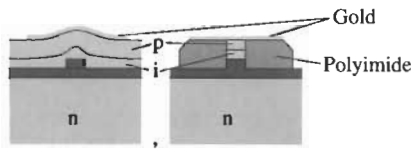


Fig. 46 Indium phosphide phase shifters.

5.2 PHASE SHIFTER

The phase shifter (see Fig. 46) can operate either by electric field or carrier injection.[88,89] For an applied electric field, the refractive index generally increases, due to the Pockels effect (linear electrooptic effect, due to the fact that InP is acentric) and carrier effects. This effect is very fast (>50 GHz) and low power. For carrier injection, the refractive index generally decreases. This effect is slower (<1 GHz) and power-inefficient. For reverse voltage electric field application, one can achieve polarization independence by tilting the phase shifter with respect to the InP crystal axes.[90] The dry-etched structure on the right in Fig. 46 usually has a lower capacitance, leading to higher-speed operation.

5.3 WAVELENGTH-SELECTIVE CROSS-CONNECT

WSCs in indium phosphide have been demonstrated using the 2 × 2 switch approach and using the interleave-chirped approach.[91,92] A 4-channel WSC using the 2 × 2 switch approach with a size of only 3 mm × 1.5 mm has been demonstrated.[93]

5.4 BEAM-STEERING SPATIAL CROSS-CONNECT

In the silica section, we described devices consisting of up to three cascaded waveguide arrays joined by star couplers. Here, we describe a device consisting of six cascaded waveguide arrays: an optical spatial cross-connect. It is nonblocking. The advantage of it being in InP is that it could potentially switch fast enough to be used for packet routing. The main idea is shown in Fig. 47. It consists of two arrays of cascaded beam steerers arranged around a large central star coupler.[94] A connection is made between a port on the left and one on the right by steering their beams at each other. The cascading serves to cover a larger steering angle with a limited available controllable

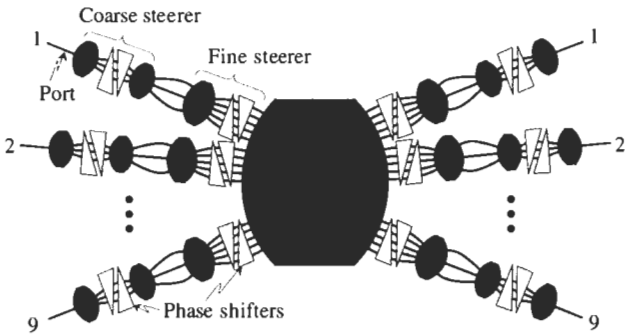


Fig. 47 Planar beam-steering spatial cross-connect.

phase shift: the coarse steerer can steer by large discrete steps in angle, and the fine steerer can stear the beam about these steps.

A version demonstrated in InP will now be described. Other strict- or wide-sense nonblocking spatial cross-connects in semiconductor material have used combinations of 1×2 or 2×2 switches or a broadcast-and-select architecture.[95] The main advantages of the beam-steering cross-connect are no waveguide crossings, which leads to a compact size, strict-sense nonblocking operation with an electrical lead number and settings number proportional to only $N \log N$ (as opposed to $\sim N^2$ for a crossbar architecture using 2×2 switches) and $N^{(1+1/\log N)} \log N$, respectively; and nonreliance on gain or loss elements.

A diagram of the detailed features of the actual waveguide layout for a 9×9 is shown in Fig. 48. The coarse and fine steerers each steer to three steering angles, giving the total of nine steering angles in the central star coupler. The lens waveguides connecting the coarse and fine steerers are uncoupled, so there are just 3 waveguides. The arrays in the coarse and fine steerers have 5 and 12 waveguides, respectively, and are weakly mutually coupled, except where the fine steerers connect to the main star coupler. There, the mutual coupling is slightly stronger. The central Brillouin zones of the central star coupler are exactly filled. The device size is 34 mm $\times$ 2 mm. The dominant contribution to the device length is the central star coupler, measuring 14.5 mm long.

The design follows the fabrication-robust guidelines for the design of the grating and lens as described in Section 2.8. The phase shifters, which perform the steering, consist of solid triangular regions of p-n junctions over the waveguides. The triangle bases and heights for the coarse and fine steering stages are 1.2 mm $\times$ 0.054 mm and 1.2 mm $\times$ 0.081 mm, respectively.

This initial demonstration used a simple fabrication technique that requires only two mask plates. First, the base wafer containing the guiding slab and rib is grown. The waveguides are patterned and wet-etched through the rib. About 0.6 μm of nonintentionally doped InP ($\sim 5 \times 10^{16}$ cm^{-3} n-type) is then

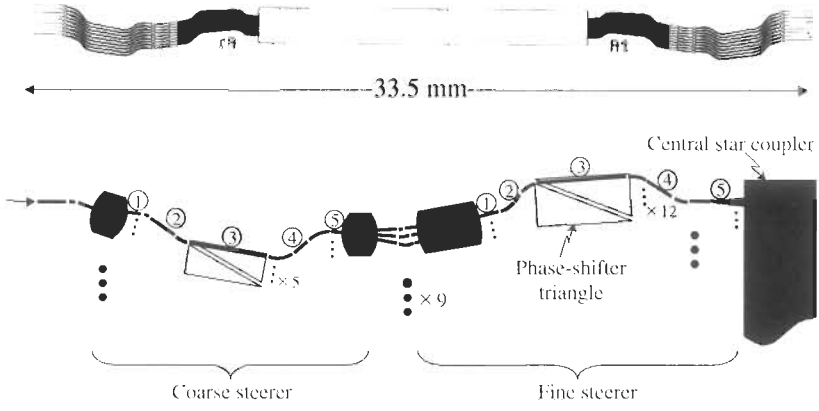


Fig. 48 Waveguide layout and details for 9 × 9 planar beam-steering cross-connect in InP.

regrown on the entire wafer. Then oxide is deposited on the wafer, and triangular openings over the beam steerer waveguide arrays are etched through the oxide. The p-doped InP is grown on the wafer, finishing with a ternary contact layer and a 15-nm InP layer. Using the same mask plate as the oxide patterning, photoresist is patterned, and gold-zinc deposited on top. However, after the photoresist patterning but before the metal deposition, the top InP layer is etched off, providing a pristine surface for the metals. Finally, the wafer backside is coated with gold. The device was not alloyed. The 72 gold triangles were not wire-bonded to anything, but instead were accessed by probe needles. No antireflection coatings were applied.

Without any electrical signal applied, each steerer was centered; so in the quiescent state all the beam steerers are pointed at the center beam steerer (port 5) of the opposite side. The beam steerers steer ~1 port/125 mA and ~1 port/50 V with reverse voltage in a triangle (i.e., the coarse stage steers by one waveguide or the fine stage steers by one port). Because using reverse bias to do the steering, the preferred method because of its zero power dissipation and high speed, requires a voltage greater than the breakdown voltage, forward current injection had to be used. This is a typical problem with using reverse voltage effects with buried waveguides—there are often high leakage and large capacitance. To achieve good steering, the injected current must be uniform over the triangle, which appears to be the case. The surprisingly good uniformity is probably due to the etching of the sacrificial InP layer before depositing the metals.

Figure 49 shows the basic switching performance and wavelength sensitivity of the device. Spontaneous emission from an Er-doped fiber amplifier was sent

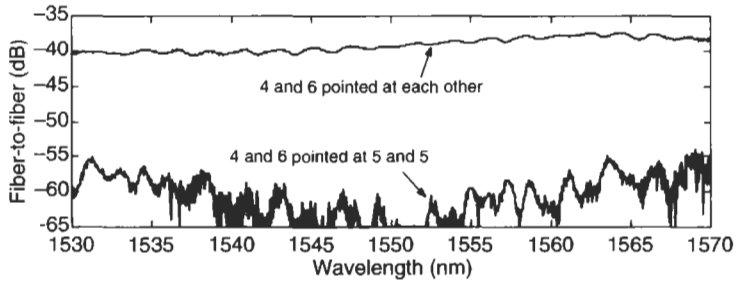


Fig. 49 Measured switching performance of the 9 × 9 spatial cross-connect.

through port 4 on the left-hand side and measured from port 6 on the right-hand side, via lensed fibers. The bottom trace shows the transmissivity when no currents are applied, and so the connection between ports 4 and 6 is not made. The top trace shows when ~125 mA was applied to the downward-steering triangle of the fine steerer of left-hand port 4, and ~125 mA is applied to the upward-steering triangle of the fine steerer of right-hand port 6, making the connection. As one can see, the extinction ratio is 15–25 dB, the highest extinction ratio occuring at ~1551 nm, near the wavelength used in designing the star couplers. Note that this is measuring a worst-case extinction ratio, which is when ports are aimed only one port away from making the connection. We tried 45 other connections (there are a total of 9! = 362880 connections) and measured crosstalk levels of −5 to −25 dB. The fiber-to-fiber insertion loss is so high partly because the phase distortion in the mutually coupled inlets in the central star coupler was not compensated for. Also, employing segmentation in all the star couplers could make a dramatic reduction in the loss.

5.5 SEMICONDUCTOR OPTICAL AMPLIFIER

Because indium phosphide is a direct bandgap semiconductor, it can give efficient stimulated emission. A semiconductor optical amplifier (SOA) is basically indium phosphide with gallium and arsenic to set the bandgap near the emission wavelength inside a p-n junction. SOAs are key elements in lasers, described next.

5.6 FAST TUNABLE LASERS

The main advantages of InP are optical gain and fast switching. The main disadvantages are polarization dependence and nonlinearities. A fast tunable laser is a highly suitable application for InP because it uses both advantages and does not care about polarization dependence or nonlinearities.

5.6.1 Distributed Bragg Reflector Lasers

Distributed Bragg reflector (DBR) lasers consist of a gain section, one or two tunable Bragg grating mirrors, and sometimes a coarse tunable filter. There are three main types.

The first is the basic DBR laser,[96] consisting of a tunable Bragg reflector, a gain section, a phase tuning section, and a reflective facet. The tunable Bragg reflector is usually a 1.4 μm bandgap Bragg grating. When current is injected, the index of the grating changes, moving the reflection peak wavelength. The phase shifter is needed to position a cavity mode wavelength under the reflection peak to give a high side-mode suppression ratio. The tuning range is limited to about 10 nm.

The second is the vertical coupler DBR laser,[97,98] in which the Bragg grating is missing teeth at periodic intervals, making the reflection spectrum consist of a periodic train of peaks. The vertical coupler has Bragg gratings that couple vertically between stacked waveguides. Such a transmissivity spectrum is very broad but also has a large tuning range. By putting both the reflection comb filter and the coarse filter in the same cavity, one can achieve different wavelengths by different combinations of the two tuning currents. Again, a phase shifter is also required to achieve a high side-mode suppression ratio.

The third is the sampled grating DBR laser,[99] in which, there are two sampled Bragg gratings, each with a different period of missing teeth, one at each end of the cavity. By adjusting the two tuning currents, different reflection peaks overlap giving different laser wavelengths. Once again, a phase shifter is required for achieving a high side-mode suppression ratio.

All of these lasers are tuned via analog currents, which can be difficult to generate by high-speed electronics, and the optimum currents are often hysteretic and sensitive to aging and temperature.

5.6.2 Multifrequency Laser

By combining a WGR with SOAs on one or both sides, one can make a series of lasers which share the same grating[100] (see Fig. 50). This concept is termed a multifrequency laser or MFL. The concept was originally demonstrated using a reflective diffraction grating.[101] The conventional MFL consists of a WGR with a shared port on one side comprising one end of the Fabry–Perot cavity, and N array ports containing SOAs comprising the other.[102] Turning on an SOA commences laser oscillation at a particular wavelength as determined by the router. The conventional MFL thus has N possible oscillation wavelengths.

By using N_1 ports with SOAs on one side and N_2 ports with SOAs on the other, one can have N_1N_2 oscillation wavelengths, taking advantage of the $N_1 \times N_2$ properties of the WGR (see Sec. 2.7.2). A hybrid $N_1 \times N_2$ laser (32 channels) using a silica WGR, mechanical switches, and discrete SOAs has been demonstrated.[103] However, this laser had to use a large WGR with an FSR of N_1N_2, had high side modes from unwanted grating orders and did

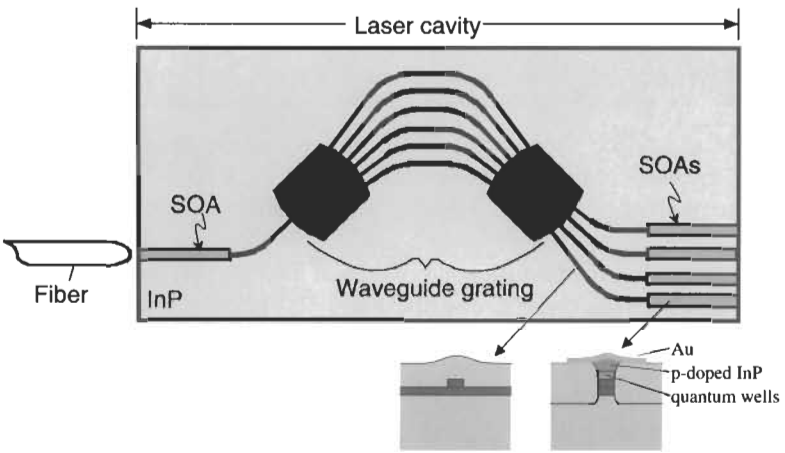


Fig. 50 Concept of WGR multifrequency laser.

not achieve single-mode operation. One can use chirp in the WGR design to significantly reduce the WGR size and suppress unwanted grating orders, greatly facilitating the achievement of single-mode operation.[104]

There are countless ways one can choose the arrangement of the ports on both sides of the router to make an $N_1 \times N_2$ laser. Suppose one wishes to make a laser with 40 evenly spaced channels (with no gaps), a small router physical size, and a reasonable method for coupling all the channels into a single output waveguide.

A necessary condition to prevent gaps is that at least two ports on one side must be separated by only one channel spacing, and we can improve the efficiency of the output coupling scheme by having all the ports on one side be separated by only one channel spacing. Thus the geometry is N_2 ports spaced by one channel spacing on one side of the router and N_1 ports spaced by N_2 channels on the other. The total number of channels is N_1N_2, and the simplest design would be to make the router FSR N be $\geq (N_1 - 1)N_2$. However, if we wish to maintain a narrow passband width, the number of grating arms must remain in proportion to N, and thus the router will be very large physically.

We can linearly chirp the router, causing each order of the grating to have its own focal line. We can have different ports access different grating orders by placing their inlets on different focal lines, allowing for a channel span larger than the router FSR. We must avoid having a port on a focal line closer to the router block the radiated mode from a port on a more distant line. Two solutions for N that maximize the distances between port inlets, minimizing the blocking are

$$N = \frac{N_1 N_2}{P} \quad \text{where} \quad \frac{N_1}{P} \neq (\text{integer} < N_1) \tag{39}$$

or

$$N = \frac{(N_1 - 1)N_2}{P} \quad \text{where} \quad \frac{N_1 - 1}{P} \neq (\text{integer} < N_1 - 1) \tag{40}$$

The latter should be used if the user is willing to have channels on the very edges of the star-coupler Brillouin zones. P is a user-chosen integer $\leqq N_1, N_2$ and is the number of focal lines used by the N_1 ports.

Consider a design with $N_1 = 5, N_2 = 8$, and $P = 2$ and no channels on the very edges of the Brillouin zone. The inlet layout is shown in Fig. 51. The channel spacing is 100 GHz, and there are 100 grating arms. More design details are given in [104].

For the output port, the design uses a waveguide that is next to the grating arm inlets in star coupler 2. It is tapered to a wide width at the free-space region so that its radiation pattern is centered on the 8-port inlets of star coupler 2. The coupling from any of the 8 ports to the output is estimated to be −20 dB. The output has an SOA to boost the power. The waveguide layout is shown in Fig. 51.

Figure 52 shows the oscillation spectra of each of the 40 upper-lower SOA combinations, taken one at a time, as measured from the output of an unpackaged chip using a lensed fiber. Note that in this initial device, the gain peak of the SOAs is at too high a wavelength, ~1590 nm, so the higher wavelength channels have significantly more power.

The device was fully packaged with a high-reflection coating and an antireflection coating on the left and right facets, respectively. We then studied the longitudinal mode behavior. It was found that the residual reflection from the

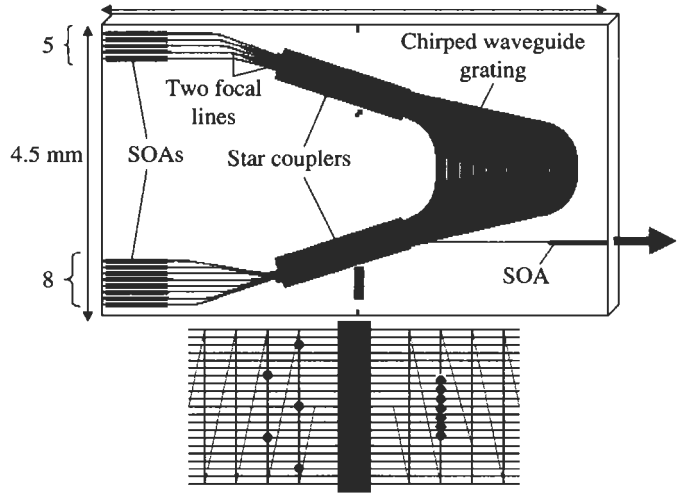


Fig. 51 Waveguide and inlet layouts of 40-channel multifrequency laser. From Ref. 104.

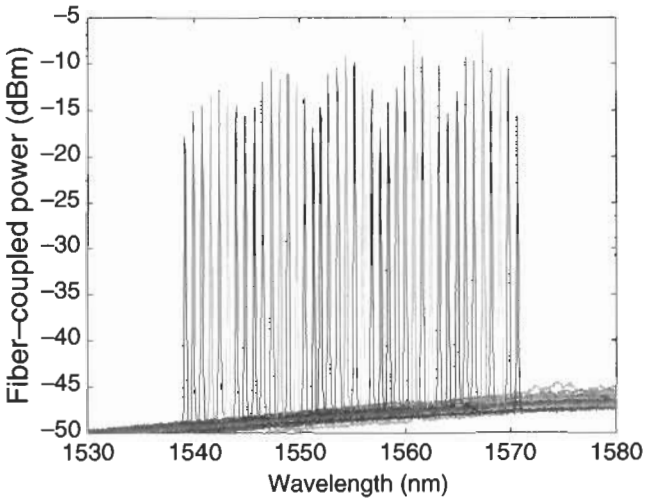


Fig. 52 Laser oscillation spectra from 40-channel multifrequency laser. From Ref. 104.

output facet can either enhance or destroy single-longitudinal mode operation, depending on the phase of the returning extracavity reflection, adjustable via the output SOA current, as analyzed in a hybrid integration study. To guarantee single-mode operation, one could either try to eliminate the reflection, or, since this laser is best suited for operation of only one channel at a time (turning on multiple channels simultaneously leads to uncontrollable behavior), one could have a lookup table for the output SOA current that depended on the activated channel. The single-mode current range was large enough (~40 mA) that the output power levels can still be reasonably matched.

5.7 NONLINEARITIES IN InP

SOAs are highly nonlinear and exhibit gain compression with the fast time constant of typically 2 ns. Even passive InP waveguides are nonlinear. Passive waveguides exhibit a strong and ultrafast $\chi^{(3)}$, with a typical n_2 of $+1 \times 10^{-12}$ cm^2/W,[105] most likely due to two-photon absorption. Wave mixing between channels in an MFL can cause instabilities when a mixing product has the same wavelength as a channel undergoing laser oscillation.[105]

A devastating slower nonlinearity triggered by two-photon absorption can occur in passive InP waveguides at powers as low as 25 mW.[106] Even though 1.55 μm photons have an energy lower than the bandgap of the 1.3 μm-bandgap guiding layer, two photons can act together to be absorbed. This generates a high-energy ("hot") carrier. One hot carrier seems to be able to be involved in the absorption of many other photons. The energy from most of these photons end up in lattice vibrations. This results in signal distortion and

crosstalk. Thus InP devices that handle multiple wavelength channel signals will be very challenging to make without nonlinear crosstalk.

5.8 ADVANTAGES OF InP

The main advantages of indium phosphide derive from its having a direct bandgap. The direct bandgap allows for stimulated emission, high-speed low-voltage electroabsorption and electrorefraction, and so on. However, it has a high nonlinearity and large polarization dependence depending on the waveguide structure. Thus for WDM applications, it is currently best suited for lasers and modulators. It is also excellent as a receiver, both surface normal and waveguide.

6 Lithium Niobate

The substrate is single-crystal lithium niobate, and the waveguides are usually part of the substrate with indiffused titanium. Lithium niobate has relatively low propagation loss in the telecommunications band (~0.2 dB/cm) yet has extremely high linear electrooptic coefficients (it is acentric). It can provide very fast and efficient phase shifters and polarization converters.

6.1 WAVEGUIDE STRUCTURE

The most common $LiNbO_3$ waveguide is the Ti-diffused waveguide.[107] Onto a substrate of $LiNbO_3$ (pure crystal), Ti strips are deposited in the desired waveguide pattern. Then the device is heated, and the Ti diffuses into the $LiNbO_3$ making a semicircular waveguide. This core is not buried, so care must be taken to keep electrodes on top from absorbing the light. The typical waveguide structure is shown in Fig. 53.

6.2 PHASE SHIFTER

The phase shift arises solely from the Pockels effect. One simply places two gold lines, around the waveguide (see Fig. 54). The electric field then changes the refractive index. 40 Gb/s modulators have been demonstrated.

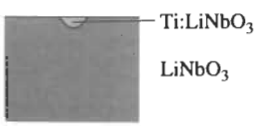


Fig. 53 Ti-diffused $LiNbO_3$ waveguide.

Fig. 54 $LiNbO_3$ phase shifter.

6.3 DIGITAL SWITCH

One can make a digital 2 × 2 switch in $LiNbO_3$ by placing electrodes on the waveguides with the same width in an adiabatic coupler (see Sec 2.10).[108] When no voltage is applied, the coupler acts as a 50/50 coupler, as expected. When a voltage is applied to one waveguide, the switch is in either a bar or cross state, and vice versa with application to the other waveguide. Once sufficient voltage is applied to make the two waveguides no longer identical, a further increase in voltage does not further affect the switching. Such switches have been combined in fabrics to make fully integrated spatial cross connects as large as 16 × 16.[109]

6.4 ACOUSTOOPTIC FILTER

A device that has both optical and acoustic waveguides is the acoustooptic filter in $LiNbO_3$. $LiNbO_3$ waveguides have different effective indices for TE- and TM-polarized light. However, by applying an acoustic wave over the waveguide, the periodic nature of the wave can cause the TE and TM lights to be phase-matched at a certain wavelength. Then a diagonal-oriented birefringence becomes significant, causing polarization conversion at that wavelength.

One way to couple the acoustic wave to the $LiNbO_3$ is to use a surface acoustic waveguide, made up of a thin film of a transparent material such as SiO_2 or In_2O_3-doped SiO_2.[110] So if one makes such an acoustic waveguide over an optical waveguide and launches a traveling acoustic wave with a transducer, the polarization in the optical waveguide will be converted at a certain wavelength. The wavelength response will have large sidelobes, however, without apodization. Apodization can be accomplished by tilting the acoustic waveguide with respect to the optical waveguide.

The polarization-converted light is, unfortunately, Doppler-shifted by the acoustic wave. One way to avoid this is to have the shifted light pass through another acoustic wave that is traveling in the opposite direction, in a second stage. Several stages must be used to achieve low crosstalk anyway.

One can launch a sum of acoustic frequencies in the waveguide making a series of filters, such as for dropping multiple channels in a WDM system. When there is only acoustic frequency, then light that was polarization converted and Doppler shifted has the opposite Doppler shift if it happens

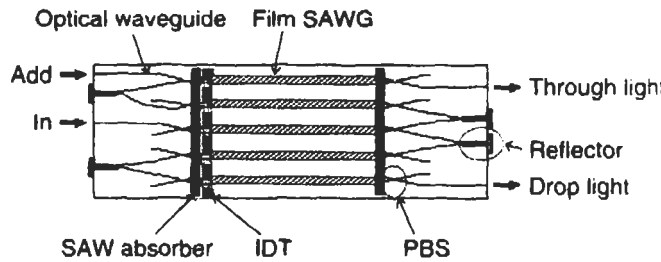


Fig. 55 Acoustooptic tunable add-drop filter. From Ref. 110.

to couple back into the original field. However, with multiple acoustic frequencies, polarization-converted light could be Doppler shifted back into the original field with a different Doppler shift, resulting in intensity beating noise.

Figure 55 shows the waveguide layout and filter characteristics of an add-drop made using acoustooptic polarization converters.

6.5 QUASI-PHASE MATCHED WAVELENGTH CONVERSION

If two wavelengths in the 1.5 μm band are sent through $LiNbO_3$, a difference frequency between the two 1.5 μm-band wavelengths is efficiently generated. However, because of the chromatic dispersion in lithium niobate, the mixing product wavelength soon falls out of phase with newly generated photons and the net wavelength conversion efficiency is poor. However, by periodically changing the sign of the electrooptic effect, the phase of the mixing wavelength can be brought back in phase, allowing its power to build up, resulting in efficient wavelength conversion. This is the idea behind quasi-phase matched wavelength conversion in lithium niobate and can be used to make efficient nonlinear devices.

6.6 ADVANTAGES OF LITHIUM NIOBATE

The primary advantage of $LiNbO_3$ is its large electrooptic effect. There are polymers which are reported to have a larger electrooptic coefficient, but they currently cannot match the reliability of $LiNbO_3$ and generally have high loss.

7 Conclusion

The main strengths of integrated devices are the ability to have long interaction lengths, their compact size, their often low insertion loss, and, most importantly, their large-volume manufacturability. Although integrated optics is about saving cost, its other attributes are making it currently successful. To build a user's confidence in a device, one must make many of them. That is usually easy with integrated optics. Thus, we will see that integrated devices

that are easy to make with consistent performance, plus have low insertion loss and a compact size, will become prevalent in optical networks.

Acknowledgments

The author is indebted to many researchers at Bell Laboratories, Lucent Technologies and Agere Systems. They include Larry Stulz, Martin Zirngibl, Charles Joyner, Corrado Dragone, Ivan Kaminow, S. Chandrasekhar, Lou Gomez, Mark Cappuzzo, Rob Pafchek, Gail Bogert, Alex Paunescu, Ed Laskowski, Alice White, Sanjay Patel, Hyang-Kyun Kim, John Gates, Pierre Schiffer, Pietro Bernasconi, and Dries Van Thourhout, but there are many others. He especially thanks Dries and Pietro for their proofreading help, Jeanette Fernandes for her assistance, and he is indebted to his family, Nenko and Hanako.

References

[1] R. Ryf, et al., "1296-port MEMS transparent optical crossconnect with 2.07 Petabit/s switch capacity," Optical Fiber Communication Conference, PD28, 2001.

[2] A. R. Ranalli, B. A. Scott, and J. P. Kondis, "Liquid crystal-based wavelength selectable cross-connect," European Conference on Optical Communications, I-68, 1999.

[3] R. C. Alferness, Plenary talk, ECIO 2001, Paderborn, Germany.

[4] H. Kogelnik, "Theory of optical waveguides," in *Guided-Wave Optoelectronics*, Springer-Verlag, New York, 1990.

[5] C. Dragone, "Optimum planar bends," Electron. Lett., vol. 29, pp. 1121–1122, 1993.

[6] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2nd edn, New York: Academic, 1991.

[7] C. H. Henry and B. H. Verbeck, "Solution of the scalar wave equation by two-dimensional Fourier analysis," J. Lightwave Technol., vol. 7, pp. 1379–1385, 1989.

[8] R. Pregla and W. Pascher, "The method of lines," in *Numerical Techniques for Microwave and Millimeter Wave Passive Structures*, T. Itoh, ed., Wiley, New York, 1989, pp. 381–446.

[9] C. L. Xu, W. P. Huang, and S. K. Chaudhuri, "Efficient and accurate vector mode calculations by beam propagation method," J. Lightwave Technol., vol. 11, pp. 1209–1215, July 1993.

[10] S. Suzuki, S. Sumida, Y. Inoue, M. Ishii, and Y. Ohmori, "Polarization-insensitive array-waveguide gratings using dopant-rich silica-based glass with thermal expansion adjusted to Si substrate," Electron. Lett., vol. 33, pp. 1173–1174, 1997.

[11] E. Wildermuth, C. H. Nadler, M. Loaker, W. Hunziker, and H. Melchior, "Penalty-free polarization compensation of SiO_2/Si arrayed waveguide grating multiplexers using stress release grooves," Electron. Lett., vol. 34, pp. 1661–1663, 1998.

[12] Y. Inoue, M. Itoh, Y. Hashizume, Y. Hibino, A. Sugita, and A. Himeno, "Novel birefringence compensating AWG design," Optical Fiber Communication Conference, WB4, 2001.

[13] K. Okamoto, *Fundamentals of Optical Waveguides*, San Diego, CA: Academic, 2000.

[14] G. R. Hadley, "Wide-angle beam propagation using Pade approximant operators," Opt. Lett., vol. 17, pp. 1426–1428, 1992.

[15] R. Scarmozzino and R. M. Osgood, Jr., "Comparison of finite-difference and Fourier-transform solutions of the parabolic wave equation with emphasis on integrated optics applications," J. Opt. Soc. Am. A, vol. 8, pp. 724–731, 1991.

[16] C. R. Doerr, "Beam propagation method tailored for step-index waveguides," IEEE Photon. Technol. Lett., vol. 13, pp. 130–132, 2001.

[17] P. G. Bernasconi, personal communication.

[18] C. Dragone, "Efficient N × N star coupler based on Fourier optics," Electron. Lett., vol. 24, pp. 942–944, 1988; C. Dragone, "Efficient N × N star couplers using Fourier optics," J. Lightwave Technol., vol. 7, pp. 479–489, March 1989.

[19] C. Dragone, "Optimum design of a planar array of tapered waveguides," J. Opt. Soc. Am. A, vol. 7, pp. 2081–2093, 1990.

[20] C. Dragone, "An N × N optical multiplexer using a planar arrangement of two star couplers," IEEE Photon. Technol. Lett., vol. 3, pp. 812–815, Sep. 1991.

[21] C. van Dam, A. A. M. Staring, E. J. Jansen, J. J. M. Binsma, T. van Dongen, M. K. Smit, and B. H. Verbeek, "Loss-reduction for phase-array demultiplexers using a double etch technique," in *Integrated Photonics Research 1996*, Boston, MA, 1996, pp. 52–55.

[22] A. Sugita, A. Kaneko, K. Okamoto, M. Itoh, A. Himeno, and Y. Ohmori, "Very low insertion loss arrayed-waveguide grating with vertically tapered waveguides," IEEE Photon. Technol. Lett., vol. 12, pp. 1180–1182, Sep. 2000.

[23] Y. P. Li, "Optical device having low insertion loss," U. S. Patent 5 745 618, April 28, 1998.

[24] W. K. Burns and A. F. Milton, "Waveguide transitions and junctions," in *Guided-Wave Optoelectronics*, Springer-Verlag, New York, 1990.

[25] M. K. Smit, "New focusing and dispersive planar component based on an optical phased array," Electron. Lett., vol. 24, pp. 385–386, 1988.

[26] A. R. Vellekoop and C. Dragone, "Four-channel integrated-optic wavelength demultiplexer with weak polarization dependence," J. Lightwave Technol., vol. 9, pp. 310–314, 1991.

[27] H. Takahashi, S. Suzuki, K. Kato, and I. Nishi, "Arrayed-waveguide grating for wavelength division multi/demultiplexer with nanometre resolution," Electron. Lett., vol. 26, pp. 87–88, 1990.

[28] C. Dragone, C. A. Edwards, and R. C. Kistler, "Integrated optics N × N multiplexer on silicon," IEEE Photon. Technol. Lett., vol. 3, pp. 896–899, 1991.

[29] G. Lenz, B. J. Eggleton, C. R. Giles, et al., "Dispersive properties of optical filters for WDM systems," IEEE J. Quant. Electron., vol. 34, pp. 1390–1402, 1998.

[30] P. Bernasconi, C. Doerr, C. Dragone, M. Cappuzzo, E. Laskowski, and A. Paunescu, "Large N × N waveguide grating routers," J. Lightwave Technol., vol. 18, pp. 985–991, 2000.

[31] K. Okamoto, K. Moriwaki, and S. Suzuki, "32 × 32 arrayed-waveguide grating multiplexer with uniform loss and cyclic frequency characteristics," Electron. Lett., vol. 33, pp. 1865–1866, 1997.

[32] C. R. Doerr, "Optical device having equal length waveguide paths," U.S. Patent 6 219 471, 2001.

[33] C. R. Doerr, M. Cappuzzo, E. Laskowski, A. Paunescu, L. Gomez, L. W. Stulz, and J. Gates, "Dynamic wavelength equalizer in silica using the single-filtered-arm interferometer," IEEE Photon. Technol. Lett., vol. 11, pp. 581–583, 1999.

[34] C. R. Doerr, K. W. Chang, L. W. Stulz, R. Pafchek, Qingdong Guo, L. Buhl, L. Gomez, M. Cappuzzo, and G. Bogert, "Arrayed waveguide dynamic gain equalization filter with reduced insertion loss and increased dynamic range," IEEE Photon. Technol. Lett., vol. 13, pp. 329–331, April 2001.

[35] K. Jinguji, N. Takato, A. Sugita, and M. Kawachi, "Mach-Zehnder interferometer type optical waveguide coupler with wavelength-flattened coupling ratio," Electron. Lett., vol. 26, pp. 1326–1327, 1990.

[36] B. E. Little and T. Murphy, "Design rules for maximally flat wavelength-insensitive optical power dividers using Mach-Zehnder structures," IEEE Photon. Technol. Lett., vol. 9, pp. 1607–1609, 1997.

[37] R. Adar, C. H. Henry, R. F. Kazarinov, R. C. Kistler, and G. R. Weber, "Adiabatic 3-dB couplers, filters, and multiplexers made with silica waveguides on silicon," J. Lightwave Technol., vol. 10, pp. 46–50, 1992.

[38] L. B. Soldano, F. B. Veerman, M. K. Smit, B. H. Verbeek, A. H. Dubost, and E. C. M. Pennings, "Planar monomode optical couplers based on multimode interference effects," J. Lightwave Technol., vol. 10, pp. 1843–1849, 1992.

[39] S. E. Harris, E. O. Ammann, and I. C. Chang, "Optical network synthesis using birefringent crystals. I. Synthesis of lossless networks of equal-length crystals," J. Opt. Soc. Am., vol. 54, pp. 1267–1279, 1964.

[40] Y. P. Li, C. H. Henry, E. J. Laskowski, H. H. Yaffe, and R. L. Sweatt, "Monolithic optical waveguide 1.31/1.55 μm WDM with −50 dB crosstalk over 100 nm bandwidth," Electron. Lett., vol. 31, pp. 2100–2101, 1995.

[41] T. Chiba, H. Arai, K. Ohira, S. Kashimura, H. Okano, and H. Uetsuka, "Chromatic dispersion free Fourier transform-based wavelength splitters for D-WDM," OECC, paper 13B2–2, 2000.

[42] M. Okuno, N. Takato, T. Kitoh, and A. Sugita, "Silica-based thermooptic switches," NTT Rev., vol. 7, pp. 56–63, 1995.

[43] Y. Silberberg, P. Perlmutter, and J. E. Baran, "Digital optical switch," Appl. Phys. Lett., vol. 51, pp. 1230–1232, 1987.

[44] K. Moriwaki, M. Abe, Y. Inoue, M. Okuno, and Y. Ohmori, "New silica-based 8 × 8 thermo-optic matrix switch on Si that requires no bias power," in Tech. Dig. OFC '95, San Diego, CA, March 1995, paper WS1, pp. 211–212; T. Goh, M. Yasu, K. Hattori, A. Himeno, M. Okuno, and Y. Ohmori, "Low loss and high extinction

ratio strictly nonblocking 16 × 16 thermooptic matrix switch on 6-in wafer using silica-based planar lightwave circuit technology," J. Lightwave Technol., vol. 19, pp. 371–379, 2001.

[45] C. R. Doerr, L. W. Stulz, and R. Pafchek, "Compact and low-loss manner of waveguide grating router passband flattening and demonstration in an 64-channel blocker/multiplexer," to appear in IEEE Photon. Technol. Lett.

[46] K. Inoue, T. Kominato, and H. Toba, "Tunable gain equalization using a Mach-Zehnder optical filter in multistage fiber amplifiers," IEEE Photon. Technol. Lett., vol. 3, pp. 718–720, 1991.

[47] C. R. Doerr, C. H. Joyner, and L. W. Stulz, "Integrated WDM dynamic power equalizer with potentially low insertion loss," IEEE Photon. Technol. Lett., vol. 10, pp. 1443–1445, 1998.

[48] C. R. Doerr, L. W. Stulz, R. Pafchek, L. Gomez, M. Cappuzzo, A. Paunescu, E. Laskowski, L. Buhl, H. K. Kim, and S. Chandrasekhar, "An automatic 40-wavelength channelized equalizer," IEEE Photon. Technol. Lett., vol. 12, pp. 1195–1197, 2000.

[49] G. Wilfong, B. Mikkelsen, C. Doerr, and M. Zirngibl, "WDM cross-connect architectures with reduced complexity," J. Lightwave Technol., vol. 17, pp. 1732–1741, 1999.

[50] C. R. Doerr, L. W. Stulz, M. Cappuzzo, L. Gomez, A. Paunescu, E. Laskowski, S. Chandrasekhar, and L. Buhl, "2 × 2 wavelength-selective cross connect capable of switching 128 channels in sets of 8," OFC PD, 2001.

[51] K. Okamoto, M. Okuno, A. Himeno, and Y. Ohmori, "16-channel optical add/drop multiplexer consisting of arrayed-waveguide gratings and double-gate switches," Electron. Lett., vol. 32, pp. 1471–1472, 1996.

[52] H. G. Bukkems, C. G. P. Herben, M. K. Smit, F. H. Groen, and I. Moerman, "Minimization of the loss of intersecting waveguides in InP-based photonic integrated circuits," IEEE Photon. Technol. Lett., vol. 11, pp. 1420–1422, 1999.

[53] C. G. P. Herben, X. J. M. Leijtens, P. Maat, H. Blok, and M. K. Smit, "Crosstalk performance of integrated optical cross-connects," J. Lightwave Technol., vol. 17, pp. 1126–1134, 1999.

[54] H. Li, C.-H. Lee, W. Lin, S. Didde, Y.-J. Chen, and D. Stone, "8-wavelength photonic integrated 2 × 2 WDM cross-connect switch using 2 × N phased-array waveguide grating (PAWG) multi-demultiplexers," Electron. Lett., vol. 33, pp. 592–594, 1997.

[55] C. R. Doerr, "Proposed WDM cross connect using a planar arrangement of waveguide grating routers and phase shifters," IEEE Photon. Technol. Lett., vol. 10, pp. 528–530, 1998.

[56] C. R. Doerr, L. W. Stulz, J. Gates, M. Cappuzzo, E. Laskowski, L. Gomez, A. Paunescu, A. White, and C. Narayanan, "Arrayed waveguide lens wavelength add-drop in silica," IEEE Photon. Technol. Lett., vol. 11, pp. 557–559, 1999.

[57] C. R. Doerr, C. H. Joyner, L. W. Stulz, and R. Monnard, "Wavelength-division multiplexing cross connect in InP," IEEE Photon. Technol. Lett., vol. 10, pp. 117–119, 1998.

[58] C. R. Doerr, B. W. Mikkelsen, G. Raybon, P. Schiffer, L. W. Stulz, M. Zirngibl, G. Wilfong, M. Cappuzzo, E. Laskowski, A. Paunescu, L. Gomez, and J. Gates, "Wavelength selective cross connect using arrayed waveguide lens multi-wavelength filters," Optical Fiber Communication Conference, postdeadline paper PD34, 1999.

[59] C. G. H. Roeloffzen, F. Horst, B. J. Offrein, R. Germann, G. L. Bona, H. W. M. Salemink, and R. M. de Ridder, "Tunable passband flattened 1-from-16 binary-tree structured add-after-drop multiplexer using SiON waveguide technology," IEEE Photon. Technol. Lett., vol. 12, pp. 1201–1203, 2000.

[60] K. Okamoto, K. Takiguchi, and Y. Ohmori, "16-channel optical add/drop multiplexer using silica-based arrayed-waveguide gratings," Electron. Lett., vol. 31, pp. 723–724, 1995.

[61] C. R. Doerr, L. W. Stulz, M. Cappuzzo, E. Laskowski, A. Paunescu, L. Gomez, J. V. Gates, S. Shunk, and A. E. White, "40-wavelength add-drop filter," IEEE Photon. Technol. Lett., vol. 11, pp. 1437–1439, 1999.

[62] R. Monnard, C. R. Doerr, A. Bing, L. W. Stulz, O. Clarke, S. Shunk, M. Cappuzzo, L. Gomez, A. Paunescu, and E. Laskowski, "Demonstration of OC-192 40-channel add-drop node using two integrated add-drop and dynamic equalization devices," European Conference on Optical Communication, paper 9.2.3, 2000.

[63] C. Dragone, "Frequency routing device having a wide and substantially flat passband," U.S. Patent 5 412 744, May 2, 1995.

[64] M. R. Amersfoort, J. B. D. Soole, H. P. LeBlanc, N. C. Andreadakis, A. Rajhel, and C. Caneau, "Passband broadening of integrated arrayed waveguide filters using multimode interference couplers," Electron. Lett., vol. 23, pp. 449–451, 1996.

[65] K. Okamoto and A. Sugita, "Flat spectral response arrayed-waveguide grating multiplexer with parabolic waveguide horns," Electron. Lett., vol. 32, pp. 1661–1662, 1996.

[66] C. R. Doerr, L. W. Stulz, R. Monnard, M. Cappuzzo, L. Gomez, E. Laskowski, A. Paunescu, S. Shunk, O. A. Clarke, A. Bing, and R. K. Vora, "40-wavelength planar channel-dropping filter with improved crosstalk," IEEE Photon. Technol. Lett., vol. 13, pp. 1008–1010, 2001.

[67] H. K. Kim, S. Chandrasekhar, T. Nielsen, C. Doerr, L. Stulz, L. Buhl, R. Monnard, S. Radic, and M. Zirngibl, "1.6 Tbit/s (40 x 40 Gbit/s) total capacity four-node optical networking using fully programmable A/D devices," Electron. Lett., vol. 36, pp. 1044–1045, 2000.

[68] T. Sato, A. Kaneko, and T. Horiguchi, "Novel method for controlling passband flatness in optical transparent networks with cascaded AWGs," European Conference on Optical Communication, pp. 166–167, 1999.

[69] C. R. Doerr, H. K. Kim, L. W. Stulz, M. Cappuzzo, L. Gomez, A. Paunescu, E. Laskowski, L. Boivin, and S. Shunk, "40-channel multi/demultiplexer with dynamic passband shape compensation," IEEE Photon. Technol. Lett., vol. 13, pp. 690–692, 2001.

[70] C. Madsen and G. Lenz, "Optical allpass filters for phase response design with applications for dispersion compensation," IEEE Photon. Technol. Lett., vol. 10, pp. 994–996, 1998.

[71] C. Madsen, G. Lenz, A. Bruce, M. Cappuzzo, L. Gomez, and R. Scotti, "Integrated tunable allpass filters for adaptive dispersion and dispersion slope compensation," IEEE Photon. Technol. Lett., vol. 11, pp. 1623–1625, 1999.

[72] C. K. Madsen, S. Chandrasekhar, E. J. Laskowski, K. Bogart, M. A. Cappuzzo, A. Paunescu, L. W. Stulz, and L. T. Gomez, "Compact integrated tunable chromatic dispersion compensator with a 4000 ps/nm tuning range," Optical Fiber Communication, PD9, 2001.

[73] F. Horst, C. Berendsen, R. Beyeler, G.-L. Bona, R. Germann, H. W. M. Salemink, and D. Wiesmann, "Tunable ring resonator dispersion compensators realized in high-refractive-index contrast SiON technology," European Conference on Optical Communication, PD, 2000.

[74] K. O. Hill, B. Malo, F. Bilodeau, and D. C. Johnson, "Photosensitivity in optical fibers," Ann. Mater. Sci., vol. 23, pp. 125–127, 1993.

[75] R. Adar, C. H. Henry, R. C. Kistler, and R. F. Kazarinov, "Polarization independent narrow band Bragg reflection gratings made with silica-on-silicon waveguides," Appl. Phys. Lett., vol. 60, pp. 1779–1781, 1992.

[76] L. Eldada and L. W. Shacklette, "Advances in polymer integrated optics," IEEE J. Sel. Top. Quant. Electron., vol. 6, pp. 54–68, 2000.

[77] N. Keil, H. H. Yao, C. Zawadzki, J. Bauer, M. Bauer, C. Dreyer, and J. Schneider, "Athermal polarization-independent all-polymer arrayed waveguide grating (AWG) multi/demultiplexer," Optical Fiber Communications Conference, PD7, 2001.

[78] M. Abe, T. Kitagawa, K. Hattori, A. Himeno, and Y. Ohmori, "Electrooptic switch constructed with a poled silica-based waveguide on a Si substrate," Integrated Photonics Research Conference, paper ITUDS, 1996.

[79] D. Chen, H. R. Fetterman, A. Chen, W. H. Steier, L. R. Dalton, W. Wang, and Y. Shi, "Demonstration of 110 GHz electro-optic polymer modulators," Appl. Phys. Lett., vol. 70, pp. 3335–3337, 1997.

[80] L. Eldada, R. Blomquist, M. Maxfield, D. Pant, G. Boudoughian, C. Poga, and R. A. Norwood, "Thermooptic planar polymer Bragg grating OADM's with broad tuning range," IEEE Photon. Technol. Lett., vol. 11, pp. 448–450, 1999.

[81] L. Eldada, S. Yin, C. Poga, C. Glass, R. Blomquist, and R. A. Norwood, "Integrated multichannel OADM's using polymer Bragg grating MZI's," IEEE Photon. Technol. Lett., vol. 10, pp. 1416–1418, 1998.

[82] U. Fischer, T. Zinke, J.-R. Kropp, F. Arndt, and K. Petermann, "0.1 dB/cm waveguide losses in single-mode SOI rib waveguides," IEEE Photon. Technol. Lett., vol. 8, pp. 647–649, 1996.

[83] R. A. Soref, J. Scmidtchen, and K. Petermann, "Large single-mode rib waveguides in GeSi-Si and Si-on-SiO2," IEEE J. Quant. Electron., vol. 27, pp. 1971–1974, 1991.

[84] P. D. Trinh, S. Yegnanarayanan, F. Coppinger, and B. Jalali, "Silicon-on-insulator (SOI) phased-array wavelength multi/demultiplexer with extremely low polarization sensitivity," IEEE Photon. Technol. Lett., vol. 9, pp. 940–942, 1997.

[85] G. V. Treyz, P. G. May, and J.-M. Halbout, "Silicon Mach-Zehnder waveguide interferometers based on the plasma dispersion effect," Appl. Phys. Lett., vol. 59, pp. 771–773, 1991.

[86] U. Fischer, T. Zinke, B. Schuppert, and K. Petermann, "Singlemode optical switches based on SOI waveguides with large cross section," Electron. Lett., vol. 30, pp. 406–408, 1994.

[87] P. Dansai, L. Thevenaz, and Ph. Robert, "Two different Mach-Zehdner interferometers in silicon on insulator technology using plasma dispersion effect," European Conference on Integrated Optics, pp. 175–178, 2001.

[88] K. Ishida, H. Nakamura, H. Matsumura, T. Kadoi, and H. Inoue, "InGaAs/InP optical switches using carrier induced refractive index change," Appl. Phys. Lett., vol. 50, pp. 141–142, 1987.

[89] J.-F. Vinchant, J. A. Cavailles, M. Erman, P. Jarry, and M. Renaud, "InP/GaInAsP guided-wave phase modulators based on carrier-induced effects: Theory and experiment," J. Lightwave Technol., vol. 10, pp. 63–69, 1992.

[90] R. Krähenbühl, R. Kyburz, W. Vogt, M. Bachman, T. Brenner, E. Gini, and H. Melchior, "Low-loss polarization-insensitive InP-InGaPcP optical space switches for fiberoptical communication," IEEE Photon. Technol. Lett., vol. 8, pp. 632 634, 1996.

[91] C. R. Doerr, C. H. Joyner, L. W. Stulz, and R. Monnard, "Wavelength-division multiplexing cross connect in InP," IEEE Photon. Technol. Lett., vol. 10, pp. 117–119, 1998.

[92] C. R. Doerr, C. H. Joyner, and L. W. Stulz, "Programmable 16-channel add/drop using an interleave-chirped waveguide grating router," IPR, PD1, 1998.

[93] C. Herben, *Compact Integrated Cross Connects for Wavelength-Division Multiplexing Networks*, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 2000.

[94] C. R. Doerr and C. Dragone, "Proposed optical cross connect using a planar arrangement of beam steerers," IEEE Photon. Technol. Lett., vol. 11, pp. 197–199, 1999.

[95] See, for example, M. Gustavsson, B. Lagerstrom, L. Thylen, M. Janson, L. Lundgren, A. C. Morner, M. Rask, and B. Stoltz, "Monolithically integrated 4 × 4 InGaAsP/InP laser amplifier gate switch arrays," Electron. Lett., vol. 28, pp. 2223–2225, 1992 or G. A Fish, B. Mason, L. A. Coldren, and S. P. DenBaars, "Compact, 4 × 4 InGaAsP-InP optical crossconnect with a scaleable architecture," IEEE Photon. Technol. Lett., vol. 10, pp. 1256–1258, 1998.

[96] T. L. Koch and U. Koren, "Semiconductor lasers for coherent optical fiber communications," J. Lightwave Technol., vol. 8, p. 274, 1990.

[97] I. Kim, R. C. Alferness, U. Koren, L. L. Buhl, B. I. Miller, M. G. Young, M. D. Chien, T. L. Koch, H. M. Presby, G. Raybon, and C. A. Burrus,

"Broadly tunable vertical-coupler filtered tensile-strained InGaAs/InGaAsP multiple quantum-well laser with 74.4 nm tuning range," Appl. Phys. Lett., 1993.

[98] M. Oberg, S. Nilsson, K. Streubel, J. Wallin, L. Backbom, and T. Klinga, "74 nm wavelength tuning range of an InGaAsP/InP vertical grating assisted codirectional coupler laser with rear sampled grating reflector," IEEE Photon. Technol. Lett., vol. 5, pp. 735–737, 1993.

[99] B. Mason, G. A. Fish, S. P. DenBaars, and L. A. Coldren, "Widely tunable sampled grating DBR laser with integrated electroabsorption modulator," IEEE Photon. Technol. Lett., vol. 11, pp. 638–640, 1999.

[100] B. Glance, I. P. Kaminow, and R. W. Wilson, "Applications of the integrated waveguide grating router," J. Lightwave Technol., vol. 12, pp. 957–962, June 1994.

[101] J. B. D. Soole, K. Poguntke, A. Schere, H. P. LeBlanc, C. Chang-Hasnain, J. R. Hayes, C. Caneau, R. Bhat, and M. A. Koza, "Multiple-stripe array grating in a cavity (MAGIC) laser: A new semiconductor laser for WDM applications," Electron. Lett., vol. 28, pp. 1805–1807, 1992.

[102] M. Zirngibl, C. H. Joyner, L. W. Stulz, U. Koren, M.-D. Chien, M. G. Young, and B. I. Miller, "Digitally tunable laser based on the integration of waveguide grating multiplexer and an optical amplifier," IEEE Photon. Technol. Lett., vol. 6, pp. 516–518, 1994.

[103] Y. Tachikawa and K. Okamoto, "Arrayed-waveguide grating lasers and their applications to tuning-free wavelength routing," Proc. Inst. Elect. Eng., vol. 143, pp. 322–328, 1996.

[104] C. R. Doerr, C. H. Joyner, and L. W. Stulz, "40-wavelength rapidly digitally tunable laser," IEEE Photon. Technol. Lett., vol. 11, pp. 1348–1350, 1999.

[105] C. R. Doerr, R. Monnard, C. H. Joyner, and L. W. Stulz, "Simultaneous CW operation of shared angular dispersive element WDM lasers," IEEE Photon. Technol. Lett., vol. 10, pp. 501–503, 1998.

[106] D. Van Thourhout, C. R. Doerr, C. H. Joyner, and J. L. Pleumeekers, "Observation of WDM crosstalk in passive semiconductor waveguides," IEEE Photon. Technol. Lett., vol. 13, pp. 457–459, 2001.

[107] R. V. Schmidt and I. P. Kaminow, Appl. Phys. Lett., vol. 25, pp. 458–460, 1974.

[108] Y. Silberberg, P. Perlmutter, and J. E. Baran, "Digital optical switch," Appl. Phys. Lett., vol. 51, pp. 1230–1232, 1987.

[109] T. Murphy, S. -Y. Suh, B. Comissiong, A. Chen, R. Irvin, R. Grencavich, and G. Richards, European Conference on Optical Communications, paper 11.2.2, 2000.

[110] T. Nakazawa, S. Taniguchi, and M. Seino, "Ti: $LiNbO_3$ acousto-optic tunable filer (AOTF)," Fujitsu Sci. Tech. J., vol. 35, pp. 107–112.

Chapter 10 Fiber Grating Devices in High-Performance Optical Communications Systems

Thomas A. Strasser

Photuris Inc., Piscataway, New Jersey

Turan Erdogan

Semrock Inc., Rochester, New York

I. Introduction and Brief History

This chapter provides both an overview and a technology update of a relatively recent fiberoptic component technology known as the *fiber grating*. This component is also often referred to as a "fiber Bragg grating" (FBG), "fiber phase grating," "UV-induced fiber grating," or "in-fiber grating."

A fiber grating is a holographic diffraction grating written by ultraviolet (UV) light directly into the core of an optical fiber. It is a phase grating, in that it is comprised mainly of a periodic variation of the index of refraction. Once exposed, the grating is essentially permanent. Typically the grating diffracts the main guided mode of an optical fiber into the identical backward traveling mode—that is, it functions as a spectrally selective mirror. However, it may also diffract the main guided mode into another mode of the fiber, including modes guided by the fiber cladding, as well as into radiation modes that propagate away from the fiber. In some cases the light coupled to nonidentical modes is used, while in other cases these other modes are merely a convenient route to dispose of the light, such that the grating functions as a spectrally dependent loss filter.

Why are fiber gratings important? Fiber gratings have already proven to be critical components for many applications in fiberoptic communications and sensor systems. Some of these are described in this chapter, while countless others are described in numerous papers and in several excellent review articles [1–6] and books [7,8]. Advantages of fiber gratings over competing technologies include all-fiber geometry, low insertion loss, high return loss or extinction, and potentially low cost. Additionally, fiber gratings can be made with minimal polarization dependence, passive temperature compensation, and extremely high reliability. But the most distinguishing feature of fiber gratings is the flexibility they offer for achieving desired spectral characteristics. Numerous physical parameters can be varied, including induced index change, length, apodization, period chirp, fringe tilt, and whether

OPTICAL FIBER TELECOMMUNICATIONS,
VOLUME IVA

ISBN: 0-12-395172-0

the grating supports counterpropagating or copropagating coupling at a desired wavelength. By varying these parameters, gratings can be made with normalized bandwidths ($\Delta\lambda/\lambda$) between 0.1 and 10^{-4}, extremely sharp spectral features, and the capability to tailor dispersion characteristics.

As a result of these fundamental advantages, a number of applications are made uniquely possible by fiber grating technology. For example, it is difficult to identify a technology better suited for making ultradense wavelength division multiplexing (WDM) filters for channel spacings of 25 GHz or less and channel bandwidths of 10 Gb/s or more. No other technology can match the combined simplicity, compactness, and spectral purity afforded by fiber gratings in fiber laser applications. Fiber gratings provide an incredibly simple and inexpensive solution for wavelength stabilization of optical amplifier pump lasers. Many fiber sensor applications would end up too complex to be worth pursuing were it not for the natural ability of fiber gratings to provide a highly sensitive, accurate, and simple solution. At the most fundamental level, it is difficult to imagine any other means to achieve the distributed reflection of light over extremely long lengths with essentially arbitrary control over the spatial variation of the amplitude and phase of the reflection. In short, fiber gratings open up a regime of optical filters that is simply not practical by any other technology.

With that glowing recommendation, one might be led to believe that fiber gratings are the great optical filter panacea—a solution to all the world's fiberoptic spectral control applications. A word of caution is merited: fiber gratings are often *not* the best solution. In fact, the same features that define the key advantages of fiber gratings also represent the key limitations. For example, thin-film filters are preferable to fiber gratings in many commercial applications because, if packaged correctly, they do not require an expensive circulator or suffer from the loss associated with a coupler. The integration of the grating in the fiber causes the need for the circulator but at the same time is the very basis for the fiber grating advantages of high reliability and long, uniform interaction lengths. Although fiber gratings can theoretically provide arbitrarily precise dispersion compensation in long-haul systems, currently dispersion-compensating fiber is the preferred solution because of the inherently smooth, ripple-free spectral response. Nevertheless, so long as fiber gratings are used for the right applications, they often offer a superb solution.

In order to provide context for the development of fiber grating technology, it is helpful to provide a brief overview of the history of this development. Detailed accounts of this history can be found in several locations [1–8]. Here, we highlight only a few of the significant events. As in the development of most technologies, in hindsight one can often identify several key, paradigm-shifting discoveries, each of which served to overturn the conventional wisdom of the day and open a door that led to an unexpected direction in the subsequent stage of development. In many cases, these discoveries are themselves forgotten or not actually used, but their impact on the overall path of development is lasting.

There is no question that for fiber grating technology the first significant discovery was made by Hill and his coworkers at the Communications Research Center in Canada in 1978 [9]. While studying the nonlinear properties of backscattering in a special fiber using blue light (488 nm) from a continuous-wave (CW) argon-ion laser, they observed the reflectivity of the light increased dramatically with time. Hill hypothesized that the cause of this high reflection was in fact a grating that was written into the core of the fiber by the standing-wave interference pattern formed by the incident light and the backward-propagating light reflected off of the cleaved fiber end face. A measurement of the spectral bandwidth of the reflected light proved this hypothesis to be correct. Prior to this discovery, no one had any idea that glass fibers could function as a photosensitive, holographic material. Hill's discovery opened the first door, and experimentation began.

The next major discovery occurred just a few years later, when Lam and Garside presented data indicating that the magnitude of the photoinduced index change depended on the square of the writing power of the argon-ion laser [10]. This result suggested that the photoinduced index change was a two-photon process, implying that a more substantial effect could be obtained with a single-photon process initiated by a UV laser at half the wavelength (around 244 nm). The door was now open for someone to demonstrate UV-written gratings.

Surprisingly, eight years passed before Meltz and coworkers at the United Technologies Research Center reported successful grating writing by a UV laser [11]. The researchers used a pulsed, frequency-doubled dye laser pumped by an excimer laser. This arrangement enabled high intensity to be achieved in a relatively coherent (spatially and temporally) UV laser beam. The coherence was critical, since interfering two laser beams on the side of the fiber formed the periodic intensity pattern. With the prediction of Lam and Garside, and the demonstration by the United Technologies team, a tremendous paradigm shift occurred: substantial refractive index changes could now be achieved with essentially arbitrary grating periods. A new door was opened now that practically interesting devices could be made with strong grating reflections at nearly any wavelength.

Applications focused research took off after the United Technologies report. However, the technology was still hampered by a few nagging problems. It seemed that only certain special fibers exhibited interesting levels of photosensitivity; these fibers were neither widely available; nor were they necessarily compatible with standard fiber systems. Additionally, the achievable index change in even the most photosensitive fibers was still quite limited—few fibers could achieve index changes in excess of 10^{-4}. The next major discovery that pushed the thinking beyond these limitations was made at AT&T Bell Labs—Lemaire, Atkins, and Mizrahi showed that by diffusing molecular hydrogen into a standard fiber at room temperature prior to grating writing, an enhancement of the photosensitivity of up to two orders of magnitude was

possible [11–13]! This brought two realizations: first, extremely large index changes comparable to or exceeding the core-cladding index difference in most fibers (10^{-2}) could be achieved, thus enabling ultrawideband and strongly chirped grating devices; and, second, almost any fiber could be rendered photosensitive, including desirable, standard, low-loss transmission fibers.

While from a research standpoint the field was progressing at full speed, skeptics questioned whether fiber gratings could ever be a manufacturable technology. A key discovery addressing this question was the simultaneous demonstration by Hill, et al. [14], and Anderson, et al. [15], of the use of a phase mask to form a master interference pattern that could be repeatably photoreproduced in fiber grating after fiber grating. With its direct analogy to lithography, this approach appeared to offer a highly manufacturable route to large-scale grating production and hence opened the door for serious product development. The phase mask discovery is a good example of how a particular historical event can have a substantial impact, while not itself forming the foundation for all future development. While it is true that currently many fiber gratings are manufactured using the phase mask approach, an appreciable fraction of commercially manufactured gratings are produced using interferometric techniques, despite the original perception that this approach would not lend itself to mass production.

It should be emphasized that while the history of fiber grating technology contains countless significant developments, not all of these will be listed here. Although the field has matured to the point of being a true manufacturable component technology, there is still great potential for many more exciting developments to come.

The remainder of this chapter is organized as follows. In Section II, an overview of the current understanding of fiber photosensitivity and the major approaches to grating fabrication is given. Section III focuses on a fundamental description of the optical properties of fiber gratings. Section IV covers both the achievable properties of fiber gratings from a device standpoint and an overview of applications of this technology in optical communications systems. Finally, some conclusions and future predictions are presented in Section V.

II. Fiber Photosensitivity and Grating Fabrication

A. CONVENTIONAL PHOTOSENSITIVITY IN GERMANOSILICATE GLASS

Conventional UV photosensitivity is defined as real refractive index changes that occur in the waveguide core of a silica fiber that is lightly doped with germanium (<10 mol%) after exposure to light at wavelengths around 242 nm. The UV light is usually from a pulsed KrF excimer (248 nm) or excimer-pumped dye laser (240–250 nm) and has a fluence on the order of

100 mJ/cm^2/pulse. The presence of Ge-doping is critical because these atoms form a defect, the germanium oxygen deficiency center (GODC), which absorbs UV light around 242 nm in a silica glass that would otherwise be transparent. The photoinduced index change is generally present as a uniform index increase throughout the exposed regions of the Ge-doped core.

The research of photosensitivity in germanosilicate fibers has generated a significant body of literature covering the mechanisms and other defects of the process [1,16,17], so only a general explanation will be included here. There are many opinions regarding the causes of UV-induced refractive index changes. It is worth noting that these reactions are very complex and dependent on defects in the atomic glass structure as well as the intensity of the UV exposure. In general, however, conventional photosensitivity is typically due to a combination of electronic excitation of defects in the glass and a change in density by compaction or dilation of the atomic glass structure. The electronic defects change the absorption of the glass at wavelengths below 300 nm, which in turn changes the real refractive index from the UV through the 1.55 μm telecommunications window via the Kramers-Kronig relationship [18–20]. The density changes probably have a contribution that scales with the initial stress present in the fiber [17]. The role of density changes can clearly dominate in highly doped fibers with abnormally high stress; however, results are inconclusive as to whether this role is the dominant mechanism in more conventional fibers.

B. NONSTANDARD TYPES OF PHOTOSENSITIVITY

Several photoinduced index changes have been reported that have characteristics that indicate they are not related to conventional photosensitivity. The different processes fall into two classes: those that induce different refractive index characteristics, and those that use a different UV excitation wavelength.

There are two alternate types of photosensitivity that induce different index characteristics than the conventional Type I photosensitivity described above. One is a strong photoinduced index change referred to as Type II that is achieved with a single very high fluence pulse (>1000 mJ/cm^2/pulse) rather than several hundred pulses for a Type I grating [21]. This type of grating has high insertion loss and polarization dependence because the high-energy pulse damages one side of the core creating a nonuniform index change. Another class of photosensitivity is most easily observed in very highly doped fibers and is referred to as Type IIa, because it was reported to be an earlier stage of Type II growth [22]. The photoinduced index change evolves over many pulses and is negative rather than positive. In general, Type IIa gratings are difficult to reproduce because the magnitude of the effect depends on the residual stress from drawing the fiber, and also because more conventional Type I effects often simultaneously contribute to the total index change.

Photosensitivity in germanosilicates has also been reported at both shorter and longer UV wavelengths. The shorter wavelengths used are typically 193 nm

or less, and are absorbed by the band edge of the doped silica glass [23]. These index changes can be very large but typically have more polarization sensitivity due to asymmetry caused by band edge absorption that is much stronger than in defects that are infrequently distributed throughout the structure. There are also limited reports of large photosensitive index changes at longer wavelengths (300 nm) and higher CW fluences ($\sim$kW/cm^2) [24,25]; however these results apparently require hydrogen loading sensitization as described in the next section.

Conventional Type I photosensitivity is by far the most commonly used fabrication technology in communications applications because the index changes are very reproducible, have minimal polarization dependence, and suffer very low insertion loss.

C. HYDROGEN LOADING

The maximum photoinduced refractive index changes in optical fibers were about 5×10^{-4} before sensitization techniques were discovered. This placed severe limitations on the coupling strength available from a given length grating, or in some cases required gratings to be longer than practical to achieve a particular purpose. Although other sensitization techniques have been reported [26,27], presensitizing virtually any silica fiber by loading it with molecular hydrogen [11] is by far the most widely used approach. Hydrogen (or deuterium) increases the sensitivity by reacting and stabilizing many more defects than previously accessible, with a maximum index change available of $>10^{-2}$ [13]. This is accomplished by UV-induced reactions in the fiber core that were correlated with the total number of germanium atoms in the glass, instead of being limited by only germanium-oxygen-deficiency defect sites as in standard photosensitivity. Because reacted hydrogen is typically incorporated as a loss OH radical with loss at 1.40 μm, deuterium is frequently employed to shift this loss beyond the standard communication window ($\sim$1.9 μm). Hydrogen sensitization has the additional benefit that it relaxes the fiber selection criteria for an application because large index changes can be realized in virtually any fiber. In addition, molecular hydrogen is a temporary sensitization means with no impact on the long-term stability of the fiber after the unreacted molecular hydrogen has diffused from the fiber. Therefore, this approach is used in almost all world-class grating devices requiring low loss, low birefringence, and/or large index changes.

D. PHOTOSENSITIVITY IN OTHER SILICA GLASSES

Photosensitivity has also been observed in many other silica, fluoride, and chalcogenide glass fibers and waveguides. There does not appear to be any widespread commercial use of any of these materials for the near future; however a brief review of the silicate glasses is relevant since these glasses could be readily compatible with the mode field size, reliability, and infrastructure

of current fiber technology. Probably the most interesting alternate glass is phosphorus-doped silicates, which have been used to promote efficient Yb-Er energy transfer in laser and amplifier structures [28,29], and also to fabricate cascaded Raman resonators utilizing the larger frequency phonon P shift [30]. Phosphosilicate glasses can exhibit very large index changes ($>10^{-3}$) when photosensitized with hydrogen and exposed at the near bandgap wavelength of 193 nm [26,31]. Perhaps the most intriguing possibility for the phosphosilicates comes from a natural photosensitive saturation similar to that of photographic film, which makes fabrication of high-contrast refractive index gratings much easier than with germanosilicates [32]. The other silicate classes of potential interest are those that are designed to obtain large photoinduced index changes (without hydrogen loading) by lightly doping with boron [33], tin [34], or cerium [35]. These approaches do show improvement in the photosensitivity; however this has come with negative impact on the reliability and/or transmission loss. Therefore these remain only a subject of research.

E. BIREFRINGENCE

The refractive index change induced in the core of an optical fiber by UV irradiation is generally anisotropic with respect to the polarization of light propagating in the fiber. This anisotropy has been used in such devices as polarization mode converters and rocking filters [36,37]. A similar anisotropy has been observed in the index change induced by the interference of visible light launched through the core of Ge-doped fiber [38–41] and has been attributed to the anisotropy of the polarization of the writing light.

The anisotropic nature of the UV-induced index change has significant implications for many applications of fiber phase gratings. For nonresonant light propagating through such a grating, the anisotropy manifests itself as simple birefringence. For resonant propagation, the birefringence can result in substantial polarization dependence of resonant grating properties, such as the amplitude or phase of the reflectivity. For example, in a fiber grating laser, two orthogonally polarized modes can experience significantly different cavity Qs as a result of the grating birefringence, forcing the laser to operate in a single polarization state [42]. Furthermore, in very long fiber gratings designed for dispersion compensation, the anisotropy can manifest itself as polarization mode dispersion (PMD).

Based on symmetry arguments, there are two tangible causes of the birefringence in the UV-induced refractive index. First, because of the asymmetric geometry associated with UV exposure through the side of the fiber, nonuniform UV absorption across the fiber core might be responsible for the birefringence. An asymmetric transverse profile has been indirectly shown to be a possible cause of birefringence in certain cases [43], but in other cases very uniform index changes have been observed across the cores of optical

fibers loaded with hydrogen in direct index-profiling measurements of sectioned fibers [11]. In the latter cases, it is unlikely that this asymmetry is the main cause of birefringence. The second possible cause is the polarization of the UV light. Recent experiments on ultralow birefringence fibers point to this second cause as the more likely explanation of most observed birefringence effects [44]. A reasonable model is based on treating the relevant defects as oriented dipoles that are preferentially bleached by the particular polarization of the writing laser beam. The resulting glass state has a preferential axis, since either the defects remaining are oriented in a certain direction, or the induced defects were derived from initial defects of a certain orientation, or both of these effects contribute.

To provide a sense for the magnitude of UV-induced birefringence, Fig. 1 shows the growth of the induced birefringence as a function of time in two different fibers during and after uniform exposure to frequency-doubled pulsed dye lasers [45]. The birefringence is measured directly by a probe laser beam that passes through the exposed section of fiber and is analyzed by a commercial polarimeter. Fiber *a* is AT&T Accutether fiber, with an NA of 0.21, and fiber *b* is an Er-doped fiber used in making short fiber lasers, with an NA of 0.27. The light was *s* polarized, or polarized perpendicular to the axis of the optical fiber. The birefringence appears to track the total induced index

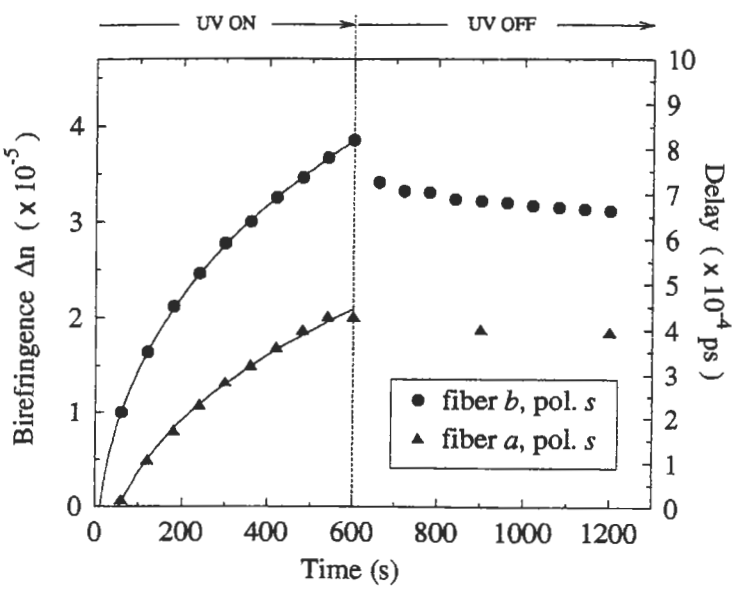


Fig. 1 Measured birefringence and estimated delay as a function of UV exposure time for fibers *a* and *b* irradiated by a single UV beam. The solid curves are theoretical fits to the data using a power law function.

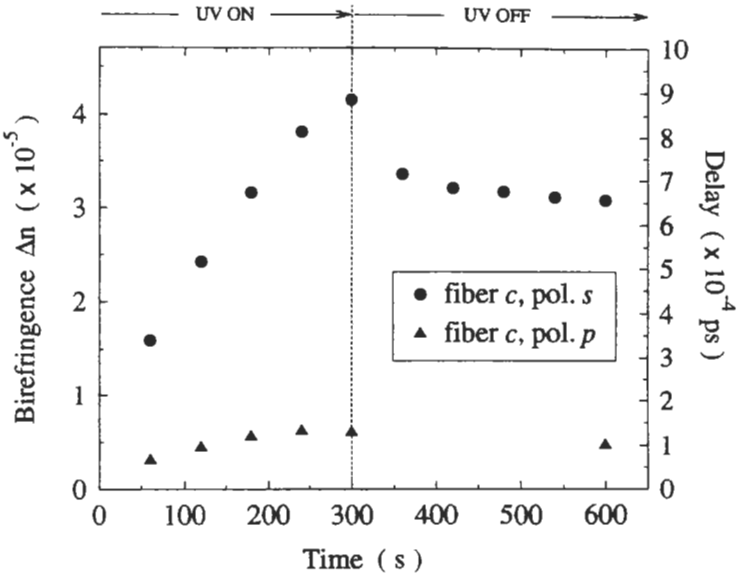


Fig. 2 Measured birefringence and estimated delay as a function of UV exposure time for fiber *c* irradiated by a single UV beam with both *s* and *p* polarizations.

change, with the exception of a slight negative component occurring at the very beginning of the exposure. This effect is presumably tied to the influence of the UV light on internal stresses in the fiber. Using estimated total induced index changes for each of these fibers, fiber *a* exhibits a relative birefringence (normalized to the induced index change) of 5% and while the relative birefringence of fiber *b* is about 8%. Note the relative delay of a polarized signal passing through such a section of fiber is extremely small, indicating that nonresonant PMD is extremely small.

Figure 2 shows a similar measurement of the birefringence in a hydrogen-loaded, standard communications fiber (NA = 0.13 and loaded with about 3 mol% of H_2). Two things can be learned from this result. First, the relative birefringence resulting from UV exposure of a hydrogen-loaded fiber is significantly smaller than that of an unloaded fiber—here the relative birefringence for *s* polarized writing light is only about 1.2%. Second, the birefringence is substantially smaller for exposure to *p* polarized light (here about 0.2%) than it is for *s* polarized light. These results clearly demonstrate the correlation between the polarization of the writing light and UV-induced birefringence.

F. LIFETIME AND ANNEALING

While fiber gratings are often described as "permanent" structures, in fact they are known to decay over time even at standard operating temperatures.

Long-term grating reliability has serious implications for the viability of this technology in high-performance WDM optical communications systems. Additionally, the thermal decay properties are of critical importance to fiber grating sensor applications that sec elevated temperatures.

Many studies have been reported that empirically demonstrate thermal decay. Rather than attempt to cover all such studies, here we highlight a simple method of analysis that makes it possible for the thermal decay properties of a photosensitive fiber to be characterized and extrapolated for accelerated aging stabilization [46]. Experiments that led to this method of analysis show that the decay of the UV-induced index change initially occurs very rapidly, but the rate of decay decreases dramatically with time. Figure 3 shows an example of the decay of a grating written in a high-NA Er-doped fiber for fiber laser applications. It is characterized in terms of the normalized integrated coupling constant (ICC), which is related to the minimum transmission $T_{\min}$ in a lossless grating by $ICC = \operatorname{atanh}\left(\sqrt{1 - T_{\min}}\right)$, and which is directly proportional to the induced index modulation. The decay is also a very strong function of temperature. Even at room temperature, the decay over time scales of interest (years) is typically several percent. Sample decay predictions based on extrapolations of data resulting from measurements like those shown in Fig. 3 are plotted in Fig. 4. Fortunately, the decay process is found to be accelerable: for applications that require very low decay over long times, it is possible to anneal the device, wiping out the portion of the UV-induced index change that

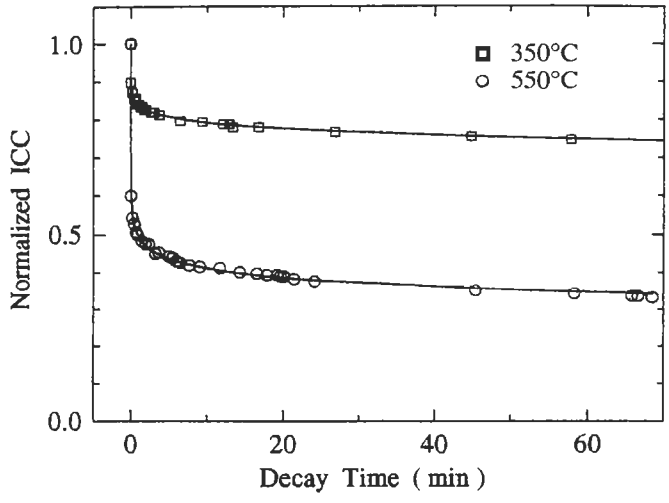


Fig. 3 Measured integrated coupling constant (proportional to the induced index change) normalized to the starting value for two gratings heated to 350 and 550°C as a function of decay time. Solid lines are power law fits to the data.

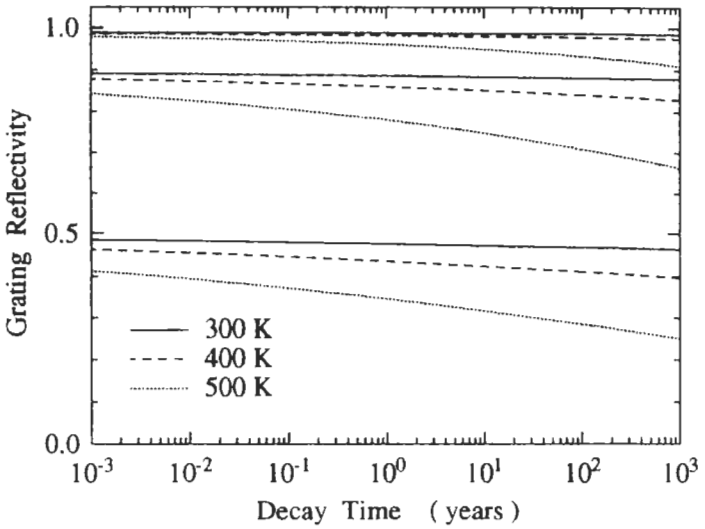


Fig. 4 Predicted decay of the maximum grating reflectivity for gratings with initial reflectivities of 99, 90, and 50%, held at temperatures of 300, 400, and 500°K.

would decay substantially over the lifetime of interest, and leaving only the highly stable portion of the index change.

It has been shown [46] that the typical decay behavior could be fit to a power-law function; however, this fit was merely a mathematical convenience that actually oversimplifies the much more powerful approach suggested by the decay model put forth in that article. In this model, the induced index change is assumed to be proportional to the number of trapped carriers in the glass system. Carriers initially liberated by the UV light are subsequently trapped in a broad energy distribution of trap states, where the detrapping rate is an activated function of the trap depth. The system is schematically illustrated in Fig. 5. In this sort of system, markedly different decay behavior occurs than in a system described by a single activation energy. At any given time t after trapping occurs, the model assumes that all states with energy below some demarcation energy E_d have been depopulated, whereas those states with energy above E_d are still populated. The demarcation energy is given by $E_d = k_B T \ln(\nu_0 t)$, where k_B is Boltzmann's constant, T is the temperature, and ν_0 is an empirically determined constant with units of frequency.

The characterization procedure is straightforward. A series of measurements of the decay of the normalized index change vs. time is made at different temperatures. All of these curves are then plotted as a function of E_d on a single graph, and ν_0 is adjusted to make the curves overlap. A curve fit to the data then yields a master curve that predicts the decay for any possible combination of time and temperature that corresponds to a particular value of E_d on the

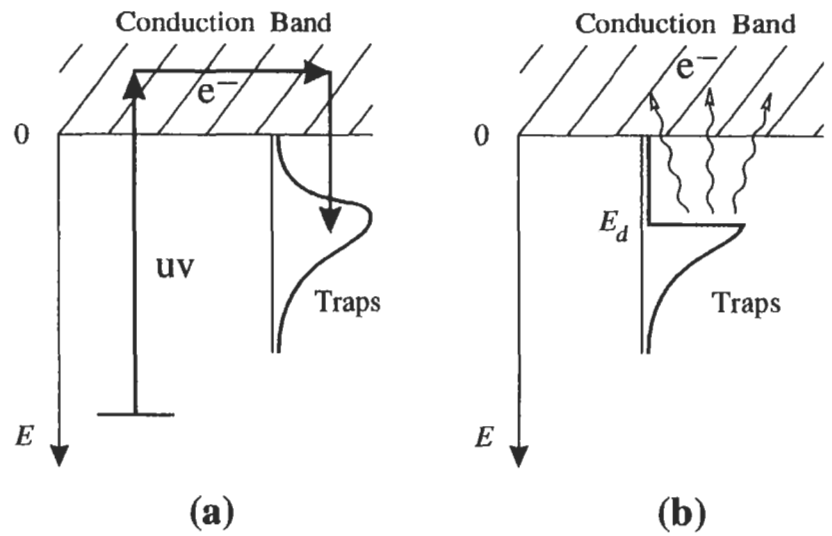


Fig. 5 Diagram of the proposed model for the decay of the UV-induced index change in which (a) electrons excited by the UV irradiation are trapped in a continuous distribution of traps, and (b) thermal depopulation of the traps at a given time and temperature approximately corresponds to shallower traps ($E < E_d$) being emptied and deeper traps ($E > E_d$) remaining full.

curve. The results are most reliable for values of E_d that lie within the range of measured data, but the range can even be extrapolated beyond this measured range. An example of the application of this procedure is shown in Fig. 6.

Once the characterization is complete, the master curve contains all of the information needed to predict the decay dynamics at any temperature and to compute accelerated aging conditions. The decay always follows the master curve from left to right, but the rate at which it progresses along the curve depends on the temperature and the location on the curve. Since for a given temperature the decay occurs much more slowly the further to the right the operating point is located, the object of accelerated aging is to anneal the grating at a very high temperature, moving the operating point as far right as possible without giving up too much induced index change, such that the subsequent decay at operating temperature occurs much more slowly than it would have without the anneal.

Whereas the above procedure has been demonstrated to work extremely well in most fibers under standard photosensitivity conditions, its applicability to fiber gratings written in hydrogen-loaded fibers is more complicated [47]. The main reason is that there appears to be more than one significant trapping mechanism in this case, which in the context of the model leads to multiple values of ν_0, thus making it difficult to apply the simple empirical procedure for determining ν_0. Despite these problems, with certain restrictions the method

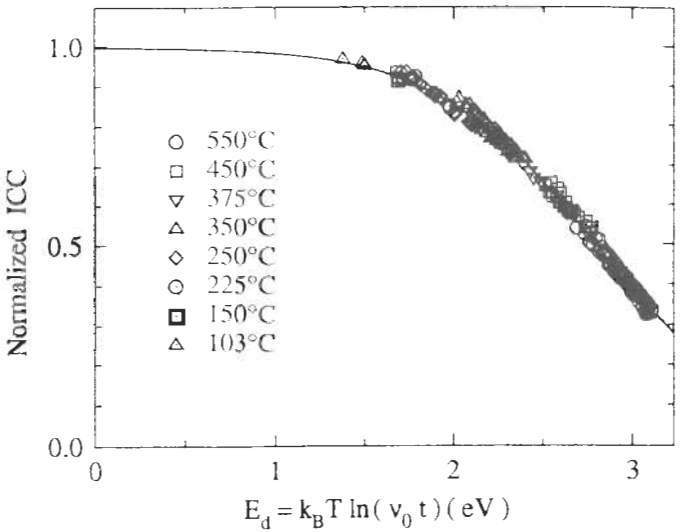


Fig. 6 Plot of the normalized Integrated Coupling Constant as a function of the demarcation energy E_d. The solid line is a curve fit.

is still applicable in the case of hydrogen-loaded fibers. Generally, the induced index change decays more rapidly in gratings written in hydrogen-loaded fibers [47,48], implying that a larger (proportional) reduction in the index change is required in the accelerated aging stabilization process.

G. GRATING WRITING

In this section, a brief overview of the actual apparatus used to write a fiber grating is provided. There are two main parts to the apparatus: the source of UV light and the means of producing periodic or quasi-periodic intensity fringes. The two main classes of the latter include interferometric approaches and phase mask approaches. Based on these means, a number of techniques have been developed to create more sophisticated grating profiles, including a variation of both the amplitude and the phase (apodization and chirp, respectively) of the grating reflection.

G.1. UV Sources

Almost any intense ultraviolet light source at the appropriate wavelength is capable of inducing a change in the refractive index of a nominally photosensitive fiber. However, some light sources are better suited than others to particular writing approaches. There are two main classes of laser sources: pulsed and CW.

The most straightforward source of intense, pulsed UV radiation is the excimer laser. Gratings can be written efficiently at both 248 nm (KrF) and 193 nm (ArF) excimer wavelengths. The main drawback of excimer lasers is their relatively short temporal coherence length and minimal spatial coherence width (both typically tens to 100 μm). This limitation makes it impractical to use amplitude or wavefront division interferometry to produce an interference pattern. But such lasers can be used directly with a phase mask. The coherence properties of excimer lasers can be dramatically improved using intracavity-filter line narrowing (temporal coherence improvement), unstable resonators (spatial coherence improvement), and a Master-Oscillator Power-Amplifier (MOPA) configuration (both temporal and spatial coherence improvement). These improvements come at the expense of intensity and often render the laser as complex as other multilaser combination solutions.

A very popular pulsed source is the excimer- or Nd:YAG-pumped dye laser, which is typically made to lase in the visible (~484 nm) and then frequency doubled to UV wavelengths (~242 nm). This combination provides a tunable, high-coherence solution, albeit with substantially lower power than that available directly from the excimer laser. But the high coherence enables the beam to be tightly focused to achieve high peak intensities.

Another pulsed source that has been demonstrated by many teams is the frequency-quadrupled Nd:YAG (266 nm) or Nd:YLF (262 nm) laser. These sources have extremely high power and are substantially more coherent than an excimer laser, but they are not tunable and happen to have wavelengths substantially longer than the ideal fiber photosensitivity wavelength of about 242 nm.

A number of other types of pulsed lasers have been described or demonstrated, including frequency-tripled pulsed Ti:sapphire lasers and frequency-tripled alexandrite lasers, frequency-doubled optical parametric oscillators (OPO), and frequency-doubled copper-vapor lasers. Many of these are tunable; but due to expense, complexity, availability, or simply comparatively inferior performance, these sources have not seen as widespread use as the sources listed above. Generally, the pulsed sources produce pulse repetition rates of single-shot to ~100 Hz, pulse widths of tens of ns, fluence per pulse of about 0.1 to 1 J/cm^2, and total exposure at the fiber of 1 to 10,000 pulses or J/cm^2.

The other main class of UV sources for fiber grating fabrication is the CW laser. The leading technology here is the frequency-doubled argon-ion laser (244 nm). Because the laser operates CW and hence does not achieve high peak powers, the doubling must either be done directly inside the laser cavity or in an external cavity. Typical powers are 50 to 500 mW, and total exposures are about 10 to 100 W/cm^2 for durations of minutes. It should also be noted that certain fibers and/or conditions allow writing to occur efficiently at around 334 nm using CW light. These wavelengths are also achieved with argon-ion lasers.

G.2. The Interferometer Method

The simplest way to achieve a periodic intensity pattern for writing a grating is to split a laser beam into two paths and then recombine the beams in the vicinity of the fiber. If the half-angle separating the two beams is α and the wavelength of the writing laser is λ_{UV}, then the resulting fringe (and hence grating) period is $\Lambda = \lambda_{UV}/(2\sin\alpha)$. Splitting can be accomplished using either amplitude-division interferometry, in which a beam splitter reflects some fraction of the beam power into one path while transmitting the rest of the power into the other path, or wavefront-division interferometry, in which the beam is literally separated by directing different transverse portions into different paths. While wavefront-division interferometry is sometimes advantageous because it can often be accomplished with an extremely simple, stable apparatus, it requires high coherence. Amplitude-division interferometry can in principle be done with an incoherent beam using a "white-light" interferometer arrangement, although practical tolerances dictate that some degree of coherence is necessary. As a result, this technique is often preferred over wavefront-division interferometry.

Figure 7 shows diagrams of four types of interferometers. The first two are amplitude-division devices, while the latter two are wavefront-division devices. The first is a nonwhite-light interferometer, since the orientation of one beam at the fiber is opposite to the orientation of the other. The second, with an even-integer difference between the number of reflections in each arm and exactly equal path lengths, is a true white-light interferometer. In such an arrangement, the incident beam may be scanned perpendicular to the optical axis, causing the intersection point on the fiber to scan while the fringe pattern remains fixed in space. The third and fourth interferometers illustrate the Lloyd's mirror (or prism) and symmetric prism interferometers. They are extremely simple and stable but require high spatial coherence.

G.3. The Phase Mask Method

An important development in grating writing was the demonstration that a master grating could be used to reproduce an unlimited number of nearly identical fiber gratings [14,15]. While numerous types of master gratings might be considered, a simple practical choice is a phase mask, which consists of a transparent glass plate with an etched, surface-relief grating pattern. The goal of the mask design is to channel all of the incident UV light into only two diffracted orders (see Fig. 8). One simple way to achieve this condition at least approximately is to design the mask for normal incidence with a square-tooth grating having a 50% duty cycle and a depth that corresponds to a π phase shift between light transmitted through the peaks and light transmitted through the valleys. In the scalar diffraction theory approximation, the diffraction efficiency into all even diffracted orders (including the zero order, or the directly transmitted beam) is zero, and the diffraction efficiency into each of the first

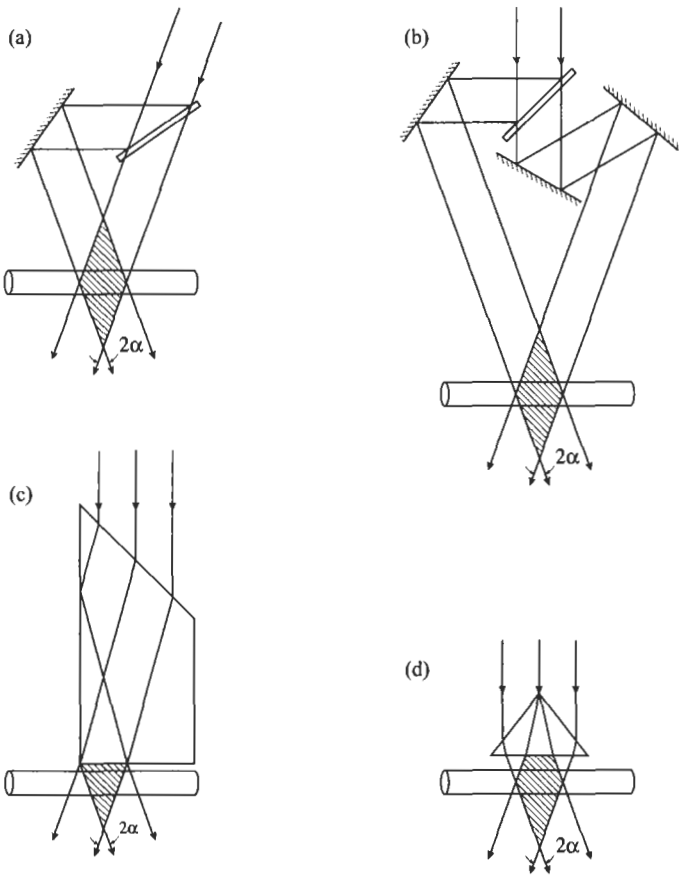


Fig. 7 Diagram of four representative interferometers showing: (a) a simple Mach-Zehnder interferometer; (b) a "white-light" interferometer; (c) a Lloyd's mirror interferometer; and (d) a symmetric prism interferometer.

diffracted orders is a high 40.5%. Because the two main beams produced by the mask are nonadjacent orders, the resulting interference pattern has a period that is half the period of the phase mask. This fortunate result simplifies the production of phase masks, since the smallest feature size for a typical mask is still over 0.5 μm.

Phase masks may be formed either holographically or using electron-beam lithography. An advantage of the latter method is that complicated patterns can be generated on the computer and directly transcribed into the final mask. A significant disadvantage is that the electron beam can only address a field of a few hundred microns in width, and hence one must stitch together many fields to make larger masks. Stitching errors are very difficult to control and lead to phase errors in the resulting fiber grating. In contrast, holographic

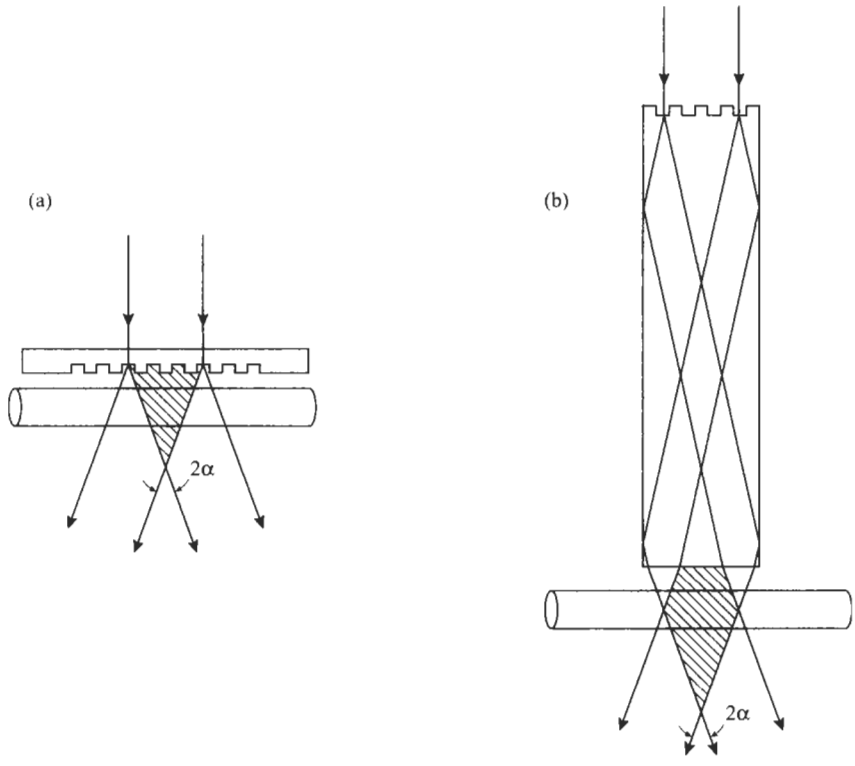


Fig. 8 Diagram of two implementations of a phase mask for writing fiber gratings. In (a) the mask is used directly, whereas in (b) it functions as a beamsplitter.

gratings of appreciable lengths (10 cm or more) can be made with extremely high optical quality, although the ability to tailor the properties along the length is quite limited.

An important property of phase masks is that they drastically reduce the temporal coherence requirements of the source. Even standard, poor-coherence excimer lasers can be used to directly write fiber gratings using a phase mask. However, spatial coherence is a concern, and the fiber must be placed as close to the mask as possible for low-coherence sources. A significant disadvantage of phase masks is that it is very difficult to tune the resulting period of the fiber grating, since even changing the UV writing wavelength does not change the fiber grating period. As a result, most phase-mask writing setups must rely on putting precise, controllable strain on the fiber so that when released, the fiber and grating relax to the appropriate length to yield a particular wavelength. It is interesting to note, however, that a phase mask can be used as an effective amplitude-division beam-splitter, as shown in Fig. 8 [20]. Like the phase mask itself, this white-light interferometer arrangement is still

insensitive to changes in the source wavelength, but unlike the phase mask, unwanted diffracted orders can be removed yielding a superior fringe pattern, and the grating period can be tuned if tilting mirrors are used in place of the prism block.

G.4. Creating Apodization and Chirp

Some applications of fiber gratings require very long and arbitrarily profiled gratings to be produced. The goal is to vary the local amplitude (apodization) and phase (chirp) of the reflection along the length of the grating. A number of approaches have been taken to achieve these sorts of structures. The most straightforward are those based on scanning the laser beam(s) along the fiber, where the beam spot size is much shorter than the resulting grating. If the fiber remains fixed in a white-light interferometer, or does not move relative to a phase mask, only the intensity is scanned—the fringe positions remain fixed. In this case, the scan speed or the laser intensity can be varied over the length of the scan and hence create apodization. However, the average (DC) index change will be apodized along with the index modulation (AC), yielding undesirable spectral features. To avoid these, gratings with constant DC index change are desired, and these can be achieved by a second scan with a single beam that does not produce interference fringes. Note that with this technique the length of the phase mask or the full aperture of the interferometer limits the full grating length. Furthermore, it does not enable chirp to be modified.

In order to arbitrarily vary the apodization and the chirp, a technique was introduced by Asseh et al. that uses an interferometer with a pulsed dye laser [49]. The fiber is fixed to an interferometrically controlled air-bearing stage that has a positioning accuracy of less than a nanometer and a travel of 50 cm. The concept is to time the laser pulses with respect to the position of the continuously moving fiber such that if the pulses arrive at just the right time, the resulting fringes exactly line up with the already written grating lines in the fiber. In this way a very long, unchirped grating can be written. However, if the pulse timing causes the fringes to be precisely out of phase with the already written grating, no grating grows. An intermediate timing causes a partial grating to grow, and hence enables apodization. Additionally, by slowly varying the phase mismatch in one direction as the beam is scanned along the fiber, an effective chirp can be built into the grating. The chirping mechanism is only approximate, and the magnitude of the achievable chirp is limited by the beam size—the smaller the beam, the larger the chirp. However, if the beam is too small, the grating takes an extremely long time to write.

A number of other approaches have been demonstrated for writing long, continuous gratings with complex profiles, although most of these are analogous to the method described by Asseh. For example, Loh et al. introduced a method based on scanning the laser beam across a uniform phase mask with a fiber behind it [50]. As the beam is scanned, the fiber or the phase mask is

scanned at a much slower velocity using a piezoelectric transducer. The greater the velocity, the greater the chirp. Ibsen et al. demonstrated a similar method [51], except instead of moving the fiber or the phase mask, the CW laser beam is modulated by an acoustooptic modulator, effectively creating pulses that are precisely timed with respect to the scanning just as in [49]. Brennan et al. showed that, at least in principle, the limitation on the grating length imposed by the interferometrically controlled stage could be removed by replacing this stage with an extremely precise fiber spool mechanically coupled to a flywheel to keep the velocity constant [52]. As in the above methods, a modulated CW laser source is used to expose the fiber to precisely timed pulsed fringe patterns to produce both apodization and chirp.

III. Optics of Fiber Gratings

In this section we focus on the optical properties of fiber phase gratings, in order to provide both intuition for understanding and tools for designing fiber gratings and related devices. The discussion below is limited to the linear optical properties of fiber gratings. Although a number of interesting nonlinear optical properties have been reported recently (see [53] for an excellent review), these are outside the scope of this section. However, the description of linear properties of fiber gratings below is fundamental for understanding these nonlinear properties. Much of the theoretical discussion below follows closely the treatment in [54].

A. EFFECTIVE INDEX MODULATION

Fiber phase gratings are produced by exposing an optical fiber to a spatially varying pattern of ultraviolet intensity. Here we assume for simplicity that what results is a perturbation to the effective refractive index n_{eff} of the guided mode(s) of interest, described by

$$\delta n_{eff}(z) = \overline{\delta n_{eff}}(z)\left(1 + \nu \cos\left[\frac{2\pi}{\Lambda} z + \phi(z)\right]\right) \tag{1}$$

where $\overline{\delta n_{eff}}$ is the DC index change spatially averaged over a grating period, ν is the fringe visibility of the index change, Λ is the nominal period, and $\phi(z)$ describes grating chirp. If the fiber has a step-index profile, and an induced index change $\delta n_{co}(z)$ is created uniformly across the core, then $\delta n_{eff} \cong \Gamma \delta n_{co}$, where Γ is the core power confinement factor for the mode of interest. For example, Fig. 9 shows the confinement factor Γ and the effective index parameter b for the LP_{01} mode. For $LP_{l\mu}$ modes, b is a solution to the dispersion

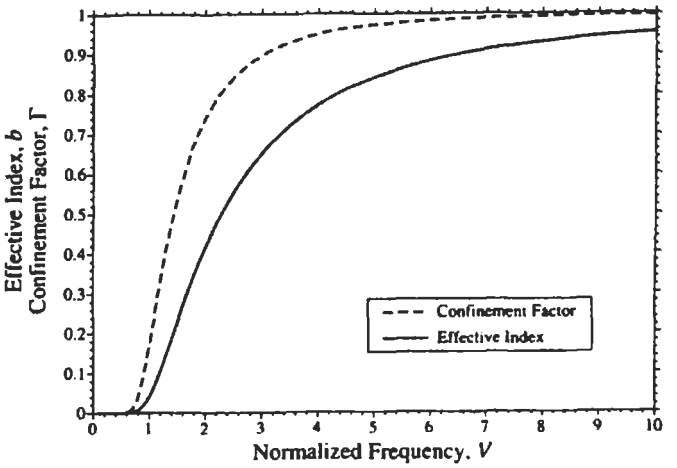


Fig. 9 Effective index parameter b and core confinement factor Γ vs. normalized frequency V for the LP_{01} mode of a step index fiber.

relation [55]

$$V\sqrt{1-b}\frac{J_{l-1}(V\sqrt{1-b})}{J_l(V\sqrt{1-b})} = -V\sqrt{b}\frac{K_{l-1}(V\sqrt{b})}{K_l(V\sqrt{b})} \tag{2}$$

where l is the azimuthal order of the mode and $V = (2\pi/\lambda)a\sqrt{n_{co}^2 - n_{cl}^2}$ is the normalized frequency, with a the core radius, n_{co} the core index, and n_{cl} the cladding index. The effective index is related to b through $b = (n_{eff}^2 - n_{cl}^2)/(n_{co}^2 - n_{cl}^2)$. Once V and b are known, the confinement factor can be determined from

$$\Gamma = \frac{b^2}{V^2}\left(1 - \frac{J_l^2\left(V\sqrt{1-b}\right)}{J_{l+1}\left(V\sqrt{1-b}\right)J_{l-1}\left(V\sqrt{1-b}\right)}\right) \tag{3}$$

The optical properties of a fiber grating are essentially determined by the variation of the induced index change δn_{eff} along the fiber axis z. Figure 10 illustrates some common variations that are discussed in this paper. The terminology used to describe these is given in the figure caption. For illustrative purposes the size of the grating period relative to the grating length has been greatly exaggerated.

B. RAY PICTURE OF FIBER GRATING DIFFRACTION

Initially we focus on the simple case of coupling between two fiber modes by a uniform grating. Before we develop the quantitative analysis using coupled-mode theory, it is helpful to consider a qualitative picture of the basic interactions of interest. A fiber grating is simply an optical diffraction grating,

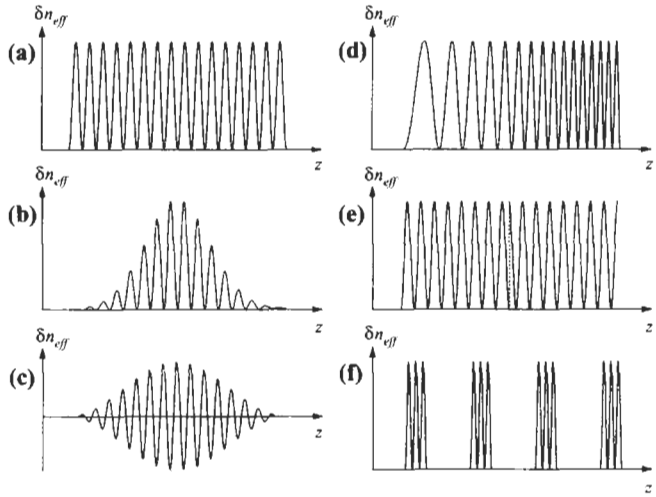


Fig. 10 Common types of fiber gratings as classified by variation of the induced index change along the fiber axis, including: (a) uniform with positive-only index change; (b) Gaussian-apodized; (c) raised-cosine-apodized with zero-DC index change; (d) chirped; (e) discrete phase shift (of π); and (f) superstructure.

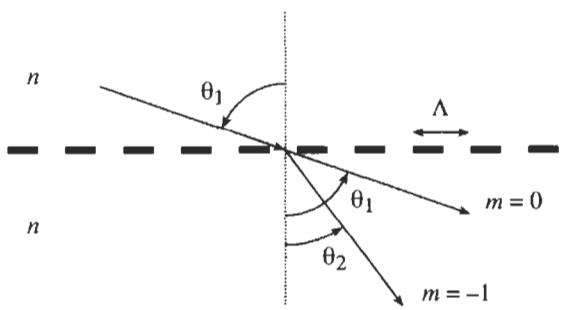


Fig. 11 The diffraction of a light wave by a grating.

and thus its effect upon a light wave incident on the grating at an angle θ_1 can be described by the familiar grating equation [56]

$$n \sin \theta_2 = n \sin \theta_1 + m\frac{\lambda}{\Lambda} \tag{4}$$

where θ_2 is the angle of the diffracted wave and the integer m determines the diffraction order (see Fig. 11). This equation predicts only the directions θ_2 into which constructive interference occurs, but it is nevertheless capable of determining the wavelength at which a fiber grating most efficiently couples light between two modes.

Fiber gratings can be broadly classified into two types: *Bragg gratings* (also called *reflection* and *short-period* gratings), in which coupling occurs between modes traveling in opposite directions; and *transmission gratings* (also called *long-period* gratings) in which the coupling is between modes traveling in the same direction. Figure 12(a) illustrates reflection by a Bragg grating of a mode with a bounce angle of θ_1 into the same mode traveling in the opposite direction with a bounce angle of $\theta_2 = -\theta_1$. Since the mode propagation constant β is

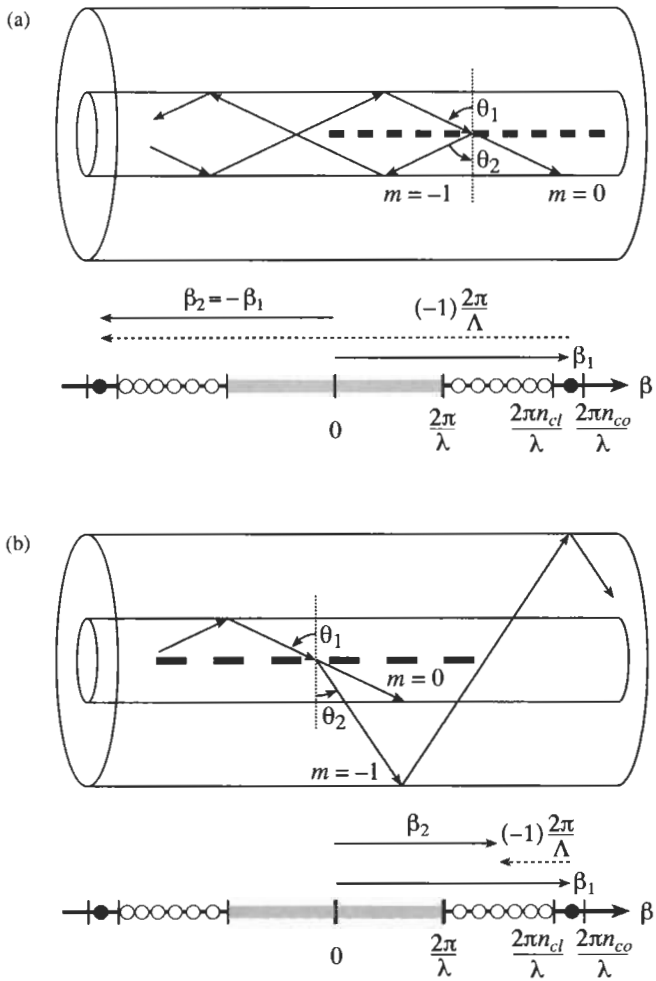


Fig. 12 Ray-optic illustration of (a) core-mode Bragg reflection by a fiber Bragg grating and (b) cladding-mode coupling by a fiber transmission grating. The β axes below each diagram demonstrate the grating condition in Eq. 5 for $m = -1$.

simply $\beta = (2\pi/\lambda)\, n_{eff}$ where $n_{eff} = n_{co} \sin\theta$, we may rewrite Eq. 4 for guided modes as

$$\beta_2 = \beta_1 + m\frac{2\pi}{\Lambda} \tag{5}$$

For first-order diffraction, which usually dominates in a fiber grating, $m = -1$. This condition is illustrated on the β axis shown below the fiber. The solid circles represent bound core modes ($n_{cl} < n_{eff} < n_{co}$), the open circles represent cladding modes ($1 < n_{eff} < n_{cl}$), and the hatched regions represent the continuum of radiation modes. Negative β values describe modes that propagate in the $-z$ direction. By using Eq. 5 and recognizing $\beta_2 < 0$, we find that the resonant wavelength for reflection of a mode of index $n_{eff,1}$ into a mode of index $n_{eff,2}$ is

$$\lambda = (n_{eff,1} + n_{eff,2})\,\Lambda \tag{6}$$

If the two modes are identical, we get the familiar result for Bragg reflection: $\lambda = 2n_{eff}\Lambda$.

Diffraction by a transmission grating of a mode with a bounce angle of θ_1 into a copropagating mode with a bounce angle of θ_2 is illustrated in Fig. 12(b). In this illustration the first mode is a core mode while the second is a cladding mode. Since here $\beta_2 > 0$, Eq. 5 predicts the resonant wavelength for a transmission grating as

$$\lambda = (n_{eff,1} - n_{eff,2})\,\Lambda \tag{7}$$

For copropagating coupling at a given wavelength, evidently a much longer grating period Λ is required than for counterpropagating coupling.

C. *COUPLED-MODE THEORY*

Coupled-mode theory is a good tool for obtaining quantitative information about the diffraction efficiency and spectral dependence of fiber gratings. While other techniques are available, here we consider only coupled-mode theory, since it is straightforward, intuitive, and accurately models the optical properties of most fiber gratings of interest. We do not provide a derivation of coupled-mode theory, as this is detailed in numerous articles and texts [57,58]. Our notation follows most closely that of Kogelnik [58]. In the ideal-mode approximation to coupled-mode theory, we assume that the transverse component of the electric field can be written as a superposition of the ideal modes labeled j (i.e., the modes in an ideal waveguide with no grating perturbation), such that

$$E_t^{\mu}(x,y,z,t) = \sum_j [A_j(z)\exp(i\beta_j z) + B_j(z)\exp(-i\beta_j z)]\mathbf{e}_{jt}(x,y)\exp(-i\omega t) \tag{8}$$

where $A_j(z)$ and $B_j(z)$ are slowly varying amplitudes of the jth mode traveling in the $+z$ and $-z$ directions, respectively. The transverse mode fields $e_{jt}(x,y)$ might describe the bound-core or radiation LP modes as given in [55], or they might describe cladding modes. While the modes are orthogonal in an ideal waveguide and hence do not exchange energy, the presence of a dielectric perturbation causes the modes to be coupled such that the amplitudes A_j and B_j of the jth mode evolve along the z axis according to

$$\frac{dA_j}{dz} = i\sum_k A_k\left(K_{kj}^t + K_{kj}^z\right)\exp\left[i\left(\beta_k - \beta_j\right)z\right] + i\sum_k B_k\left(K_{kj}^t - K_{kj}^z\right)\exp\left[-i\left(\beta_k + \beta_j\right)z\right] \tag{9}$$

$$\frac{dB_j}{dz} = -i\sum_k A_k\left(K_{kj}^t - K_{kj}^z\right)\exp\left[i\left(\beta_k + \beta_j\right)z\right] - i\sum_k B_k\left(K_{kj}^t + K_{kj}^z\right)\exp\left[-i\left(\beta_k - \beta_j\right)z\right] \tag{10}$$

In Eqs 9 and 10, $K_{kj}^t(z)$ is the transverse coupling coefficient between modes j and k given by

$$K_{kj}^t(z) = \frac{\omega}{4}\iint_{\infty} dxdy\Delta\varepsilon(x,y,z)\, e_{kt}(x,y)\cdot e_{jt}(x,y) \tag{11}$$

where $\Delta\varepsilon$ is the perturbation to the permittivity, approximately given by $\Delta\varepsilon \cong 2n\delta n$ when $\delta n \ll n$. The longitudinal coefficient $K_{kj}^z(z)$ is analogous to $K_{kj}^t(z)$, but generally $K_{kj}^z(z) \ll K_{kj}^t(z)$ for fiber modes, and thus this coefficient is usually neglected.

In most fiber gratings the induced index change $\delta n(x,y,z)$ is approximately uniform across the core and nonexistent outside the core. We can thus describe the core index by an expression similar to Eq. 1, but with $\overline{\delta n_{eff}}(z)$ replaced by $\overline{\delta n_{co}}(z)$. If we define two new coefficients

$$\sigma_{kj}(z) = \frac{\omega n_{co}}{2}\overline{\delta n_{co}}(z)\iint_{core} dxdy e_{kt}(x,y)\cdot e_{jt}(x,y) \tag{12}$$

and

$$\kappa_{kj}(z) = \frac{\nu}{2}\sigma_{kj}(z) \tag{13}$$

where σ is a DC (period-averaged) coupling coefficient and κ is an AC coupling coefficient, then the general coupling coefficient can be written

$$K_{kj}^t(z) = \sigma_{kj}(z) + 2\kappa_{kj}(z)\cos\left[\frac{2\pi}{\Lambda}z + \phi(z)\right] \tag{14}$$

Eqs 9–14 are the coupled-mode equations that we use to describe fiber-grating spectra below.

D. BRAGG (REFLECTION) GRATINGS

Near the wavelength for which reflection of a mode of amplitude $A(z)$ into an identical counterpropagating mode of amplitude $B(z)$ is the dominant interaction in a Bragg grating, Eqs 9 and 10 may be simplified by retaining only terms that involve the amplitudes of the particular mode, and then making the synchronous approximation [58] which amounts to neglecting terms on the right-hand sides of the differential equations that contain rapidly oscillating z dependence, since these contribute little to the growth and decay of the amplitudes. The resulting equations can be written

$$\frac{dR}{dz} = i\hat{\sigma}R(z) + i\kappa S(z) \tag{15}$$

$$\frac{dS}{dz} = -i\hat{\sigma}S(z) - i\kappa^* R(z) \tag{16}$$

where the amplitudes R and S are $R(z) \equiv A(z)\exp(i\delta z - \phi/2)$ and $S(z) \equiv B(z)\exp(-i\delta z + \phi/2)$. In these equations κ is the AC coupling coefficient from (13) and $\hat{\sigma}$ is a general DC self-coupling coefficient defined as

$$\hat{\sigma} \equiv \delta + \sigma - \frac{1}{2}\frac{d\phi}{dz} \tag{17}$$

The detuning δ, which is independent of z, is defined to be

$$\delta \equiv \beta - \frac{\pi}{\Lambda} = \beta - \beta_D = 2\pi n_{eff}\left(\frac{1}{\lambda} - \frac{1}{\lambda_D}\right) \tag{18}$$

where $\lambda_D \equiv 2n_{eff}\Lambda$ is the design wavelength for Bragg scattering by an infinitesimally weak grating $(\delta n_{eff} \to 0)$ with a period Λ. Note when $\delta = 0$ we find $\lambda = 2n_{eff}\Lambda$, the Bragg condition predicted by the qualitative grating picture above. The DC coupling coefficient σ is defined in Eq. 12. Absorption loss in the grating can be described by a complex coefficient σ, where the power loss coefficient is $\alpha = 2\,\mathrm{Im}(\sigma)$. Light not reflected by the grating experiences a transmission loss of $10\log_{10}(e)\alpha$ dB/cm. Finally, the derivative $(1/2)\,d\phi/dz$ describes possible chirp of the grating period, where $\phi(z)$ is defined through Eq. 1 or 14.

For a single-mode Bragg reflection grating we find the simple relations:

$$\sigma = \frac{2\pi}{\lambda}\overline{\delta n_{eff}} \tag{19}$$

$$\kappa = \kappa^* = \frac{\pi}{\lambda}\nu\overline{\delta n_{eff}} \tag{20}$$

If the grating is uniform along z, then $\overline{\delta n_{eff}}$ is a constant and $d\phi/dz = 0$, and thus κ, σ, and $\hat{\sigma}$ are constants. Thus Eqs 15 and 16 are coupled first-order ordinary differential equations with constant coefficients, for which closed-form

solutions can be found when appropriate boundary conditions are specified. The reflectivity of a uniform fiber grating of length L can be found by assuming a forward-going wave incident from $z = -\infty$ (say $R(-L/2) = 1$) and requiring that no backward-going wave exists for $z \geq L/2$ (i.e., $S(L/2) = 0$). The amplitude and power reflection coefficients $\rho = S(-L/2)/R(-L/2)$ and $r = |\rho|^2$, respectively, can then be shown [57,58] to be

$$\rho = \frac{-\kappa \sinh(\sqrt{\kappa^2 - \hat{\sigma}^2}L)}{\hat{\sigma}\sinh(\sqrt{\kappa^2 - \hat{\sigma}^2}L) + i\sqrt{\kappa^2 - \hat{\sigma}^2}\cosh(\sqrt{\kappa^2 - \hat{\sigma}^2}L)} \tag{21}$$

and

$$r = \frac{\sinh^2(\sqrt{\kappa^2 - \hat{\sigma}^2}L)}{\cosh^2(\sqrt{\kappa^2 - \hat{\sigma}^2}L) - \hat{\sigma}^2/\kappa^2} \tag{22}$$

A number of interesting features of fiber Bragg gratings can be seen from these results. Typical examples of the power reflectivity r for uniform gratings with $\kappa L = 2$ and $\kappa L = 8$ are shown in Fig. 13, plotted versus the normalized wavelength

$$\frac{\lambda}{\lambda_{max}} = \frac{1}{(1 + \hat{\sigma}L/\pi N)} \tag{23}$$

where N is the total number of grating periods ($N = L/\Lambda$), here chosen to be $N = 10{,}000$, and λ_{max} is the wavelength at which maximum reflectivity occurs. If N were larger or smaller, the reflection bandwidth would be narrower

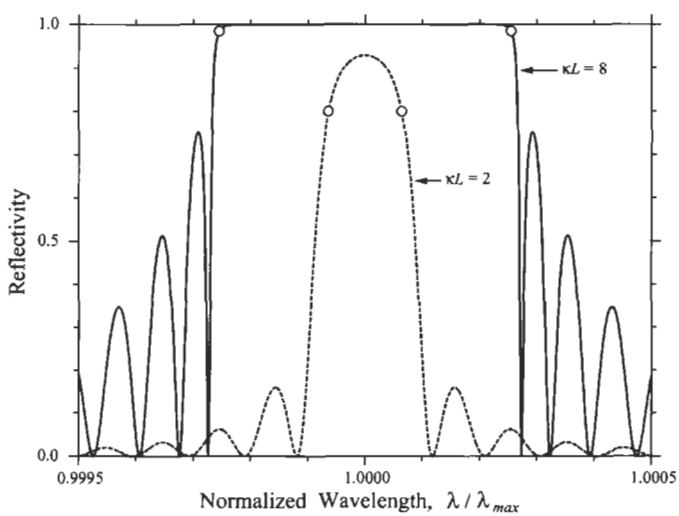


Fig. 13 Reflection spectra vs. normalized wavelength for Bragg reflection in uniform gratings with $\kappa L = 2$ (dashed line) and $\kappa L = 8$ (solid line).

or broader, respectively, for a given value of κL. From Eq. 22, we find the maximum reflectivity r_{max} for a Bragg grating is

$$r_{max} = \tanh^2(\kappa L) \tag{24}$$

which occurs when $\hat{\sigma} = 0$, or at the wavelength

$$\lambda_{max} = \left(1 + \frac{\overline{\delta n_{eff}}}{n_{eff}}\right)\lambda_D \tag{25}$$

The points on these plots denoted by open circles indicate the band edges, or the points at the edges of the band gap defined such that $|\hat{\sigma}| < \kappa$. Inside the band gap, the amplitudes $R(z)$ and $S(z)$ grow and decay exponentially along z; outside the band gap they evolve sinusoidally. The reflectivity at the band edges is

$$r_{band\ edge} = \frac{(\kappa L)^2}{1 + (\kappa L)^2} \tag{26}$$

and the band edges occur at the wavelengths

$$\lambda_{band\ edge} = \lambda_{max} \pm \frac{\nu\overline{\delta n_{eff}}}{2n_{eff}}\lambda_D \tag{27}$$

From Eq. 27 the normalized bandwidth of a Bragg grating as measured at the band edges is simply

$$\frac{\Delta\lambda_{band\ edge}}{\lambda} = \frac{\nu\overline{\delta n_{eff}}}{n_{eff}} \tag{28}$$

where $\nu\overline{\delta n_{eff}}$ is simply the AC part of the induced index change.

A more readily measurable bandwidth for a uniform Bragg grating is that between the first zeroes on either side of the maximum reflectivity. Looking at the excursion of $\hat{\sigma}$ from $\hat{\sigma} = 0$ that causes the numerator in Eq. 22 to go to zero, we find

$$\frac{\Delta\lambda_0}{\lambda} = \frac{\nu\overline{\delta n_{eff}}}{n_{eff}}\sqrt{1 + \left(\frac{\lambda_D}{\nu\overline{\delta n_{eff}}L}\right)^2} \tag{29}$$

In the weak-grating limit, for which $\nu\overline{\delta n_{eff}}$ is very small, we find

$$\frac{\Delta\lambda_0}{\lambda} \to \frac{\lambda_D}{n_{eff}L} = \frac{2}{N}\left(\nu\overline{\delta n_{eff}} \ll \lambda_D/L\right) \tag{30}$$

the bandwidth of weak gratings is said to be length-limited. However, in the strong-grating limit we find

$$\frac{\Delta\lambda_0}{\lambda} \to \frac{\nu\overline{\delta n_{eff}}}{n_{eff}}\left(\nu\overline{\delta n_{eff}} \gg \lambda_D/L\right) \tag{31}$$

In strong gratings the light does not penetrate the full length of the grating, and thus the bandwidth is independent of length and directly proportional to the induced index change. For strong gratings the bandwidth is similar whether measured at the band edges, at the first zeroes, or as the full width at half maximum.

To compare the theory to an experimental measurement, Fig. 14 shows the measured (dots) and calculated (line) reflection from a 1-mm long uniform grating with a design wavelength of 1558 nm and an induced index change of $\overline{\nu\delta n_{eff}} = 8 \times 10^{-4}$, thus yielding $\kappa L = 1.64$.

Recently there has been growing interest in the dispersive properties of fiber Bragg gratings for applications such as dispersion compensation, pulse shaping, and fiber and semiconductor laser components. Although many of these rely on the ability to tailor the dispersion in the nonuniform gratings discussed below in Section IIIF, here we introduce the basis for determining delay and dispersion from the known (complex) reflectivity of a Bragg grating. The group delay and dispersion of the reflected light can be determined from the phase of the amplitude reflection coefficient ρ in Eq. 21. If we denote $\theta_\rho \equiv \text{phase}(\rho)$, then at a local frequency ω_0 we may expand θ_ρ in a Taylor series about ω_0. Since the first derivative $d\theta_\rho/d\omega$ is directly proportional to the frequency ω, this quantity can be identified as a time delay. Thus the delay time τ_ρ for light reflected off of a grating is

$$\tau_\rho = \frac{d\theta_\rho}{d\omega} = \frac{\lambda^2}{2\pi c}\frac{d\theta_\rho}{d\lambda} \tag{32}$$

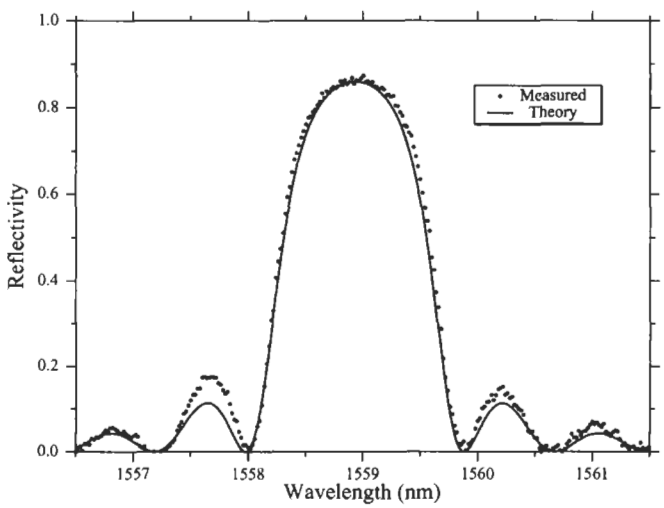


Fig. 14 Measured (dots) and calculated (line) reflection spectra for Bragg reflection in a 1-mm long uniform grating with $\kappa L = 1.64$.

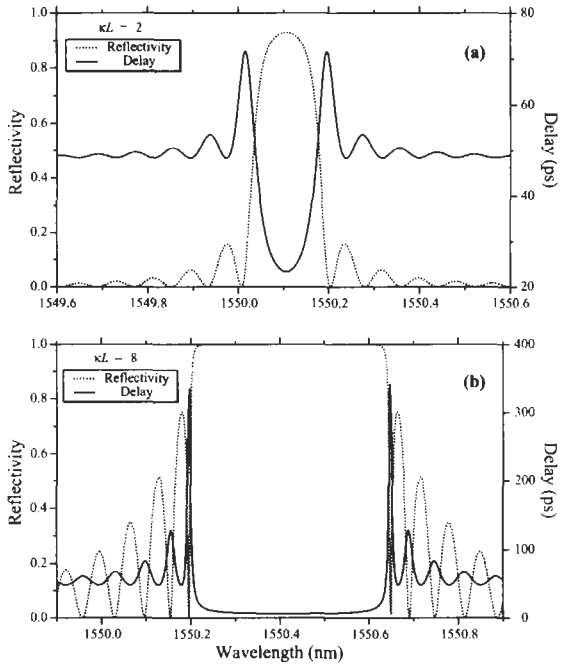


Fig. 15 Calculated reflection spectra (dotted line) and group delay (solid line) for uniform Bragg gratings with (a) $\kappa L = 2$ and (b) $\kappa L = 8$.

τ_ρ is usually given in units of ps. Figure 15 shows the delay τ_ρ calculated for the two example gratings from Fig. 13. Here the grating length is $L = 1\,\text{cm}$, the design wavelength is $\lambda_D = 1550\,\text{nm}$, $n_{eff} = 1.45$, and the fringe visibility is $v = 1$. For the weaker grating in Fig. 15(a), $\overline{\delta n_{eff}} = 1 \times 10^{-4}$, while for the stronger grating in Fig. 15(b), $\overline{\delta n_{eff}} = 4 \times 10^{-4}$. We see that for unchirped uniform gratings both the reflectivity and the delay are symmetric about the wavelength λ_{max}.

Since the dispersion d_ρ (in ps/nm) is the rate of change of delay with wavelength, we find

$$d_\rho = \frac{d\tau_\rho}{d\lambda} = \frac{2\tau_\rho}{\lambda} - \frac{\lambda^2}{2\pi c}\frac{d^2\theta_\rho}{d\lambda^2} = -\frac{2\pi c}{\lambda^2}\frac{d^2\theta_\rho}{d\omega^2} \tag{33}$$

In a uniform grating the dispersion is zero near λ_{max} and becomes appreciable only near the band edges and side lobes of the reflection spectrum, where it tends to vary rapidly with wavelength. Qualitatively, this behavior of the delay and dispersion (along with numerous other characteristics of fiber gratings) can be nicely explained by an effective medium picture developed by Sipe, et al. [59]. For wavelengths outside the bandgap, the boundaries of the uniform

grating (at $z = \pm L/2$) act like abrupt interfaces, thus forming a Fabry–Perot-like cavity. The nulls in the reflection spectrum are analogous to Fabry–Perot resonances—at these frequencies light is trapped inside the cavity for many round trips, thus experiencing enhanced delay.

E. LONG-PERIOD (TRANSMISSION) GRATINGS

Near the wavelength at which a mode 1 of amplitude $A_1(z)$ is strongly coupled to a copropagating mode 2 with amplitude $A_2(z)$, Eqs 9 and 10 may be simplified by retaining only terms that involve the amplitudes of these two modes, and then making the usual synchronous approximation. The resulting equations can be written

$$\frac{dR}{dz} = i\hat{\sigma}R(z) + i\kappa S(z) \tag{34}$$

$$\frac{dS}{dz} = -i\hat{\sigma}S(z) + i\kappa^* R(z) \tag{35}$$

where the new amplitudes R and S are $R(z) \equiv A_1 \exp[-i(\sigma_{11} + \sigma_{22})z/2] \cdot \exp(i\delta z - \phi/2)$ and $S(z) \equiv A_2 \exp[-i(\sigma_{11} + \sigma_{22})z/2] \cdot \exp(-i\delta z + \phi/2)$, and σ_{11} and σ_{22} are DC coupling coefficients defined in Eq. 12. In Eqs 34 and 35 $\kappa = \kappa_{21} = \kappa_{21}^*$ is the AC crosscoupling coefficient from Eq. 13 and $\hat{\sigma}$ is a general DC self-coupling coefficient defined as

$$\hat{\sigma} \equiv \delta + \frac{\sigma_{11} - \sigma_{22}}{2} - \frac{1}{2}\frac{d\phi}{dz} \tag{36}$$

Here the detuning, which is assumed to be constant along z, is

$$\delta \equiv \frac{1}{2}(\beta_1 - \beta_2) - \frac{\pi}{\Lambda} = \pi \Delta n_{eff}\left(\frac{1}{\lambda} - \frac{1}{\lambda_D}\right) \tag{37}$$

where $\lambda_D \equiv \Delta n_{eff}\Lambda$ is the design wavelength for an infinitesimally weak grating. As for Bragg gratings, $\delta = 0$ corresponds to the grating condition predicted by the qualitative picture of grating diffraction, or $\lambda = \Delta n_{eff}\Lambda$.

For a uniform grating $\hat{\sigma}$ and κ are constants. Unlike for the case of Bragg reflection of a single mode, here the coupling coefficient κ generally cannot be written simply as in Eq. 20. For coupling between two different modes in both Bragg and transmission gratings the overlap integrals in Eqs 12 and 13 must be evaluated numerically (for untilted gratings, closed-form expressions exist, but these still require numerical evaluation of Bessel functions). Like the analogous Bragg grating equations, Eqs 34 and 35 are coupled first-order ordinary differential equations with constant coefficients. Thus closed-form solutions can be found when appropriate initial conditions are specified. The transmission can be found by assuming only one mode is incident from $z = -\infty$ (say $R(0) = 1$ and $S(0) = 0$). The power *bar* and *cross* transmission,

$t_= = |R(z)|^2 / |R(0)|^2$ and $t_\times = |S(z)|^2 / |R(0)|^2$, respectively, can be shown to be [10]

$$t_= = \cos^2\left(\sqrt{\kappa^2 + \hat{\sigma}^2} z\right) + \frac{1}{1 + \kappa^2/\hat{\sigma}^2} \sin^2\left(\sqrt{\kappa^2 + \hat{\sigma}^2} z\right) \tag{38}$$

$$t_\times = \frac{1}{1 + \hat{\sigma}^2/\kappa^2} \sin^2\left(\sqrt{\kappa^2 + \hat{\sigma}^2} z\right) \tag{39}$$

Typical examples of the cross transmission $t_\times$ for two uniform gratings of length L with $\kappa L = \pi/2$ (dashed line), $\kappa L = 3\pi/2$ (solid line), and $\sigma_{11} = \sigma_{22} = 0$ are shown in Fig. 16, plotted versus the normalized wavelength defined in Eq. 23. Here the total number of grating periods is $N = 100$. If N were larger or smaller, the bandwidth would be narrower or broader, respectively, for a given value of κL. The maximum cross transmission (which occurs when $\hat{\sigma} = 0$) is

$$t_{\times,max} = \sin^2(\kappa L) \tag{40}$$

and it occurs at the wavelength

$$\lambda_{max} = \frac{1}{1 - (\sigma_{11} - \sigma_{22})\Lambda/2\pi} \lambda_D \tag{41}$$

For coupling between a core mode 1 and a cladding mode 2 (see Fig. 12(b)) with an induced index change in the core only, $\sigma_{11} = \sigma$ from Eq. 19, and we

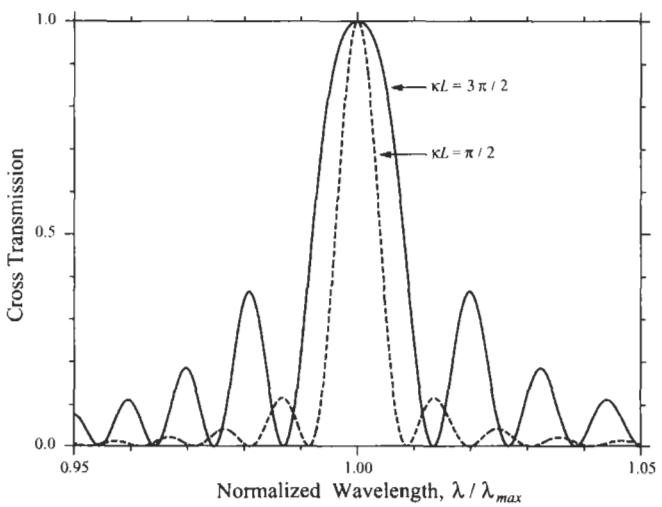


Fig. 16 Calculated cross transmission $t_\times$ through uniform transmission gratings with $\kappa L = \pi/2$ (dashed line) and $\kappa L = 3\pi/2$ (solid line).

find that $\sigma_{22} \ll \sigma_{11}$ (since the core confinement factor for the cladding mode is small). Therefore we can approximate Eq. 41 as

$$\lambda_{max} \cong \left(1 + \frac{\overline{\delta n_{eff}}}{\Delta n_{eff}}\right) \lambda_D \tag{42}$$

where we have assumed that $\overline{\delta n_{eff}}$, the induced change in the core-mode effective index, is much smaller than Δn_{eff}. Comparing Eqs 42 and 25, we see that the wavelength of maximum coupling in a long-period cladding-mode coupling grating shifts (toward longer wavelengths) as the grating is being written $n_{eff}/\Delta n_{eff}$ times more rapidly than the shift that occurs in a Bragg grating.

A useful measure of the bandwidth of a long-period grating is the separation between the first zeroes on either side of the spectral peak. Looking at the excursion of $\hat{\sigma}$ from $\hat{\sigma} = 0$ that causes the numerator in Eq. 39 to go to zero, we find

$$\frac{\Delta\lambda_0}{\lambda} = \frac{2\lambda}{\Delta n_{eff} L}\sqrt{1 - \left(\frac{\kappa L}{\pi}\right)^2} \tag{43}$$

for a grating in which at most one complete exchange of power between the two modes occurs ($\kappa L \leq \pi$). For very weak gratings the normalized bandwidth is simply $\Delta\lambda_0/\lambda = 2/N$, just as for weak Bragg gratings. For strongly overcoupled gratings, where $\kappa L \gg \pi$, the sidelobes become significantly more pronounced and hence a better measure of the bandwidth is the full width at half maximum (FWHM) of the envelope traced by the peaks of the sidelobes. Looking at the first factor in the expression for $t_\times$ in Eq. 39, we find

$$\frac{\Delta\lambda_{env}}{\lambda} = \frac{2\lambda\kappa}{\pi\Delta n_{eff}} \tag{44}$$

Since sidelobes are usually undesirable, most transmission gratings are designed such that $\kappa L \leq \pi/2$, where for the strongest gratings (large cross transmission) $\kappa L \sim \pi/2$.

To compare the theory to an experimental measurement, Fig. 17 shows the measured (dots) and calculated (line) bar transmission $t_=$ for a relatively weak grating that couples the LP_{01} core mode to the lowest-order cladding mode in a standard dispersion-shifted fiber ($\Delta n_{eff} = 0.0042$). The grating is 50 mm long and has a coupling-length product of $\kappa L = 0.39$.

Whereas the bandwidth estimates above are generally accurate for practical Bragg gratings, Eqs 43 and 44 for transmission gratings can be poor estimates if the effective index dispersion is large near the resonant wavelength. The estimates assume Δn_{eff} is independent of wavelength. But if, for example, $\Delta n_{eff} \propto \lambda$ over an appreciable span of wavelengths, then from Eq. 37 the detuning δ could remain small and thus allow strong coupling over the entire span, giving rise to a much broader grating! Dispersion is more of a concern for

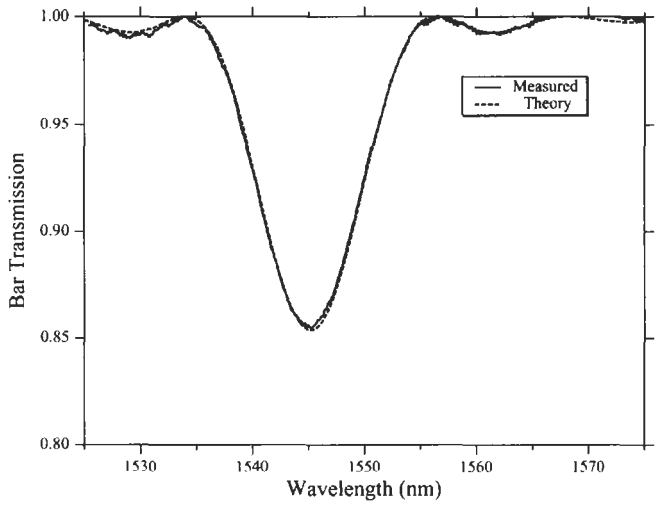


Fig. 17 Measured (dotted line) and calculated (solid line) bar transmission $t_{=}$ through a uniform cladding-mode transmission grating.

transmission gratings mainly because these tend to have broader bandwidths than Bragg gratings even in the absence of dispersion.

While the results in this section apply rigorously only to coupling between two modes by a uniform grating, many of these rules of thumb prove to be excellent approximations even for nonuniform gratings.

F. NONUNIFORM GRATINGS

In this section we investigate the properties of nonuniform gratings in which the coupling occurs predominantly between two modes. We consider approaches to modeling both Bragg and transmission gratings and look at examples of several types of nonuniformity.

Most fiber gratings designed for practical applications are not uniform gratings. Often the main reason for choosing a nonuniform design is to reduce the undesirable sidelobes prevalent in uniform-grating spectra; but there are many other reasons to adjust the optical properties of fiber gratings by tailoring the grating parameters along the fiber axis. It has been known for some time that apodizing the coupling strength of a waveguide grating can produce a reflection spectrum that more closely approximates the often-desired "top-hat" shape [60–62]. Sharp, well-defined filter shapes are rapidly becoming critical characteristics for passive components in dense wavelength-division multiplexed (DWDM) communications systems. Chirping the period of a grating enables the dispersive properties of the scattered light to be tailored [60]. Chirped fiber gratings are useful for dispersion compensation [63], for controlling and shaping reflections in fiber lasers [64], and for creating stable continuous-wave

(CW) and tunable mode-locked external-cavity semiconductor lasers [65,66]. Sometimes it is desirable to create discrete, localized phase shifts in an otherwise periodic grating. Discrete phase shifts can be used to open an extremely narrow transmission resonance in a reflection grating or to tailor the passive filter shape. Perhaps the most well-known application of discrete phase shifts is the use of a "quarter-wave" or π phase shift in the center of a distributed-feedback laser to break the threshold condition degeneracy for the two lowest-order laser modes, thus favoring single-mode lasing [67,68]. Recently interest has grown in gratings with periodic superstructure, in which the coupling strength, DC index change, or grating period are varied periodically with a period much larger than the nominal grating period Λ. Such sampled gratings have been proposed and demonstrated for a number of applications [69], including use as a wavelength reference standard for DWDM systems [70]. Understanding effects of discrete phase shifts and superstructure has become critical recently with the advent of meter-long Bragg gratings for dispersion compensation produced by stitching together exposure regions formed with multiple phase masks [71,72].

We consider two standard approaches for calculating the reflection and transmission spectra that result from two-mode coupling in nonuniform gratings. The first is direct numerical integration of the coupled-mode equations. This approach has several advantages, but it is rarely the fastest method. The second approach is a piecewise-uniform approach, in which the grating is divided into a number of uniform pieces. The closed-form solutions for each uniform piece are combined by multiplying matrixes associated with the pieces. This method is simple to implement, almost always sufficiently accurate, and generally the fastest. Other approaches are also possible, such as treating each grating half-period like a layer in a thin-film stack (Rouard's method) [73]. Like the piecewise-uniform approach, this method amounts to multiplying a string of matrixes; but because the number of matrixes scales with the number of grating periods, this approach can become intractable for fiber gratings that are centimeters long with 10^5 periods or more.

The direct-integration approach to solving the coupled-mode equations is straightforward. The equations have been given above: Eqs 15 and 16 apply to Bragg gratings and Eqs 34 and 35 are used for transmission gratings. Likewise, the boundary conditions have been described above. For a Bragg grating of length L, one generally takes $R(L/2) = 1$ and $S(L/2) = 0$, and then integrates backwards from $z = L/2$ to $z = -L/2$, thus obtaining $R(-L/2)$ and $S(-L/2)$. Since the transmission grating problem is an initial-value problem, the numerical integration is done in the forward direction from $z = -L/2$ to $z = L/2$, starting with the initial conditions $R(-L/2) = 1$ and $S(-L/2) = 0$, for example. Typically, adaptive-stepsize Runge-Kutta numerical integration works well.

For modeling apodized gratings by direct numerical integration, we simply use the z-dependent quantities $\sigma_{kj}(z)$ and $\kappa_{kj}(z)$ in the coupled-mode equations,

which give rise to a $\hat{\sigma}(z)$ that also depends on z. For some apodized grating shapes, we need to truncate the apodization function. For example, fiber gratings are frequently written by a Gaussian laser beam, and thus have an approximately Gaussian profile of the form

$$\overline{\delta n_{eff}}(z) = \overline{\delta n_{eff}} \exp\left(-4 \ln 2 z^2 / fwhm^2\right) \tag{45}$$

where $\overline{\delta n_{eff}}$ is the peak value of the DC effective index change and *fwhm* is the full width at half maximum of the grating profile. Typically, Eq. 45 is truncated at several times the *fwhm*; i.e., we choose $L \sim 3\,fwhm$. Another common profile is the "raised-cosine" shape

$$\overline{\delta n_{eff}}(z) = \overline{\delta n_{eff}} \tfrac{1}{2}[1 + \cos(\pi z / fwhm)] \tag{46}$$

This profile is truncated at $L = fwhm$, where it is identically zero. Many other apodized profiles are of interest as well, such as the "flattop raised-cosine."

Chirped gratings can be modeled using the direct integration technique by simply including a nonzero z-dependent phase term $(1/2)d\phi/dz$ in the self-coupling coefficient $\hat{\sigma}$ defined in Eqs 17 and 36. In terms of more readily understandable parameters, the phase term for a linear chirp is

$$\frac{1}{2}\frac{d\phi}{dz} = -\frac{4\pi n_{eff} z}{\lambda_D^2}\frac{d\lambda_D}{dz} \tag{47}$$

where the "chirp" $d\lambda_D/dz$ is a measure of the rate of change of the design wavelength with position in the grating, usually given in units of nm/cm. Linear chirp can also be specified in terms of a dimensionless "chirp parameter" F [62], given by

$$F = \frac{fwhm^2}{z^2}\phi(z) = -4\pi n_{eff}\frac{fwhm^2}{\lambda_D^2}\frac{d\lambda_D}{dz} \tag{48}$$

F is a measure of the fractional change in the grating period over the whole length of the grating. It is important to recognize that because chirp is simply incorporated into the coupled-mode equations as a z-dependent term in the self-coupling coefficient $\hat{\sigma}$, its effect is identical to that of a DC index variation $\sigma(z)$ with the same z dependence. This equivalence has been used to modify dispersion of gratings without actually varying the grating period [62].

Incorporating discrete phase shifts and superstructure into the direct-integration approach is straightforward. For example, as the integration proceeds along z, each time a discrete phase shift is encountered a new constant phase shift is added in Eq. 1 or 14. In the coupled-mode Eqs 15 and 16, we thus multiply the current value of κ by $\exp(i\phi)$ where ϕ is the shift in grating phase. Superstructure is implemented through the z dependence in $\sigma(z)$ and $\kappa(z)$. For example, for sampled gratings we simply set $\kappa = 0$ in the nongrating regions.

The often-preferred piecewise-uniform approach to modeling nonuniform gratings is based on identifying 2×2 matrices for each uniform section of the grating, and then multiplying all of these together to obtain a single 2×2 matrix that describes the whole grating [74]. We divide the grating into M uniform sections and define R_i and S_i to be the field amplitudes after traversing the section i. Thus for Bragg gratings we start with $R_0 = R(L/2) = 1$ and $S_0 = S(L/2) = 0$ and calculate $R(-L/2) = R_M$ and $S(-L/2) = S_M$, while for transmission gratings we start with $R_0 = R(-L/2) = 1$ and $S_0 = S(-L/2) = 0$ and calculate $R(L/2) = R_M$ and $S(L/2) = S_M$. The propagation through each uniform section i is described by a matrix $\mathbf{F}_i$ defined such that

$$\begin{bmatrix} R_i \\ S_i \end{bmatrix} = \mathbf{F}_i \begin{bmatrix} R_{i-1} \\ S_{i-1} \end{bmatrix} \tag{49}$$

For Bragg gratings the matrix $\mathbf{F}_i^B$ is given by

$$\mathbf{F}_i^B = \begin{bmatrix} \cosh(\gamma_B \Delta z) - i\dfrac{\hat{\sigma}}{\gamma_B}\sinh(\gamma_B \Delta z) & -i\dfrac{\kappa}{\gamma_B}\sinh(\gamma_B \Delta z) \\ i\dfrac{\kappa}{\gamma_B}\sinh(\gamma_B \Delta z) & \cosh(\gamma_B \Delta z) + i\dfrac{\hat{\sigma}}{\gamma_B}\sinh(\gamma_B \Delta z) \end{bmatrix} \tag{50}$$

where Δz is the length of the ith uniform section, the coupling coefficients $\hat{\sigma}$ and κ are the local values in the ith section, and

$$\gamma_B \equiv \sqrt{\kappa^2 - \hat{\sigma}^2} \tag{51}$$

Note γ_B is imaginary at wavelengths for which $|\hat{\sigma}| > \kappa$. For transmission gratings the matrix $\mathbf{F}_i^t$ is

$$\mathbf{F}_i^t = \begin{bmatrix} \cos(\gamma_t \Delta z) + i\dfrac{\hat{\sigma}}{\gamma_t}\sin(\gamma_t \Delta z) & i\dfrac{\kappa}{\gamma_t}\sin(\gamma_t \Delta z) \\ i\dfrac{\kappa}{\gamma_t}\sin(\gamma_t \Delta z) & \cos(\gamma_t \Delta z) - i\dfrac{\hat{\sigma}}{\gamma_t}\sin(\gamma_t \Delta z) \end{bmatrix} \tag{52}$$

where in this case

$$\gamma_c \equiv \sqrt{\kappa^2 + \hat{\sigma}^2} \tag{53}$$

Once all of the matrices for the individual sections are known, we find the output amplitudes from

$$\begin{bmatrix} R_M \\ S_M \end{bmatrix} = \mathbf{F} \begin{bmatrix} R_0 \\ S_0 \end{bmatrix}; \quad \mathbf{F} = \mathbf{F}_M \cdot \mathbf{F}_{M-1} \cdots \mathbf{F}_i \cdots \mathbf{F}_1 \tag{54}$$

The number of sections needed for the piecewise-uniform calculation is determined by the required accuracy. For most apodized and chirped gratings

$M \sim 100$ sections is sufficient. For quasi-uniform gratings like discrete-phase-shifted and sampled gratings, M is simply determined by the number of actual uniform sections in the grating. M may not be made arbitrarily large, since the coupled-mode-theory approximations that lead to Eqs 15–16 and 34–35 are theoretically not valid when a uniform grating section has too few grating periods [74]. Thus we require $\Delta z \gg \Lambda$, which means we must maintain

$$M \ll \frac{2n_{eff}L}{\lambda_D} \tag{55}$$

To implement the piecewise-uniform method for apodized and chirped gratings, we simply assign constant values σ, κ, and $(1/2)d\phi/dz$ to each uniform section, where these might be the z-dependent values of $\sigma(z)$, $\kappa(z)$, and $(1/2)d\phi/dz$ evaluated at the center of each section. For phase-shifted and sampled gratings, we insert a phase-shift matrix $\mathbf{F}_{pi}$ between the factors $\mathbf{F}_i$ and $\mathbf{F}_{i+1}$ in the product in Eq. 54 for a phase shift after the ith section. For Bragg gratings the phase-shift matrix is of the form

$$\mathbf{F}_{pi}^{B} = \begin{bmatrix} \exp(-i\phi/2) & 0 \\ 0 & \exp(i\phi_i/2) \end{bmatrix} \tag{56}$$

and for transmission gratings $\mathbf{F}_{pi}^{t} = (\mathbf{F}_{pi}^{B})^{*}$, since we propagate the field amplitudes in the $-z$ direction for Bragg gratings but the $+z$ direction for transmission gratings. Here ϕ_i is the shift in the phase of the grating itself for discrete phase shifts, and for sampled gratings (see Fig. 10(f))

$$\frac{\phi_i}{2} = \frac{2\pi n_{eff}}{\lambda}\Delta z_0 \tag{57}$$

where Δz_0 is the separation between two grating sections.

Having described two basic approaches for calculating reflection and transmission spectra through nonuniform gratings, we now give several examples that demonstrate the effects of apodization, chirp, phase shifts, and superstructure on the optical properties of fiber gratings. For most of the examples in the remainder of this section the piecewise-uniform method was used because of its speed, but the results are indistinguishable from those obtained by direct integration.

To demonstrate the effects of apodization, Fig. 18 shows the reflection and group delay vs. wavelength for gratings similar to those described in Fig. 15, only here the gratings have a Gaussian profile as illustrated in Fig. 10(b) and described by Eq. 45. The maximum index change values $\overline{\delta n_{eff}}$ and the *fwhm* = 1 cm for the gratings in Fig. 18(a) and (b) are the same as the uniform index change values and the length of the uniform gratings in Fig. 15(a) and (b), respectively. We see that the spectra are similar, except there are no sidelobes on the long-wavelength side and different sidelobes on the short-wavelength

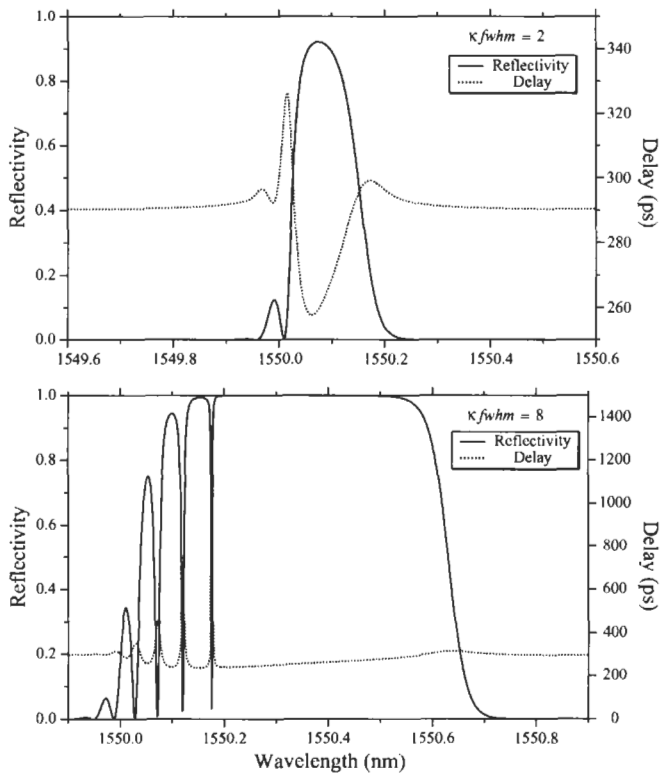


Fig. 18 Reflection and group delay vs. wavelength for Gaussian gratings similar to the uniform gratings in Fig. 7: (a) $\kappa fwhm = 2$, and (b) $\kappa fwhm = 8$.

side of the Gaussian spectra. The short-wavelength structure is caused by the nonuniform DC index change and has been described in detail elsewhere [59,75]. Short wavelengths lie inside the local band gap ($|\hat{\sigma}| < \kappa$) associated with the wings of the grating and thus are strongly reflected there, but lie outside the local band gap near the center of the grating where they are only weakly reflected; the wings of the grating thus act like a Fabry–Perot cavity at short wavelengths. Note that the difference in nonresonant delay (at wavelengths far from the Bragg resonance) between Figs. 15 and 18 is not significant. The *relative* delay is more relevant than the absolute delay, which is sensitive to the identification of time zero (or $z = L/2$ in this case).

In Fig. 19(a) the reflection spectra for several nonuniform gratings are plotted, where for each the DC index change is assumed to be zero ($\sigma \to 0$), as illustrated in Fig. 10(c). In the models described above, the DC index change is set to a small value ($\overline{\delta n_{eff}} \to 0$) while the AC index change is maintained at the desired value, here $\nu\overline{\delta n_{eff}} = 1 \times 10^{-3}$, by assuming a large fringe visibility $\nu \to \infty$. The grating profiles include a Gaussian, a raised-cosine, and a flattop

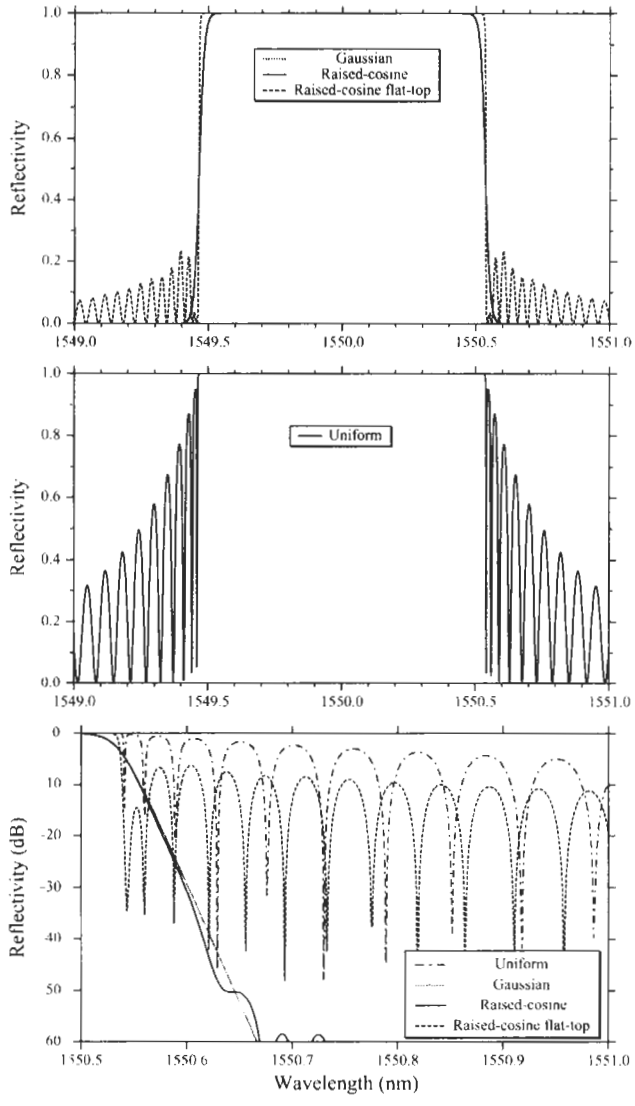


Fig. 19 Calculated reflectivity of uniform and nonuniform gratings with AC index change of 1×10^{-3}, zero DC index change, and *fwhm* = 10 mm. (a) Gaussian (dotted line), raised-cosine (solid line), and flattop raised-cosine (dashed line). (b) Uniform grating. (c) Expanded view of filter edge in dB.

raised-cosine (in which the length of the uniform region is twice the *fwhm* of the cosine region) with the total *fwhm* = 10 mm for each. The Gaussian and raised-cosine gratings are indistinguishable on this plot. Note that the side lobes on the short-wavelength side of the Gaussian and raised-cosine spectra (cf. Fig. 18)

have been eliminated by making the DC index change uniform, and the spectra are symmetric about λ_{max}. For comparison, Fig. 19(b) shows a 10 mm uniform grating with zero DC index change. To investigate how well the apodized gratings approximate a top-hat spectrum, Fig. 19(c) shows a close-up view of the filter edges on a logarithmic scale. The (minimal) differences between the Gaussian and raised-cosine gratings are evident here; although the inherently truncated raised-cosine does exhibit side lobes, they are below −50 dB.

An example of gratings with (solid line) and without (dotted line) a discrete phase shift is shown in Fig. 20. The gratings have a raised-cosine shape with *fwhm* = 10 mm, maximum AC index change $\nu\overline{\delta n_{eff}} = 2 \times 10^{-4}$, and no DC index change ($\overline{\delta n_{eff}} \to 0$). The π phase shift of the grating phase at the center opens a narrow transmission resonance at the design wavelength, but also broadens the reflection spectrum and diminishes r_{max}.

The reflection spectrum of a sampled grating exhibiting periodic superstructure is shown in Fig. 21 (solid line). As illustrated in Fig. 10(f), the grating is uniform except for regions where $\nu\overline{\delta n_{eff}}$ is set to zero. Here the AC index change in the nonzero regions is $\nu\overline{\delta n_{eff}} = 5 \times 10^{-4}$, the total length of the grating is 100 mm, and there are 50 sections with an on-off duty cycle of 10%. To facilitate an understanding of the resulting spectrum, the dashed line in Fig. 21 shows the reflection spectrum produced by a single section of this grating with the same index change; since its length is 200 μm, $\kappa L = 0.2$ for this section. The zeros in the envelope of the sampled-grating spectrum coincide with the zeroes of the single-section spectrum. Also shown in Fig. 21 (dotted line) is a plot of the spectrum that results when the sampled grating is compressed,

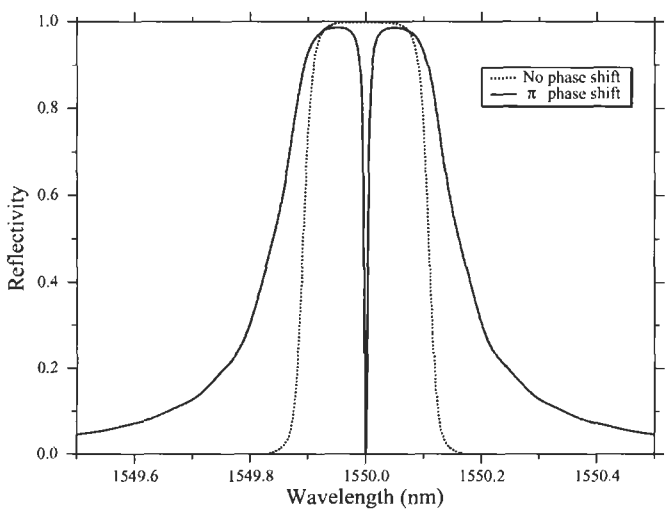


Fig. 20 Calculated reflectivity of a raised-cosine grating with (solid line) and without (dotted line) a π phase shift in the phase of the grating at the center.

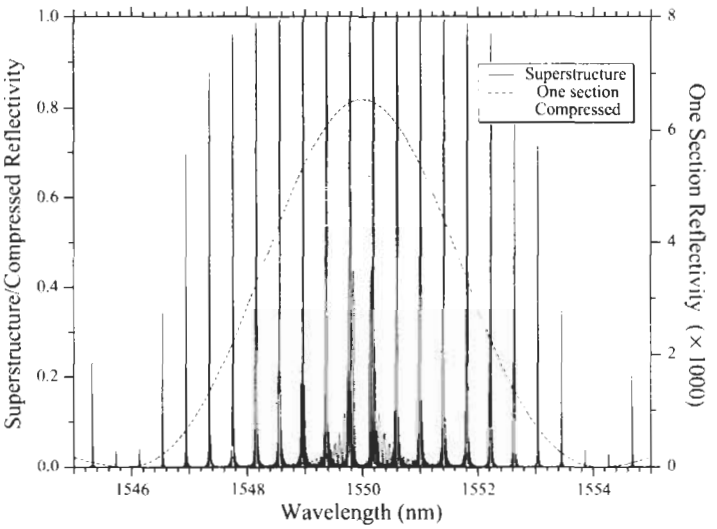


Fig. 21 Calculated reflectivity of a uniform grating with periodic superstructure (solid line), where the 100-mm-long grating has 50200-μm-long grating sections spaced 1.8 mm apart. Also shown are the reflectivity of a single 200-μm-long section (dashed line) and of the compressed grating (dotted line). in which the 1.8 mm grating-free sections are removed.

such that the total length of the modulated portion of the grating is maintained (10 mm), but the unmodulated regions are removed.

An example of a chirped grating is shown in Fig. 22. The grating has a zero-DC raised-cosine profile with $fwhm = 1$ cm, $\nu\overline{\delta n_{eff}} = 5 \times 10^{-4}$, and a chirp of $d\lambda_D/dz = -1$ nm/cm. The plot shows the group delay calculated from Eq. 32 (dashed line), the dispersion calculated from Eq. 33 (solid line), and the reflection spectrum (inset). The bandwidth of a similar unchirped grating as estimated from Eq. 31 is 0.53 nm. A commonly used estimate of dispersion in a linearly chirped grating when variation of grating period along z is the dominant source of chirp is

$$d_\rho \sim 100 \left(\frac{d\lambda_D}{dz} \right)^{-1} \left(\frac{ps}{nm} \right) \tag{58}$$

where $d\lambda_D/dz$ is in nm/cm, and we have approximated $2n_{eff}/c \cong 100$ ps/cm. According to Eq. 58, the grating in Fig. 22 should exhibit a dispersion of $d_\rho = -100$ ps/nm and thus a delay of $\tau_\rho = 200$ ps between 1549 and 1551 nm. These simple estimates disagree with the actual values by a factor of almost 2, mainly because the effective chirp resulting from apodization is comparable to the grating-period chirp.

Several examples of bar-transmission spectra through nonuniform transmission gratings are shown in Fig. 23. The solid line shows transmission

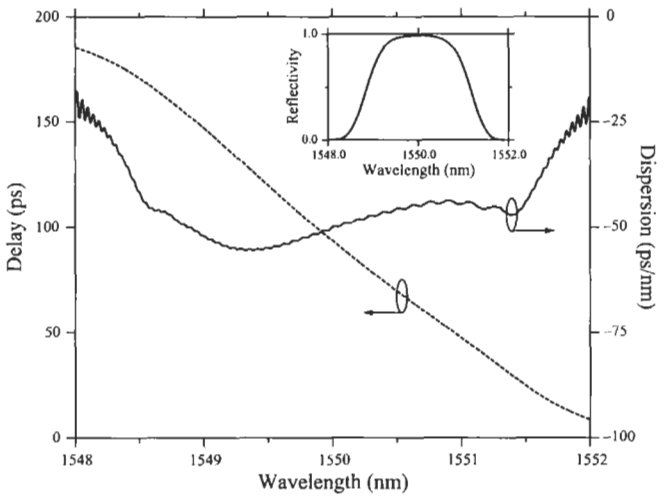


Fig. 22 Calculated group delay (dashed line) and dispersion (solid line) of a raised-cosine grating with AC index change of 5×10^{-4}, zero DC index change, *fwhm* = 10 mm, and a chirp of −1 nm/cm. The inset shows the reflectivity spectrum.

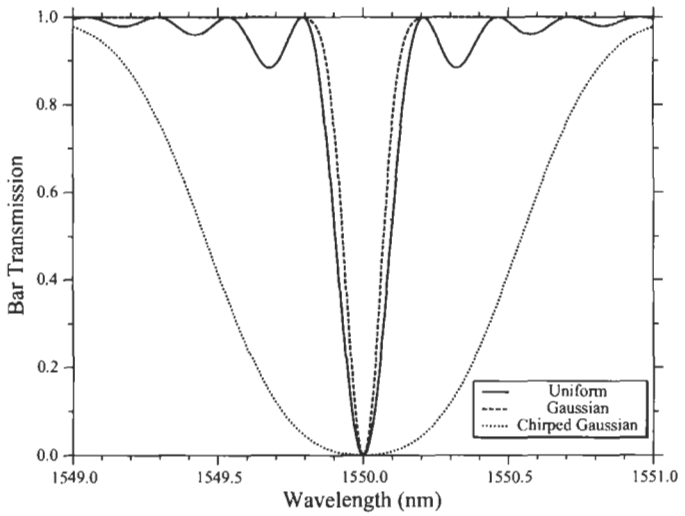


Fig. 23 Calculated bar transmission $t_=$ through a uniform grating (solid line), a Gaussian grating (dashed line), and a chirped Gaussian grating (dotted line) with grating parameters given in the text.

through a uniform grating with a coupling-length product of $\kappa L = \pi/2$, zero-DC index change ($\sigma_{11} = \sigma_{22} = 0$), an effective-index difference of $\Delta n_{eff} = 0.1$, and a length of 10 cm. The dashed line represents the transmission through a similar grating but with a Gaussian profile, where the

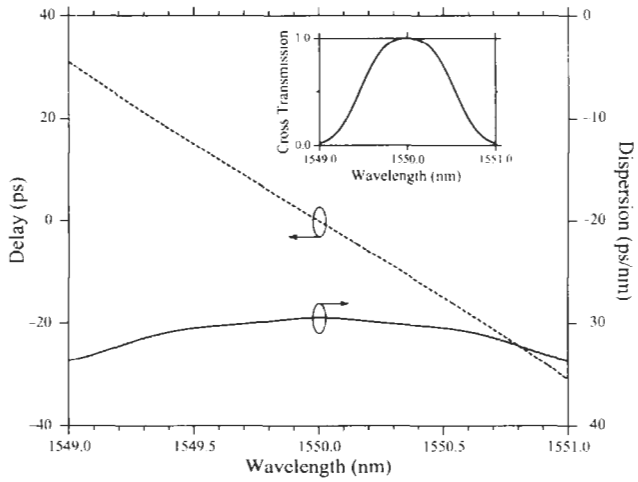


Fig. 24 Group delay and dispersion for the cross amplitude of light transmitted through the chirped Gaussian grating described in Fig. 15. Cross transmission t_x is shown in the inset.

coupling-length product has been slightly reduced to $\kappa \cdot fwhm = 0.48\pi$ to maintain maximum coupling. Note the reduction of the side lobes, just as occurs for Bragg gratings. Although no example is shown, if the DC index change were nonzero, as shown in Fig. 10(b), the Gaussian grating spectrum would exhibit Fabry–Perot structure on the short-wavelength side of the spectrum similar to that shown in Fig. 18. The dotted line in Fig. 23 is the transmission spectrum obtained for a similar Gaussian grating with a linear chirp of $d\lambda_D/dz = -0.025\,\text{nm/cm}$ and $\kappa \cdot fwhm = 1.4\pi$. Figure 24 shows the spectrum of the group delay and dispersion associated with this grating; the cross transmission is shown in the inset. Although the dispersion is relatively small for this example, note how linear the delay is over the full bandwidth of the grating.

G. TILTED GRATINGS

Another parameter a designer of fiber gratings can use is the tilt of the grating fringes in the core of the fiber [76,77]. The main effect of grating tilt in a single-mode Bragg grating is to effectively reduce the fringe visibility ν defined in Eqs 1 and 13–14. However, grating tilt can also dramatically affect the coupling to radiation modes, which we discuss briefly in Section J, below, as well as enable otherwise unallowed coupling between discrete bound modes of fiber. The latter interaction has been demonstrated in both Bragg [78] and transmission gratings [79] that couple the LP_{01} mode to the LP_{11} mode of a dual-mode fiber.

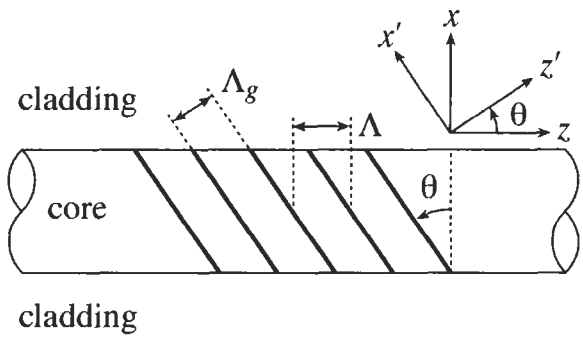


Fig. 25 Diagram of the parameters associated with a tilted phase grating in the core of an optical fiber.

To see how grating tilt affects single-mode Bragg reflection in a fiber grating, suppose that the induced index change in the core of the fiber δn_{co} is rotated by an angle θ, such that it is

$$\delta n_{co}(x,z) = \overline{\delta n_{co}}(z')\left(1 + v\cos\left[\frac{2\pi}{\Lambda_g}z' + \phi(z')\right]\right) \tag{59}$$

where the z' axis, shown in Fig. 25, is defined to be $z' = x\sin\theta + z\cos\theta$. The grating period along the z (fiber) axis, which determines the resonant wavelengths for coupling, is thus $\Lambda = \Lambda_g/\cos\theta$. For the slowly varying functions $\overline{\delta n_{co}}(z')$ and $\phi(z')$, we take $z' \cong z\cos\theta$; i.e., we simply take the projection of these functions onto the fiber axis. The general coupling coefficient in Eq. 14 then becomes

$$K'_{\mp\pm}(z) = \sigma(z) + 2\kappa_{\mp\pm}(z)\cos\left[\frac{2\pi}{\Lambda}z + \phi(z\cos\theta)\right] \tag{60}$$

where we recognize that the subscripts j and k describe the same mode, except when one is associated with the forward-going mode (+) the other describes the backward-going mode (−). The self- and cross-coupling coefficients now become

$$\sigma(z) = \frac{\omega n_{co}}{2}\overline{\sigma n_{co}}(z\cos\theta)\iint\limits_{core} dxdy\boldsymbol{e}_{\mp t}(x,y)\cdot\boldsymbol{e}_{\pm t}(x,y) \tag{61}$$

and

$$\kappa_{\mp\pm}(z,\theta) = \frac{v}{2}\frac{\omega n_{co}}{2}\overline{\delta n_{co}}(z\cos\theta) \times \iint\limits_{core} dxdy\exp\left(\pm i\frac{2\pi}{\Lambda}x\tan\theta\right)\boldsymbol{e}_{\mp t}(x,y)\cdot\boldsymbol{e}_{\pm t}(x,y) \tag{62}$$

Notice that $\kappa_{+-} = (\kappa_{-+})^*$. Except for scaling of the slowly varying functions, the effects of tilt can be lumped into an "effective fringe visibility" $v_{\mu\pm}(\theta)$, defined such that

$$\frac{v_{\mp\pm}(\theta)}{v} = \frac{\iint\limits_{core} dxdy \exp\left(\pm i \frac{2\pi}{\Lambda} x \tan\theta\right) \boldsymbol{e}_{\mp t}(x,y) \cdot \boldsymbol{e}_{\pm t}(x,y)}{\iint\limits_{core} dxdy \boldsymbol{e}_{\mp t}(x,y) \cdot \boldsymbol{e}_{\pm t}(x,y)} \tag{63}$$

Therefore we may write

$$\kappa_{\mp\pm}(z,\theta) = \frac{v_{\mp\pm}(\theta)}{2}\sigma(z) \tag{64}$$

in direct analogy to Eq. 13. This important result states that the effect of grating tilt on single-mode Bragg reflection is simply to reduce the effective fringe visibility by an amount given in Eq. 63. As an example, Fig. 26 shows a plot of the normalized effective fringe visibility as a function of tilt angle θ for the LP_{01} mode in a fiber with a cladding index of 1.44, a core-cladding $\Delta = 0.0055$, a core radius of $a = 2.625\,\mu m$, and at a wavelength of 1550 nm.

The Bragg reflection spectrum for a tilted grating can still be calculated from Eqs 15 and 16, where now $\kappa \equiv \kappa_{-+}$. Figure 27 shows reflection spectra over a range of tilt angles for a Gaussian grating with a *fwhm* = 5 mm, a maximum index change of $\overline{\delta n_{eff}} = 7.25 \times 10^{-4}$, and zero-tilt visibility $v = 1$. Because

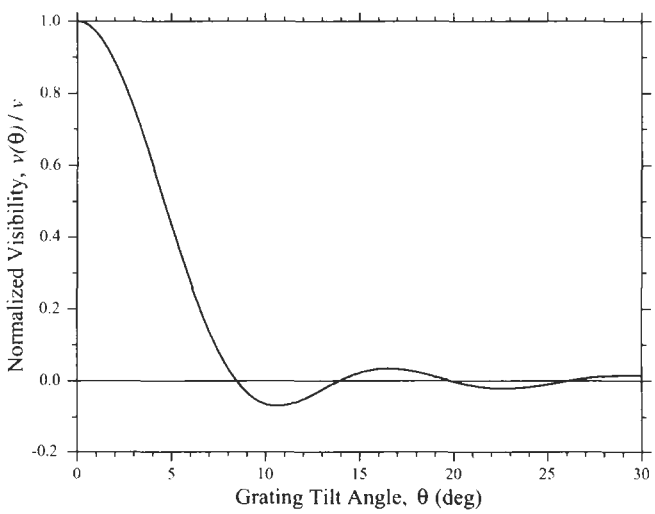


Fig. 26 Plot of the normalized effective fringe visibility associated with single-mode Bragg reflection in a tilted grating in the core of a fiber with parameters described in the text.

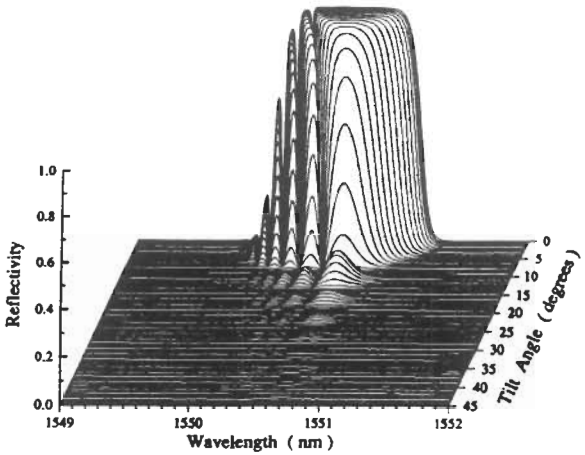


Fig. 27 Calculated reflectivity spectrum over a range of grating tilt angles for a Gaussian grating.

tilt affects Bragg reflection by effectively reducing the fringe visibility, this plot also demonstrates how the reflection spectrum of a Gaussian grating changes as the visibility is reduced (here by keeping the DC index change constant). Proper use of tilt for reducing the effective visibility and hence Bragg reflection can be practically useful in cases where large radiation or other bound mode coupling is desired but Bragg reflection is undesirable [78].

H. COUPLING TO CLADDING MODES

There are numerous applications of fiber gratings that involve coupling of light into and out of the core mode(s) of a fiber. Figure 28 shows measured spectra of the transmission of an LP_{01} core mode through an untilted Gaussian fiber grating that is 5 mm long and has a maximum index change of $\overline{\delta n_{eff}} = 2 \times 10^{-3}$. In Fig. 28(a) the bare (uncoated) section of fiber that contains the grating is immersed in index-matching fluid to simulate an infinite cladding. What results is a smooth transmission profile for $\lambda < 1540$ nm demonstrating loss due to coupling of the core mode to the continuum of radiation modes. Radiation mode coupling is briefly discussed in Section I and is described in detail in [75,76]. In Fig. 28(b) the fiber is immersed in glycerin, which has a refractive index greater than the cladding index. The transmission spectrum now exhibits fringes that are caused by Fabry–Perot-like interference resulting from partial reflection of the radiation modes off the cladding-glycerin interface. In Fig. 28(c) the bare fiber is surrounded by air; for the wavelengths shown in this plot, the LP_{01} core mode couples into distinct backward-going cladding modes

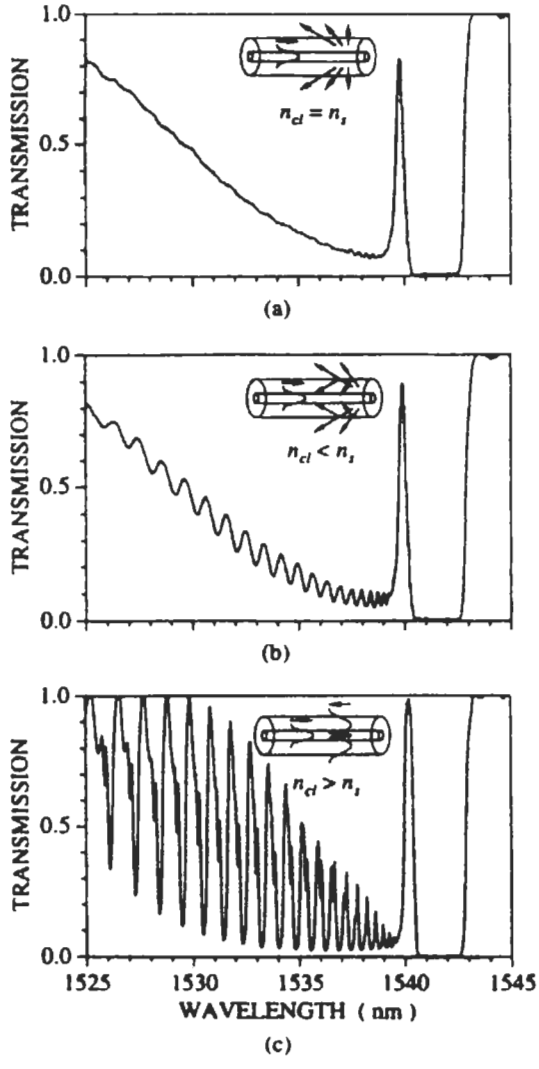


Fig. 28 Measured transmission through a Bragg grating where (a) the uncoated fiber is immersed in index-matching liquid to simulate an infinite cladding; (b) the fiber is immersed in glycerin; and (c) the bare fiber is surrounded by air and thus supports cladding modes.

with a well-defined resonance peak for each mode. Here we briefly examine this last type of interaction. A more detailed description can be found in [80].

Figure 12(b) illustrates coupling of a core mode to a cladding mode from a ray point of view, where cladding modes are bound modes of the total fiber with $1 \leq n_{eff} \leq n_{cl}$ (assuming the fiber is surrounded by air). It is shown above

that the grating period required for copropagating coupling between a core mode of index n_{eff}^{co} and a cladding mode of index n_{eff}^{cl} is $\Lambda = \lambda/(n_{eff}^{co} - n_{eff}^{cl})$. In analogous fashion, counterpropagating modes ($\theta_2 < 0$ in Fig. 12(b)) can be coupled by a grating with a much shorter period $\Lambda = \lambda/(n_{eff}^{co} + n_{eff}^{cl})$.

Using the coupled-mode theory outlined above, the equations [80] that describe coupling among the LP_{01} core mode and the (exact) cladding modes of order 1 μ by a reflection grating are

$$\frac{dA^{co}}{dz} = i\sigma_{01-01}^{co-co}A^{co} + i\sum_{\mu}\kappa_{1\mu-01}^{cl-co}B_{\mu}^{cl}\exp\left(i2\delta_{1\mu-01}^{cl-co}z\right) \tag{65}$$

$$\sum_{\mu}\left(\frac{dB_{\mu}^{cl}}{dz} = -i\kappa_{1\mu-01}^{cl-co}A^{co}\exp\left(+i2\delta_{1\mu-01}^{cl-co}z\right)\right) \tag{66}$$

where the detuning is

$$\delta_{1\mu-01}^{cl-co} \equiv \frac{1}{2}\left(\beta_{01}^{co} + \beta_{1\mu}^{cl} - \frac{2\pi}{\Lambda}\right) \tag{67}$$

Here σ_{01-01}^{co-co} is the self-coupling coefficient for the LP_{01} mode given by Eq. 19, $\kappa_{1\mu-01}^{cl-co}$ is the cross-coupling coefficient defined through Eqs 12 and 13, and we have neglected terms that involve self- and cross-coupling between cladding modes, since the associated coupling coefficients are very small, or $\sigma_{1\mu-1\mu}^{cl-cl}, \kappa_{1\mu-01}^{cl-cl} \ll \kappa_{1\mu-01}^{cl-co} \ll \sigma_{01-01}^{co-co}$. In Eqs 65 and 66 we have also not included Bragg reflection of the LP_{01} core mode, making these equations valid only at wavelengths far away from this Bragg resonance.

The common synchronous approximation was used in obtaining Eqs 65 and 66, but unlike Eqs 15 and 16, the resulting equations contain coefficients that oscillate rapidly along z. This is because there are now multiple values of δ, which are not easily removed by defining R and S as we did for Eqs 15 and 16. Hence we must integrate the coupled-mode equations numerically even for a uniform grating, and for nonuniform gratings we may not employ the piecewise-uniform approach. However, if the cladding mode resonances do not overlap, then we may calculate each resonance separately, retaining only the core mode and the appropriate cladding mode for the resonance. In that case the problem reduces to simple 2-mode coupling, as described in detail in Sections III D, E, and F.

As an example of cladding-mode coupling in a Bragg grating, Fig. 29 shows the measured (a) and calculated (b) transmission spectrum through a Gaussian grating written in hydrogen-loaded [11] Corning Flexcore fiber. The grating length is *fwhm* = 4 mm, the index change is $\overline{\delta n_{eff}} = 2.8 \times 10^{-3}$, and the visibility is approximately $v = 1$. The large transmission dip at 1541 nm is the Bragg reflection resonance, while the remaining dips are cladding mode resonances. Here the resonances overlap substantially over most of the wavelength range

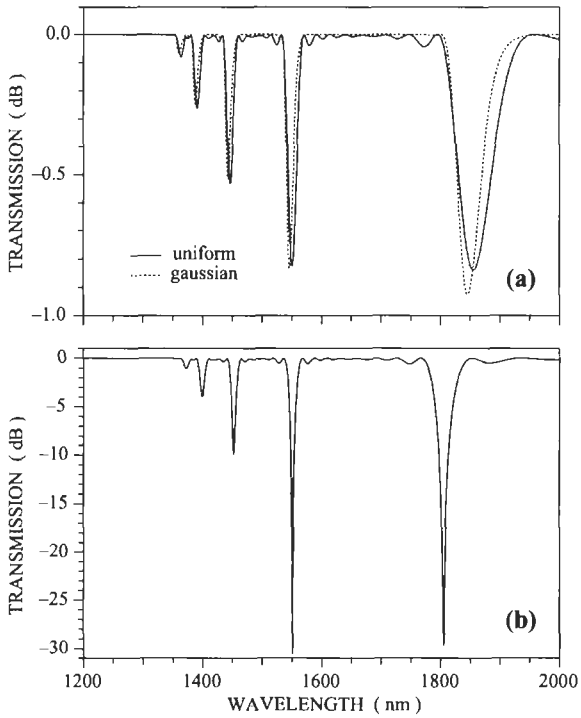


Fig. 30 Calculated transmission spectra through a relatively weak (a) and a stronger (b) transmission grating, each designed to couple the LP_{01} core mode to the $\mu = 5$ cladding mode at 1550 nm. Solid lines represent a uniform grating while the dashed line represents a Gaussian grating.

weak grating with a length (*fwhm*) of 25 mm, a maximum index change of $\overline{\delta n_{eff}} = 2 \times 10^{-4}$, and a uniform (solid line) or Gaussian (dashed line) profile. The four main dips seen in these spectra correspond to coupling to the $\mu = 1, 3, 5$, and 7 cladding modes. Figure 30(b) shows the transmission for a stronger, uniform grating with an induced index change of 1×10^{-3}. In each case the grating period is adjusted to achieve coupling at 1550 nm between the LP_{01} core mode and the $\mu = 5$ cladding mode. That is, the gratings in Fig. 30(a) and (b) have periods of $\Lambda = 545\,\mu m$ and $\Lambda = 502\,\mu m$, respectively.

In the above discussion we consider only the $l\mu$ cladding modes with azimuthal order $l = 1$. In an untilted grating, coupling may occur only between modes of the same azimuthal symmetry (as shared by the $LP_{01} = HE_{11}$ mode and the exact $HE/EH_{1\mu}$ cladding modes). In a typical 125 μm diameter fiber there are several hundred $l = 1$ cladding modes at IR communications wavelengths. In a tilted grating, in which coupling to an enormous number

of higher-azimuthal-order cladding modes is allowed, the resonances are so densely spaced that they are generally not well resolvable.

I. RADIATION-MODE COUPLING

Although numerous applications are possible for fiber gratings as waveguide-grating couplers that couple free-space beams into and out of bound fiber modes, these applications are only in the early stages of development. The effect of radiation-mode coupling as a loss mechanism for core-mode transmission has been studied [75,76], and an overview of this analysis is given below. Then we briefly describe approaches that have been taken to suppress radiation-mode coupling, followed by one technique particularly well suited to the analysis of the out-coupled radiation, and finally while applications of radiation-mode coupling gratings are detailed in Section IV B.3.

Figure 31 illustrates the coupling of an LP_{01} core mode to the continuum of backward propagating radiation modes from a ray point of view. The coupling strength to these high-order (small-θ_2) modes is quite small unless the grating is suitably blazed. But as Fig. 28(a) illustrates, if the fiber is immersed in an index matching fluid or recoated with a high-index polymer, bound cladding modes cease to exist and substantial coupling to low-order ($|\beta_2| > 2\pi/\lambda$) radiation modes can occur near the Bragg reflection resonance even in untilted gratings. A useful quantity to be aware of is the wavelength λ_{cut} at which true radiation-mode (not cladding-mode) coupling "cuts" on and/or off. Assuming the core confinement factor for the

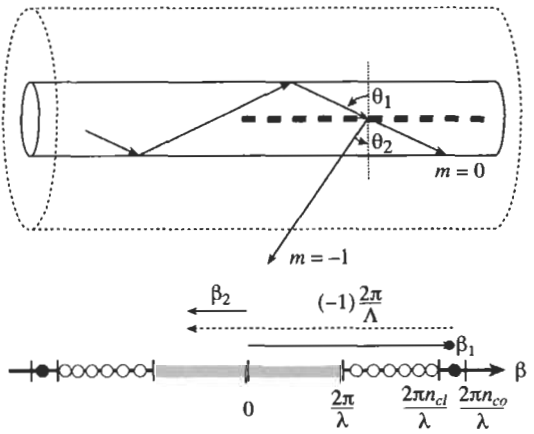


Fig. 31 Ray-optic illustration of core mode coupling to a backward-traveling radiation mode by a Bragg grating. The β axes below the diagram demonstrates the grating condition in Eq. 5 for $m = -1$.

radiation modes is much smaller than that of the bound mode of interest, then we find

$$\lambda_{cut} \cong \frac{1}{2}\left(1 \pm \frac{n}{n_{eff}}\right)\left(1 + \frac{\overline{\delta n_{eff}}}{n_{eff}}\right)\lambda_D \tag{71}$$

where $n = n_{cl}$ for the case of an infinitely clad fiber with index n_{cl} (see Fig. 28(a)), or $n = 1$ when the fiber is surrounded by air such that bound cladding modes may propagate. The "+" sign in the first factor applies to reflection, while the "−" sign corresponds to forward coupling. The second factor describes the shift of λ_{cut} with increasing DC index change.

Consider the coupling of an LP_{01} core mode to backward-propagating radiation modes labeled $LP_{p\rho}$, where the discrete index p identifies the polarization and azimuthal order, while the continuous label $\rho = \sqrt{(2\pi/\lambda)^2 n_{cl}^2 - \beta_{p\rho}^2}$ denotes the transverse wavenumber of the radiation mode ($\beta_{p\rho}$ is the usual axial propagation constant). Here we assume the fiber has a cladding of index n_{cl} and with an infinite radius. The coupled-mode equations for this case are then

$$\frac{dA}{dz} = i\sigma_{01-01}^{co-co}A + i\sum_{p}\int d\rho \kappa_{p\rho-01}^{ra-co}B_{p\rho}\exp\left(-i2\delta_{p\rho-01}^{ra-co}z\right) \tag{72}$$

$$\frac{dB_{p\rho}}{dz} = -i\kappa_{01-p\rho}^{co-ra}A\exp\left(+i2\delta_{p\rho-01}^{ra-co}z\right) \tag{73}$$

where

$$\delta_{p\rho-01}^{ra-co} \equiv \frac{1}{2}\left(\beta_{01} + \beta_{p\rho} - \frac{2\pi}{\Lambda}\right) \tag{74}$$

and where A is the amplitude of the core mode, $B_{p\rho}$ is the amplitudes of the (continuous spectrum of) radiation modes, and the usual summation now includes an integral in Eq. 72. Also, σ_{01-01}^{co-co} is the LP_{01} self-coupling coefficient given by Eq. 19, and $\kappa_{01-p\rho}^{co-ra} = \left(\kappa_{p\rho-01}^{ra-co}\right)^*$ is the cross-coupling coefficient defined through Eqs. 12 and 13. By applying essentially a first-born approximation to the core mode amplitude A, we can obtain an approximate expression for $B_{p\rho}$ from Eq. 73. After inserting this result into Eq. 72 and performing the integral over ρ, it can be shown [75,76] that the core mode amplitude approximately obeys an equation of the form

$$\frac{dA}{dz} = i\sigma_{01-01}^{co-co}A - \left[\sum_{p}\frac{\pi\beta_{p\rho}}{\rho}\left|\kappa_{p\rho-01}^{ra-co}\right|^2\right]A \tag{75}$$

where the term in square brackets is evaluated at $\beta_{p\rho} = 2\pi/\Lambda - \beta_{01}$. Since this term is real, clearly it gives rise to exponential loss in the amplitude of the core mode. Notice the loss coefficient is proportional to the square of

the cross-coupling coefficient, and hence to the square of the induced index change.

Although the analysis given here is simplified, it is possible to predict the radiation-mode coupling loss spectrum even when the grating is tilted following a similar development to that in Section III G. As might be expected, as the tilt angle is increased (the grating is blazed), light can be coupled more efficiently at smaller angles $|\theta_2|$. But increasingly many radiation-mode azimuthal orders must be included in the summation in Eq. 75 to accurately model radiation modes propagating more normal to the fiber axis. As an example of some typical transmission spectra, Fig. 32 shows a calculation of the loss in transmission through the same Gaussian grating described in Fig. 27, only here $\overline{\delta n_{eff}} = 1.5 \times 10^{-3}$. Spectra are shown for a range of tilt angles between 0 and 45°, where the grating period Λ along the fiber axis is kept constant. The LP_{01} mode incident on the grating is assumed to be polarized perpendicular to both the x and z axes (see Fig. 25). The peak at the longest wavelength results from Bragg reflection, whereas the loss at other wavelengths is due to radiation-mode coupling. As expected, the efficiency for coupling to smaller and smaller angles θ_2 (which occurs at shorter wavelengths) improves as the grating is increasingly blazed.

Suppression of the radiation-mode (or cladding-mode) loss is critical when a fiber grating is to be used over a wide spectral bandwidth such that high transmission at wavelengths below the Bragg wavelength is important. Equations 72 and 73 show that the most reliable way to reduce the coupling loss is to nullify the coupling coefficients $\kappa_{p\rho-01}^{ra-co}$ and $\kappa_{01-p\rho}^{co-ra}$. If we look at Eqs 12 and 13, and recall that the bound and radiation modes are orthogonal, then the origin of the nonzero coupling coefficients is the restriction of the region of integration in the integrals to the core region only. Hence, if the photosensitivity and resultant grating perturbation extend out into the cladding, covering the full extent of the bound core mode, the coupling coefficient integrals would

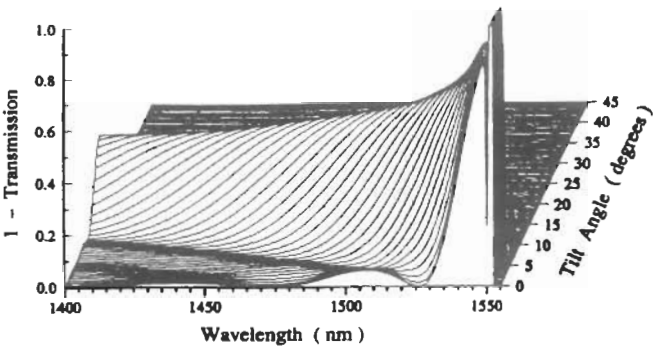


Fig. 32 Calculated loss in transmission through a Gaussian grating in a fiber with an infinite cladding over a range of grating tilt angles.

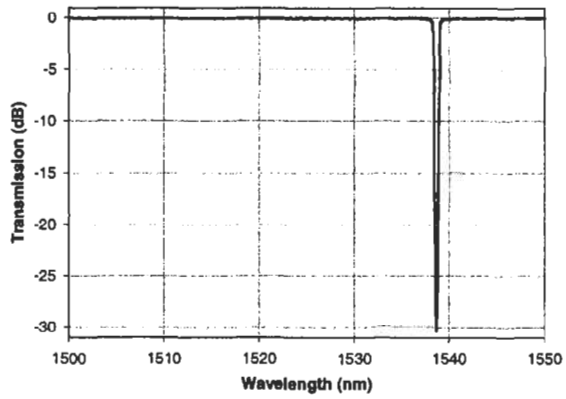


Fig. 33 Transmission through a strong grating written in a fiber designed to minimize coupling to radiation and cladding modes [36].

go to zero. A number of groups have designed and manufactured fibers with such a photosensitive cladding region. Recently, Dong, et al., demonstrated fibers capable of producing very strong gratings (dip in transmission exceeding 30 dB) while keeping the cladding mode losses less than about 0.1 dB [82]. Figure 33 shows a plot of the transmission spectrum for such a grating.

Coupled-mode theory is useful for analyzing the effects of radiation-mode coupling on the transmission of a bound core mode, but it is not as convenient for computing the actual radiated fields from a grating. A technique well suited for this purpose is the volume current method, in which a dielectric perturbation is treated as a volume current source term in Maxwell's equations. The induced vector potential field and resulting magnetic and electric fields are then computed, followed by the Poynting vector that represents the intensity of the radiated light [83]. By integrating the Poynting vector over all directions of radiation, the total radiated power and its dependence on wavelength can also be computed. Figure 34 shows an example of the measured and computed azimuthal dependence of the radiated intensity at normal incidence to the fiber from a 45-degree blazed grating written in Corning Flexcore fiber [83]. The agreement is quite good for the main peak, although there is some discrepancy far down in the sidelobes.

J. GRATING SYNTHESIS

An exciting recent development in the design and analysis of fiber gratings is the ability to theoretically synthesize very complex grating structures. The problem amounts to finding a grating amplitude and phase description, i.e., the apodization and chirp profiles, that give a specified complex spectral response. Synthesis is useful both as a design tool and for characterization of already fabricated gratings with complex profiles. Though it has been recognized for

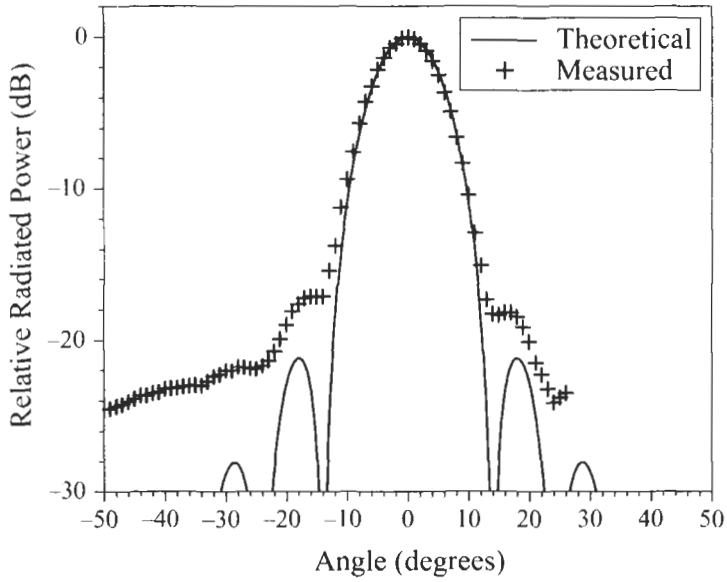


Fig. 34 Measured and calculated angular dependence of the light radiated by a 45° tilted grating normal to a fiber axis.

some time that techniques such as the iterative Fourier transform method [84–86] and the Gel'Fand-Levitan-Marchenko inverse scattering approach [87,88] could be applied to waveguide grating synthesis, it has been a common view that the inverse problem is significantly more complicated to solve than the direct, forward problem of determining a grating spectrum from a given structure. However, the recent articles by R. Feced, et al. [89] and L. Poladian [90] show that in fact the synthesis problem actually is as simple as the forward problem. One can find the grating structure from the reflection spectrum simply by propagating the fields along the grating structure while simultaneously evaluating the grating strength using a simple causality argument.

The methods developed by Feced and Poladian are quite similar, but not identical. The former is based on a discrete matrix propagation technique, whereas the latter is based on a continuous treatment of the amplitude evolution along the grating [91]. Nevertheless, both are based on the following procedure: by causality, the coupling coefficient at the front end of the grating is determined only by the leading edge of the impulse response (Fourier transform of the spectral response), since at the very beginning of the impulse response light does not have time to propagate more deeply into the grating and hence "sees" only the first layer. After computing the value of the coupling coefficient in the first layer, the fields are propagated to the next layer of the grating. This propagation is essentially the direct (forward) calculation and can be accomplished by numerical integration of the coupled-mode equations

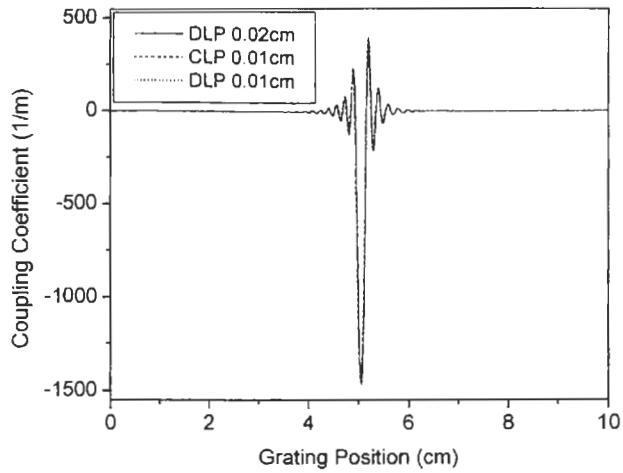


Fig. 35 Computed variation of the magnitude of the coupling coefficient for a grating synthesized to achieve a nearly square, dispersionless filter shape. Three calculations are compared, including discrete layer peeling calculations using two different layer thicknesses, and a continuous layer peeling calculation.

or by the transfer-matrix method. Now one is in the same situation as at the beginning, since the effect of the first layer is "peeled off." The process is continued to the back of the grating, so that the entire grating structure is reconstructed.

As an example, Fig. 35 shows the amplitude of the reconstructed grating coupling coefficient that corresponds to a flattop, nearly rectangular, dispersionless passband filter described by the super-Gaussian function

$$r(\delta) = \sqrt{R}\exp\left[-(\delta/\delta_{PB})^{20}\right] \tag{76}$$

where the maximum reflectivity is $R = 90\%$ and the width is determined by $\delta_{PB} = 19.2\,\mathrm{cm}^{-1}$, which corresponds to a full width at half maximum of $37.84\,\mathrm{cm}^{-1}$ in wavenumber, or 1 nm in wavelength at a center wavelength of 1550 nm. Here DLP stands for discrete layer peeling (Feced's approach) and CLP for continuous layer peeling (Poladian's approach). Also given are the assumed layer thicknesses. Figure 36 shows the resulting spectra computed from the synthesized grating structures. While similar, the DLP and CLP yield fundamentally different spectral responses. The CLP filter performance is actually a little better, but the DLP algorithm is more than an order of magnitude faster.

Finally, it is interesting to note that the synthesis techniques based on layer peeling have now been extended to long-period grating synthesis as well [92,93]. The ability to synthesize such filters is remarkable, given that the application of

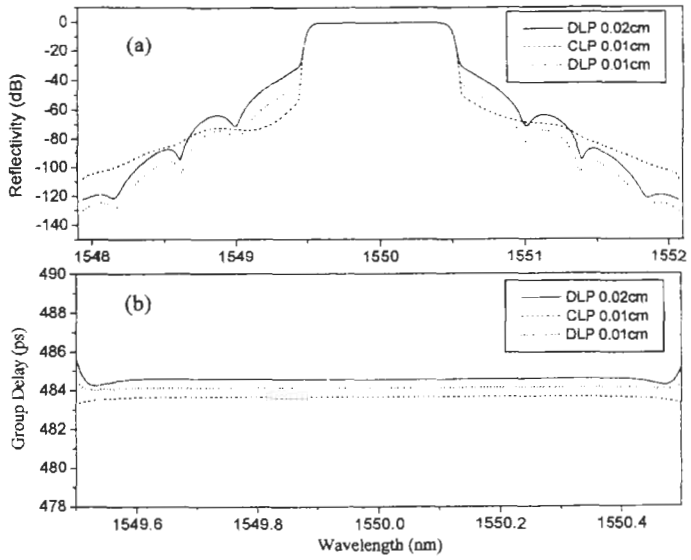


Fig. 36 Spectral responses of the three gratings synthesized in Fig. 35.

causality to the synthesis algorithm is not as simple for long-period gratings. Nevertheless, with a few restrictions the synthesis can be accomplished readily.

IV. Properties and Applications of Fiber Gratings

The properties of fiber gratings are critical to identifying and realizing successful component technologies for communication system applications. The following section will demonstrate that practical implementation of many devices requires control of grating properties to realize the optimum fiber grating device technologies. This section introduces the details of fiber grating properties and how these are exploited, together with different filter characteristics from Section III, to realize devices that are interesting for a wide range of communications system applications.

A. FIBER GRATING PROPERTIES

The primary properties that change the innate optical characteristics of a fiber grating are temperature and strain. Changes in the local temperature or stress shift the resonance wavelength of the coupling between two modes. If the change is uniform over the length of the grating, the filter spectrum typically shifts in wavelength with minimal change in filter shape. As can be deduced from the wavelength resonance formulas in the previous section, the magnitude and direction of the wavelength shift are different for reflective and transmissive

gratings and can be deduced from the basic phase matching relationships if the change in the grating period and effective refractive indices are included.

Bragg reflection gratings have temperature and strain sensitivities that are easily understood and quantified because they couple the same fundamental mode propagating in opposite directions. A change in temperature changes the grating period via thermal expansion and the effective refractive index (via the thermooptic coefficient). The net result in germanosilicate fibers is a Bragg wavelength that increases with temperature at a rate of 7–8 ppm/°C (~0.011 nm/°C at 1550 nm) [94]. Similarly, uniform axial strain directly changes the grating period and slightly changes the effective refractive index according to the stress-optic coefficient of glass. The net result is an increase in wavelength with uniform axial strain at a rate of 0.78 ppm/$\mu\varepsilon$ (~1.3 pm/$\mu\varepsilon$ at 1550 nm) [94].

The temperature and strain characteristics of transmissive gratings also generally exhibit a shift in wavelength with minimal change in filter shape. The rate of change is much less predictable, however, because the center wavelength is inversely proportional to the difference in effective indices of the two modes (Eq. 7). Therefore, small changes in effective index can create very large changes in filter wavelength. This effect can be seen in the literature for long-period gratings [95], which show temperature and strain dependencies that can change sign and vary by an order of magnitude depending on the fiber design and the particular modes coupled.

A.1. Device Technologies

There are a number of technologies that are used to adapt intrinsic fiber grating filtering properties to meet the requirements of a particular communication system application. These technologies generally either stabilize the grating or enable dynamic control of the filtering properties.

Numerous commercial athermal Bragg grating devices are now offered that control the strain and temperature properties of the fiber gratings by packaging them to offset each other ($d\lambda/dT < 0.5$ pm/°C!). This approach holds the grating under strain in a package designed to have a net negative thermal expansion, thereby reducing the fiber strain at a rate to balance the increase in wavelength from temperature (approximately 10 $\mu\varepsilon$/°C). This package can be designed with an unusual glass or ceramic material with a negative expansion [96] or from a hybrid combination of materials with different thermal expansion coefficients [97].

An alternative device technology is to use strain or temperature to provide active control of the center wavelength of a grating. In general, temperature-tuning a packaged grating is easier to ensure stable filtering properties, but strain is a better mechanism when tuning speed is important. Both approaches suffer from a difficulty in readily achieving a 40–80 nm tuning range that is sufficient to cover the broadband WDM communication systems already in

use. The useful temperature tuning range is restricted to a few nanometers by reduced grating stability at higher temperatures and by poor energy efficiency for cooling to lower temperatures. The strain tuning range is also limited to a few nanometers by the tensile strength of the fiber when tuning in tension. Much larger tuning ranges of >30 nm have been reported in compression [98]; however, the fiber tends to buckle when compressed, which can distort the grating filter shape while adding package difficulty and reliability risk. Therefore, practical tunable fiber grating devices have been limited to a few nm of tuning; there are a number of applications that are enabled by this amount of tuning.

There are some unique device technologies that have been used for temperature tuning. A grating with controllable chirp and center wavelength has been demonstrated with independently controllable thermoelectric coolers on each end of a grating [99,100]. This is quite useful in filtering applications where a filter needs to change dispersion or bandwidth; however the device is relatively energy inefficient. An alternate, more efficient method is to heat the fiber grating with an integrated resistive thin film heater that is deposited on the fiber [101]. This approach has been extended to include multiple films that control the center wavelength with one heater, and the temperature gradient with a second heater that has a resistance gradient induced by tapering the heater thickness [102].

Packaging technologies that can induce strain generally use either piezoelectric actuators [103] or electromagnetic forces [104,105]. Piezoelectric actuators generally have a relatively small range of travel and therefore use mechanical structures to amplify the motion of the actuator. Electromagnetic forces are a very interesting technology for communication systems because of the emergence of programmable magnets that can be latched into a specific configuration [105]. This device requires no power consumption except during reconfiguration, which occurs when the external magnets are reprogrammed by electric pulses through a solenoid. Figure 37 shows a series of different grating spectra tuned in 100 GHz steps by actuating a solenoid to change the magnetization.

A.2. Waveguide Design

Optimization of the waveguiding structure in the grating region is another key to realizing grating components with properties uniquely tailored for communication systems. In general, the variables that can be optimized are the chromatic dispersion of the waveguide modes, the core confinement, the distribution of photosensitive material across the mode, and the addition of new materials to change the thermal or mechanical properties. As discussed with temperature and strain dependence, transmissive gratings are more sensitive to these effects than reflective gratings.

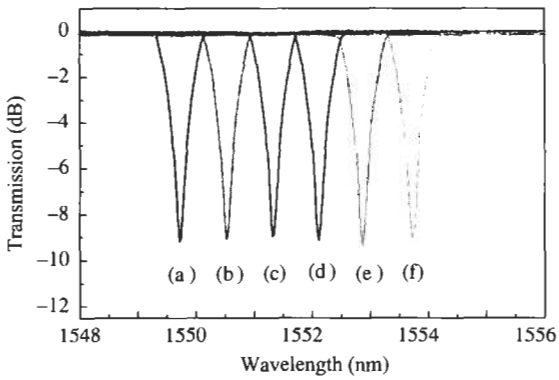


Fig. 37 The spectral transmission of a Bragg grating filter tuned in a latchable magnetic package [105]. Each curve represents 100 GHz wavelength shift between ITU channels.

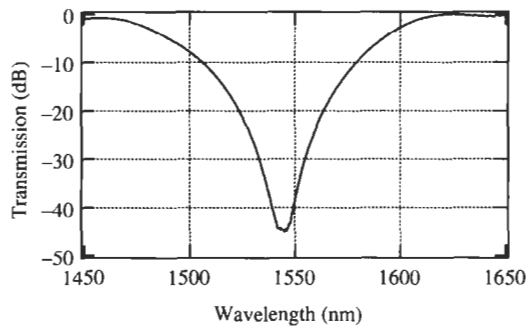


Fig. 38 A broadband long period grating which achieves LP_{01} to LP_{02} spatial mode coupling over 100 nm bandwidth via control of waveguide dispersion [107].

The waveguide refractive index profile of fibers has been optimized for many years to tailor dispersion of the fundamental mode near the 1550 nm communication window. Similarly, the refractive index profiles of fibers can be designed to elicit specific waveguide dispersion not only for the fundamental guided mode, but also for higher-order guided and cladding modes. This approach has been leveraged to design a long-period grating in a waveguide that compensates thermooptic changes with thermal expansion changes, reducing the shift in center wavelength with temperature from 100 pm/°C to 4 pm/°C [106]. Figure 38 shows the transmission spectrum of a long-period grating in a fiber optimized to increase the coupling bandwidth from 1 nm to 50 nm [107].

The design freedom to control the spectral characteristics of a grating via waveguide design is directly related to the refractive index contrast available, although this may be limited when considering the fabrication resolution. Typical low-loss germanosilicate fibers can be fabricated with a maximum

refractive index contrast of <5% (~4% > $\Delta n/n$ > −0.6%). A new class of waveguides described as microstructured fibers have seen significant interest in recent years [108,109], providing even more flexibility than the conventional examples above. These fibers incorporate well-defined regions of air into the waveguiding region, yielding an even more flexible index contrast of 30%. In addition, improvements in fabrication technology have resulted in better control over the size and location of the air regions. While this technology is not yet mature, it has already yielded interesting devices including long-period gratings with internal air cladding that is insensitive to polymer recoating [110], and a temperature-tunable long-period grating with an integrated heater deposited on the fiber surface [102,111].

A final design opportunity to control the characteristics of fiber grating devices is to use hybrid materials in the waveguide. The temperature dependence of a long-period grating with and without a specially designed polymer recoat material that reduced the temperature shift from 48 pm/°C to 4 pm/°C is shown in [112]. It is also possible to wick polymers into the air gaps of microstructured fibers, thereby creating cladding modes that have enhanced temperature sensitivity for tunable gratings [113]. While hybrid material waveguides have yet to be widely used, they will enable a new level of flexibility to implement new devices.

B. FIBER COMMUNICATION SYSTEM DEVICES

Fiber gratings have had a very significant impact on the evolution of fiber communication system technology in recent years. The flexibility that exists to create virtually any desirable transmission or reflection filter in a low-loss device has resulted in widespread use of fiber gratings to realize the latest communication technology advances.

B.1. Transmission Filters

The primary applications in communication systems for filters with wavelength-dependent transmission loss are within optical amplifiers. Filter applications ideally suited for fiber gratings include the removal of amplified spontaneous emission (ASE) noise from the signal fiber, and precise wavelength-dependent gain flattening filtering (GFF) to complement the wavelength-dependent gain intrinsic to optical amplification in EDFAs and Raman fiber amplifiers (RFAs).

a. Gain-Flattening Filters

Fiber gratings have been used extensively in demonstrations of the latest broadband EDFA and RFA technology [114–117], primarily because the alternative technology, dielectric thin film filters, requires a substantial development time to obtain accurate filters.

Gain-flattening filters are typically inserted between two stages of an optical amplifier as shown in Fig. 39, and have a loss shape designed to offset the uneven gain as a function of wavelength for an EDFA or RFA. When the filter is placed between amplification stages, the minimum insertion loss of the device is not critical; however the wavelength dependence of the loss must match the target filter shape very accurately to obtain uniform gain over a bandwidth of 40 nm or more. Three different fiber grating GFF types are discussed below.

The GFF initially used to gain-flatten EDFAs was the long period grating (LPG) [114]. A typical gain flattening filter for a 30+ nm wide C-band amplifier is shown in Fig. 40. The overall complex filter shape can be achieved very accurately by concatenating three or four gratings with a simple Gaussian loss characteristic (shown in dashed lines). Since these gratings have low transmission loss, concatenation does not present a problem, and the extremely

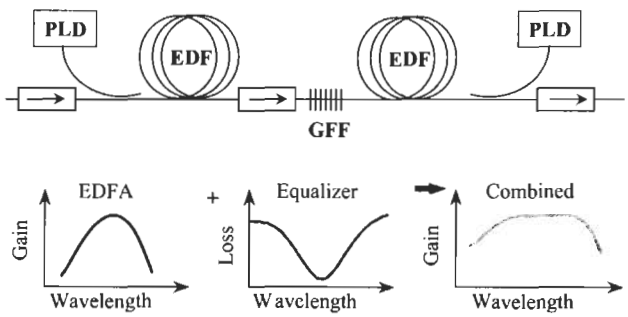


Fig. 39 The positioning of a GFF between two stages in a conventional erbium-doped fiber (EDF) optical amplifier. A wide, uniform gain bandwidth results when the GFF spectral loss complements the intrinsic amplifier gain shape.

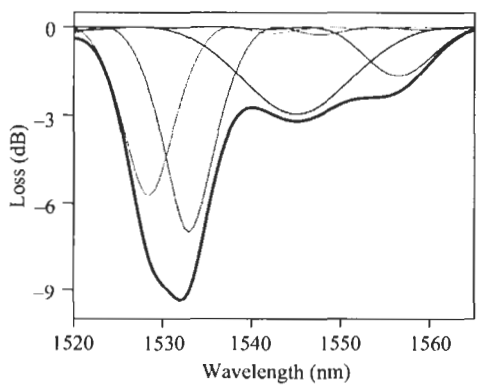


Fig. 40 The complex spectral shape of a GFF filter (thick line) for and EDF amplifier with >30 nm of gain flattened bandwidth. The net filter function is obtained by concatenating four individual low loss LPG filters (thin lines).

low reflection from these gratings (<−80 dB) prevents multiple passes of ASE noise in high-gain amplifiers. This approach has been utilized to gain-flatten wideband amplifiers within 1 dB, including simple two-stage EDFAs with 40 nm of bandwidth [115]. These filters have also been used to realize a more complex three-stage EDFA with a combined 80 nm of bandwidth in the C (1530–1560 nm) and L (1570–1610 nm) amplification bands [116].

An alternate GFF filter technology that is more widely commercially deployed is the chirped Bragg grating (CBG) filter. The exposure along the length of this grating is varied to control the coupling coefficient at the wavelength reflected locally, and therefore the reflected intensity is a function of wavelength. This approach requires an extra isolator to avoid efficiency lost to amplification of (unwanted) reflected light. In addition, radiation mode loss and wavelength-dependent ripple limit the maximum loss at a given wavelength to ~5 dB, which in some applications means an dielectric filter may be needed in addition to the grating filter. The advantage of this filter is a compact footprint and an ability to match complex gain shapes more accurately than any other technology [118].

The final fiber grating GFF technology is an acoustooptic long-period grating that offers dynamic control to adapt to individual amplifier gain shape, including changes with network loading [119]. The previous fixed GFFs only flattened the amplifier gain for one specific Er inversion level, so an attenuator must be used to maintain a constant average inversion via change in average signal power. A dynamic GFF can adjust the filter shape to compensate for input power changes, thereby avoiding a separate attenuator. As shown in Fig. 41, this device couples light from the core to a lossy cladding mode via a long-period grating induced by an acoustic horn [120]. The strength and wavelength of the LPG coupling are determined by the amplitude and frequency of the applied RF signal. Up to three different LPGs filters can be simultaneously induced with one acoustic horn by superimposing multiple RF frequencies on the horn to couple the fundamental mode to different, orthogonal cladding modes [120].

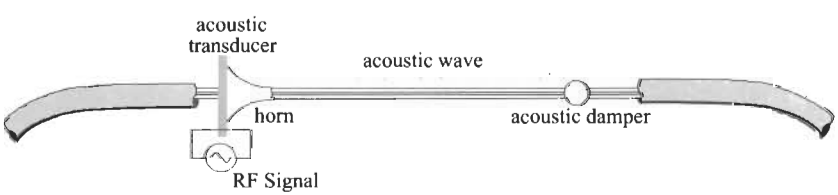


Fig. 41 A dynamic GFF with tunable spectral coupling via a LPG induced by a controllable acoustic wave. Multiple LPGs coupling to different cladding modes can be acoustically encoded in one fiber [120].

b. ASE Filters

Several different grating technologies have been used to remove optical amplifier noise and therefore improve the amplifier efficiency. Both LPGs and CBGs have been used to remove ASE near the 1530 nm EDFA peak [81,121]. Because this type of ASE filter is not typically required in gain-flattened amplifiers, neither grating has seen much commercial use. A more recent demonstration has shown that better efficiency can be obtained by filtering C-band ASE in the first stage of an L-band EDFA that is highly inverted to achieve reduce the amplifier noise figure [122]. Fiber grating filters are uniquely suited for this application because the gain shape of highly inverted erbium requires a very high loss in the C-band but a very low loss for L-band wavelengths that do not see much gain. In this application, tilted Bragg loss filters that couple C-band light into the cladding can be spliced at multiple positions in the first EDFA stage [122]. This approach has demonstrated an L-band EDFA that has an efficiency (25%) and external noise figure (<5 dB) that effectively is the same as would be expected from a C-band amplifier.

B.2. Reflection Filters

Perhaps more than any other characteristic, the reflection properties of fiber Bragg gratings (FBGs) distinguish this technology from other filter technologies. FBGs are unparalleled in their capability to make long, very uniform resonant structures with ten thousand to over one million periods. The result is a technology that can be adapted to make filters with extremely sharp filter characteristics, or large amounts of chromatic dispersion that can be imprinted into a relatively short section of fiber.

a. Fabry–Perot Guiding Filters

A fundamental reflection filter technology is the Fabry–Perot (FP) filter, which can be realized in fiber grating form by fabricating two identical Bragg gratings in a fiber with a small separation. The cavity finesse is determined by the grating reflectivity, because the loss of the gratings and the fiber are very low. In addition, chirped gratings are typically necessary for a broadband device, because unchirped gratings are limited to a few nanometers of bandwidth by the available photoinduced refractive index change. Fiber grating FP devices are useful as periodic wavelength references to align WDM channels, or as soliton guiding filters with periodic loss to remove amplifier ASE noise between channels in long distance soliton transmission [123,124]. One unique capability of a fiber grating FP is the ability to precisely control the reflectivity as a function of wavelength to obtain a well controlled wavelength-dependent finesse as shown in Fig. 42. When such a device is used as a soliton guiding filter, the dispersion variation due to the change in finesse can be used to offset the dispersion slope of the transmission fiber [124].

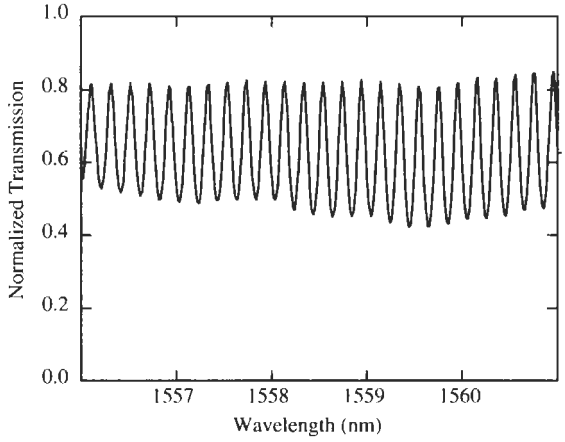


Fig. 42 A fiber grating Fabry–Perot etalon with wavelength-dependent finesse. These devices can be used for soliton filtering, where finesse change compensates the dispersion slope of the fiber [124].

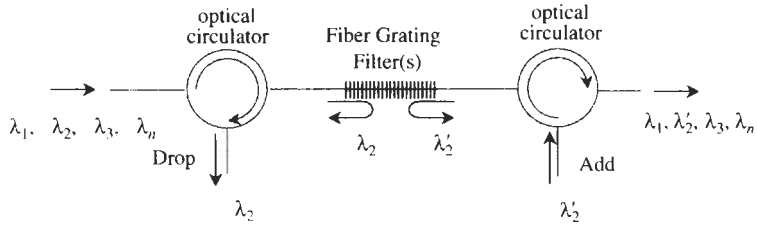


Fig. 43 A fiber grating OADM using a circulator to access a WDM "drop" channel (λ_2) reflected off a fiber grating. An "add" channel (λ_2') is multiplexed at the same wavelength using the same grating, but a second circulator [125].

b. Add/Drop Filters

Perhaps the most demanding filtering application is the requirement for an optical add/drop multiplexer that removes a single wavelength from a group of closely spaced WDM channels [125]. This is typically accomplished as shown in Fig. 43, where a fiber grating reflects a particular wavelength and a circulator is used to redirect the reflected wavelength onto a receiver. At the same time, the same wavelength can be added through another circulator and reflect off the same grating to be multiplexed into the WDM output. The circulators can be replaced with fiber couplers for applications that are less cost-sensitive, and multiple wavelengths can be reflected off the same wide grating or multiple narrowband gratings. The filtering demands of this application are particularly difficult, because the grating cannot transmit more than −30 dB (0.1%) of the dropped signal without a substantial performance penalty on the added

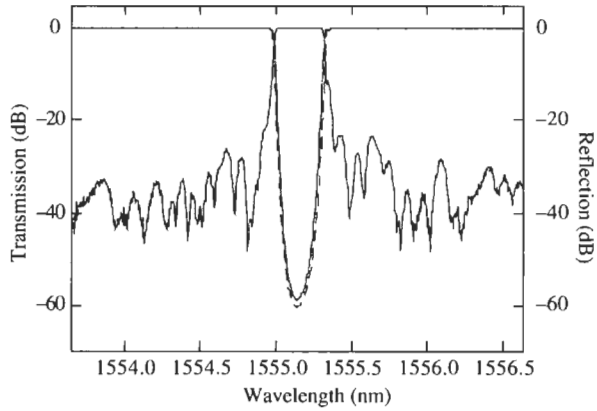


Fig. 44 The measured transmission and reflection characteristics of a fiber grating suitable for OADM at 37 GHz [126]. The usable OADM transmission bandwidth is >70% for a −30 dB transmitted isolation requirement. The solid and dashed transmission curves show the two principal states of polarization, which are nearly identical.

channel due to coherent interference with the drop channel [126]. In addition, it is desirable to reflect <−20 dB (1%) of light from the adjacent channels to prevent crosstalk at the drop channel receiver. The extraordinary filtering characteristics of fiber gratings can be seen in Fig. 44 for a 4 cm long fiber grating that meets these demanding requirements for a 37 GHz channel spacing. Such sharp, narrowband filters are only possible because the grating is very uniform with an interaction length of $>10^4$ periods, resulting in a maximum change in transmission of greater than 700 dB/nm! A frequent requirement for add/drop gratings is to minimize signal loss in channels shorter than the Bragg wavelength by suppression of radiation mode coupling (see Section III I).

While this amplitude filtering performance is unparalleled in other technologies, the phase distortion of such a filter can become a limitation at modulation rates greater than 2.5 Gb/s if the filter is a minimum phase filter [127]. Unfortunately, an unchirped grating in reflection is a minimum phase filter, which means sharp transmission changes result in chromatic dispersion limitations that reduce the usable bandwidth of the filter to less than that defined by the amplitude function [128]. Fortunately, the grating synthesis work in Section III J has shown that it is possible to fabricate a nonminimum phase filter with sharp amplitude characteristics and a relatively low chromatic dispersion (only when incident from one direction). The coupling constant along a linear phase grating and the resulting filtering characteristics are shown in Fig. 45(a) and (b) [129]. As described in Section III F, each pi phase shift represents a change in the sign of the coupling coefficient. These linear phase gratings have more recently been extended to enable 25 GHz channel spacing for 10 Gb signals [130], which would not be possible without the

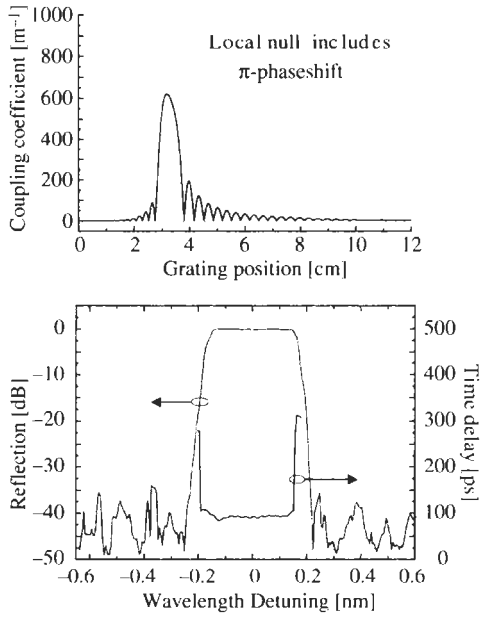


Fig. 45 A linear phase grating for sharp OADM filtering with minimal dispersion limitations at 10 Gb/s [129]. In (a) the grating strength as a function of position, and (b) shows the amplitude and dispersion response of the resulting filter.

uniform group delay characteristics of this technology. While such gratings are not easy to fabricate due to long length (12 cm) and complex index profile, this represents the only filtering technology suitable for the add/drop requirements of current high bitrate, dense channel spacing communication systems.

c. Dispersion Compensation

In WDM optical fiber communication, the maximum distance and bitrate are frequently limited by the chromatic dispersion of the transmission fiber. A light pulse that travels through the fiber is distorted because of a change in wavelength along the pulse arising from the encoding modulator interacting with the chromatic dispersion in the fiber to cause the higher-frequency components to travel slightly more quickly than the lower-frequency components. This degrades the signal integrity as fast components of one pulse overlap with slower components of an earlier pulse. In a linear transmission system, however, it is possible to compensate for the accumulated dispersion by passing the pulses through a device with an equal amount of negative dispersion [131,132]. Almost all dispersion compensation in commercial systems is done with a special dispersion compensation fiber (DCF) designed to a strong negative waveguide dispersion of up to −200 ps/nm. The advantage

of this approach is that the dispersion is broadband and compensates many WDM channels at once with no phase distortion. The disadvantage of DCF is that it takes a long length of fiber to compensate a span, entailing a large amount of space and a relatively large insertion loss. In addition, the DCF can only accept a limited signal power because the small mode field of the fiber and the long propagation length results in nonlinear pulse distortions at high signal intensity. Figure 46 shows an alternate dispersion compensation grating (DCG) technology that compensates for dispersion by reflecting different wavelengths at different positions along a chirped fiber grating [63]. This is an attractive alternative, because large amounts of dispersion can be created with low loss in a small amount of space, and the short device length guarantees no nonlinearity problems at high signal power. In addition, the capability to precisely vary the chirp rate along the grating could compensate for the wavelength dependence of the dispersion (dispersion slope). The disadvantage of a DCG is that the grating length scales linearly with the compensation bandwidth, which makes this a good technology for narrowband compensation but much more difficult for broadband devices which might ideally be 1 to 10 meters long. The advantages of DCGs have been shown in research demonstrations of devices [133–135] and transmission systems [136,137]; however, a number of difficulties have prevented broad commercialization. Perhaps the largest problem is that the dispersive properties of the gratings have small phase distortions due to imperfections in the local period or apodization [138,139]. Tremendous progress has been made in reducing these; however, the scaling of WDM and TDM bandwidths in recent years requires longer gratings *and* less

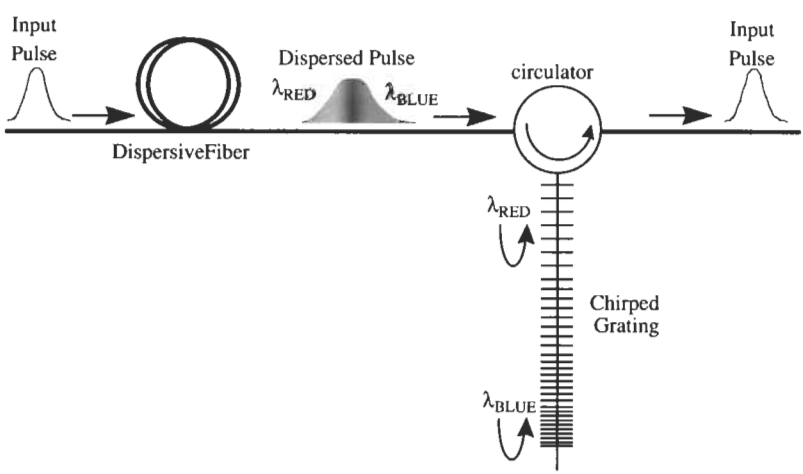


Fig. 46 A DCG reverses the broadening of a pulse that results from transmission through a dispersive fiber with a chirped grating that reflects slow-moving pulse wavelengths (red) before faster-moving ones (blue). The distorted pulse can be faithfully reconstructed in an ideal grating with appropriate chirp and apodization.

phase distortion. To date, these demands have prevented broadband DCGs from achieving commercial success.

A promising alternate application for DCG technology is as a dynamic compensation element to optimize chromatic dispersion of high bitrate signals. This can be accomplished by changing the temperature or strain of a nonlinearly chirped grating [140], or by changing the local chirp rate of a grating with a strain gradient [141] or a temperature gradient [142]. The temperature gradient implementation which chirps Bragg wavelength along the grating length is controlled by current through an integrated heater with a tapered thickness and hence resistance [142]. This arrangement increases the grating chirp rate with increasing heater current, thereby increasing the grating bandwidth and decreasing the group delay as a function of wavelength (dispersion). This device has been used to provide continuously adjustable narrowband compensation of a channel that has tight chromatic dispersion tolerances due to very high bitrate and/or an unpredictable phase distortion from nonlinear propagation at high signal powers. The benefit of optimizing the dispersion can be seen in the eye diagrams in Fig. 47, where increase in signal power before transmission can change 40 Gb/s error-free pulses to completely distorted using DCF [143]. A tunable DCG that can tune by 100 ps/nm can restore pulse distortions caused by nonlinear transmission with almost no power penalty. Although the group delay variations due to grating imperfections remain a problem with this technology, the successful operation of these

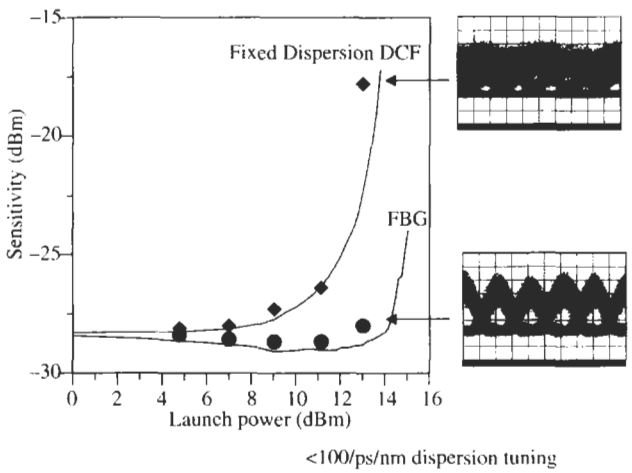


Fig. 47 The receiver sensitivity for a 40 Gb/s signal after high-power nonlinear transmission with optimized DCF, and with a tunable dispersion compensator before the receiver to adjust for differing levels of nonlinear distortion [143]. The eye diagrams shows that the dispersion tuning enables recovery of a nearly closed eye diagram.

devices at bitrates up to 160 Gb/s [144] shows the tunable optimization of the device largely mitigates the impact of these imperfections.

B.3. Outcoupling Devices

A number of applications have arisen that utilize light coupled out of the fiber by a fiber grating as a nonintrusive way to monitor the light passing through. These devices typically use tilted fiber gratings to enhance the efficiency and directivity of light coupled out one side of the fiber. In general these devices have very broad bandwidths ($\gg$10 nm) compared to the reflection bandwidth of a similar untilted grating.

a. Optical Channel Monitoring

One important application is the monitoring of channel power and wavelength for control and fault detection in complex WDM systems with many channels. There are many technologies to accomplish this, including a scanning Fabry–Perot or a bulk grating spectrometer that disperses WDM channels onto a detector array. Figure 48 shows a fiber grating variant that has been commercially successful [145]. This spectrometer uses a tilted fiber grating to diffract a portion of the guided light out of the fiber at an angle that is wavelength-dependent. As shown in Fig. 48, the diffracted light at a given wavelength can be focused in space by decreasing the period along the length of the grating. As a first approximation this focuses different wavelengths at different points in space, at which point a detector array can be positioned to detect the intensity as a function of wavelength. Fiber gratings are used as spectrometers because they are low in cost, reliable, have low polarization dependence, and have angle of incidence guaranteed by the integration of the grating in the fiber. It has also been shown that this type of device is capable of very high resolution

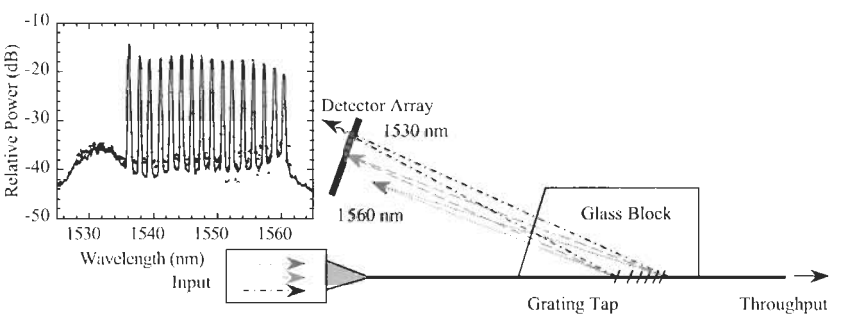


Fig. 48 A fiber grating spectrometer adapted to couple WDM channels from a fiber and focus the channels on individual detector array elements, yielding the intensity as a function of wavelength. The fiber grating operates as a tap, a dispersive grating, and a focusing element (via chirp).

monitoring of WDM channels spaced as closely as 50 GHz, as well as ASE noise floor monitoring between 100 GHz spaced channels [146]. An alternate approach to monitor the WDM spectrum has been demonstrated in the form of a Fourier-transform spectrometer [147]. This device uses the fringe interference pattern from a signal that is incident on the grating from both sides and coupled out the side of the fiber. The interference pattern is unique for each wavelength because the light diffracts out the fiber at different angles. This approach has been shown to have a resolution that is similar to those of conventional spectrometers.

b. Polarization Monitor

As communication technology pushes transmission capacity and distance limits, polarization is becoming an important factor that can be used to improve system performance. Polarization multiplexing of two channels at the same wavelength is used to double transmission capacity, and orthogonal polarization interleaving of adjacent channels is used to reduce interchannel nonlinearities. In addition, higher per-channel bitrates (10–40 Gb/s) may require optical PMD compensation, which typically must control the output polarization after a long transmission span [148]. These applications have generated interest in low-cost, low-loss, polarimeters that can provide deterministic, polarization-controlled output from a communication system [149–151]. A broadband fiber grating version of the device that is inexpensive and compact has been demonstrated as shown in Fig. 49. This device uses a fiber grating tilted at 45° as a polarization sensitive tap that reflects

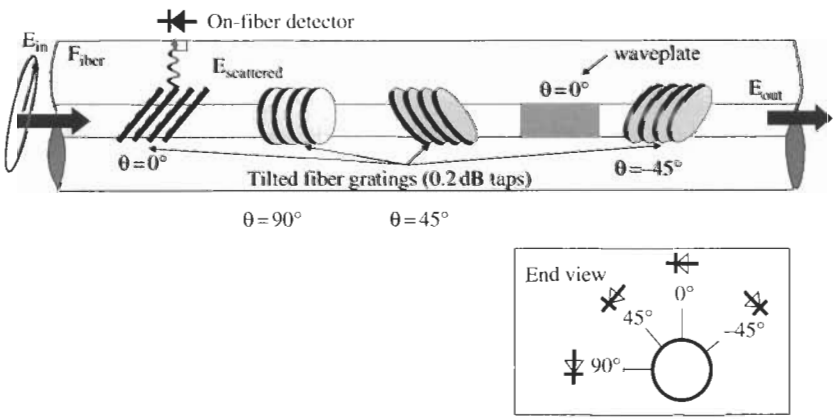


Fig. 49 An all-fiber polarimeter based on fiber grating taps [149]. The grating taps are polarization-selective by tilting the grating at 45° tilt relative to the fiber axis, and UV birefringence is used to make a waveplate between taps. Polarization-dependent coupling from four taps with different azimuthal orientation defines the unique polarization state.

only s-polarized light incident on the grating [76]. The polarimeter functions using a series of four taps with different azimuthal orientations, and a 1 cm in-fiber waveplate photoinduced by the birefringence from a uniform UV exposure [45]. It has been shown that these taps have a polarization selectivity >300 : 1 over a bandwidth in excess of 70 nm, where the Stokes parameters as calculated by the fiber polarimeter agree very well with the results simultaneously measured from a commercially available laboratory polarimeter [149]. The integration of this device with a polarization controller has been used to demonstrate active polarization demultiplexing of two 10 Gb/s channels with no penalty [152], where the two polarizations were encoded with unique RF tones to independently monitor the evolution of the polarization of both signals with the same polarimeter.

B.4. Spatial Mode Conversion Devices

Other devices of interest use fiber or waveguide structures with multiple spatial modes for device functionality. A correctly positioned periodic grating can provide wavelength- and mode-selective coupling between the spatial modes. The most critical aspect of these devices is that the signal must carefully couple only one mode back into the single-mode fiber, because other modes have traveled a different path length and therefore will cause coherent signal interference.

a. Null Fiber Coupler

One very interesting device in this class uses multiple modes in the waist of a special fiber coupler to direct light to two different single-mode outputs. The null coupler is a fused fiber coupler that couples one input fiber through the coupler waist in an even mode and the other input through an odd mode. An unperturbed waist will not have cross-coupling because the even and odd modes are orthogonal; however, wavelength-selective mode conversion can be employed in the waist to route different wavelengths to different outputs. The first device reported was dynamically controlled using an acoustic grating in the coupler waist induced by RF modulation similar to Fig. 41 [153]. This device has also been demonstrated with a photoinduced Bragg grating that is tilted (blazed) to optimize coupling between two fiber modes [154]. In either case, the use of the null coupler is a low-loss approach to couple light reflected by a grating without the cost of a fiberoptic circulator.

b. LP_{02} Mode Dispersion Compensation

An alternate approach to dispersion compensation devices is to ameliorate nonlinear pulse distortion in dispersion compensation fiber (DCF) by propagating light in a higher-order spatial mode, as shown in Fig. 50. This was originally demonstrated using the LP_{11} mode [155], which reduced device length with a higher negative dispersion, while decreasing the intensity in

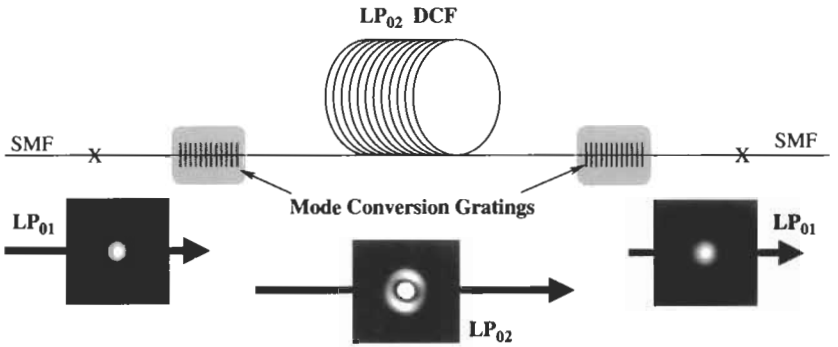


Fig. 50 Fiber grating mode converters enable the use of a DCF that operates in the higher-order LP_{02} mode (mode intensity shown), yielding smaller, lower loss, and less nonlinear DCF modules [157]. These mode conversion gratings can operate over very broad wavelength ranges with very high discrimination.

the fiber by increasing the effective area of the guided mode. The shorter fiber length may also help by reducing size, cost, and loss of a DCF module. The initial implementation of LP_{11} DCF had significant loss because of inefficient mode conversion technology and polarization dependence from coupling between nearly degenerate spatial modes of the asymmetric mode. More recently, it has been shown that DCF propagating in the symmetric LP_{02} mode eliminates most of the polarization dependence [156,157], however, performance to date has been limited by the mode conversion technology. A promising and interesting technology is the use of transmissive LP_{01}/LP_{02} photosensitive gratings to provide broadband, selective-mode conversion. The LP_{01} transmission spectrum of such a mode converter is shown in Fig. 38, where fiber design has been utilized to create greater than 99% (20 dB) mode conversion in excess of 30 nm [107]. Such a technology is critical for higher-order mode DCF to achieve very low loss and very high spatial mode discrimination. These mode conversion gratings have been demonstrated as part of a LP_{02} DCF module with very large effective area (65 μm^2), very high figure of merit (177 ps/nm/dB), and high dispersion slope compensation ratio ($D'/D \sim 0.01\ nm^{-1}$). The difficulty with these approaches is the need for a low-loss mode conversion device, to prevent undesired intermode coupling within the DCF fiber, and couple only the higher-order spatial mode back into a single-mode fiber.

c. *Fiber Lasers*

A number of fiber laser devices use fiber Bragg gratings as resonator reflectors. Fiber gratings are interesting in these applications because they have very low return loss in addition to very good reflectivity, chromatic dispersion, and wavelength stability. These advantages have been exploited to create a number

of high-power, wavelength-accurate, single-frequency fiber laser transmitters for WDM systems, including master oscillator power amplifier (MOPA) configurations [158], distributed feedback (DFB) lasers [68], and hybrid fiber grating-semiconductor sources [73]. The commercial success of these lasers has been limited by the lack of an attractive integrated modulator technology; however, the following pair of high-power fiber laser sources demonstrate the advantages of Raman amplification in communication systems.

d. Raman Lasers and Amplifiers

Fiber gratings enable the fabrication of very high-Q, low-loss laser resonators that can be exploited to improve the efficiency of weak nonlinear processes whose efficiency improves with intensity. Raman scattering is one nonlinear process that is of great interest in communication systems [159]. Raman lasers use gain created by nonlinear Raman scattering, where a photon guided in the fiber scatters a fraction of its energy into a lattice molecule vibration (phonon), with the remaining energy available as gain or spontaneous emission for lower energy photons. The high optical power in a high-Q cavity dramatically increases the efficiency of this process, enabling the use of this process to decrease the wavelength of an input laser by the energy of one or more phonons. One unique aspect of fiber gratings is the capability to nest several high-Q cavities at different wavelengths in the same fiber with multiple gratings at different wavelengths. This approach has been used to cascade multiple phonon shifts within Raman resonators as shown in Fig. 51. The ability to precisely choose the wavelength and the reflectivity of fiber gratings gives precise control over the frequency shifting process and thus the output wavelength. Therefore, when combined with high-power cladding-pumped laser sources around 1.1 μm, the low loss of the fiber and the high reflectivity gratings provide efficient transfer to output at virtually any wavelength from 1.1 to 1.6 μm [160], with single-mode fiber-coupled output power up to almost 10 W [161].

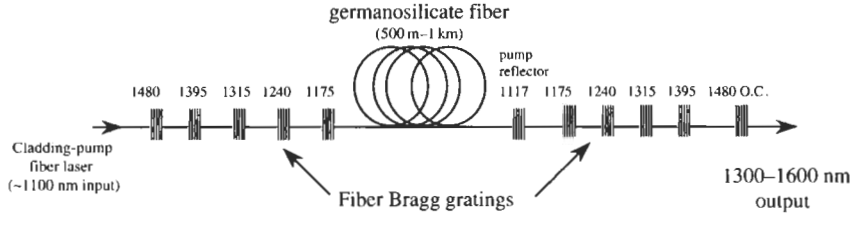


Fig. 51 A cascaded Raman resonator with fiber grating reflectors. The low-loss, high-reflection gratings enable efficient, well-controlled Raman shifting through many phonon shifts, resulting in high-power fiber-coupled output at almost any wavelength from 1.3 to 1.6 μm.

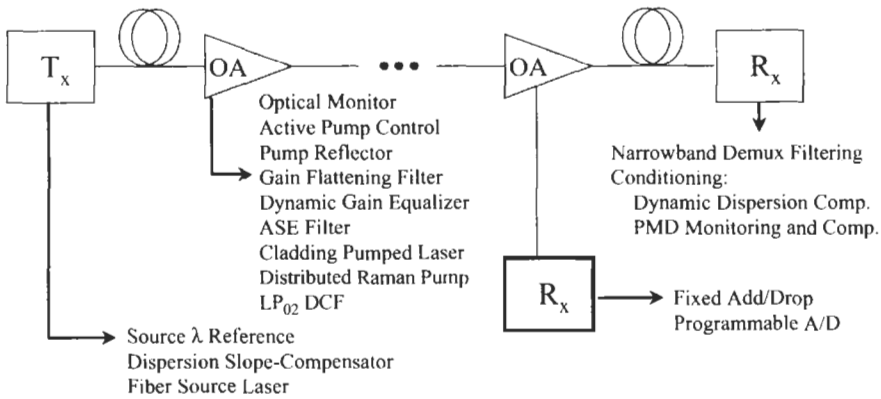


Fig. 52 The devices of fiber grating technology that have impacted almost every critical part of a WDM communication system are shown. In the recent past many of the initial device demonstrations used fiber grating technology.

V. Conclusions

In its short history, fiber grating technology has seen many exciting and significant developments, not all of which could be listed here. The field has matured to the point of being a true manufacturable component technology—fiber gratings are being used to improve the performance and functionality of optical communications systems today. But there is still significant potential for new developments and novel devices. Figure 52 shows some of the anticipated applications for fiber gratings in future optical communications systems.

Fiber gratings can be designed to obtain virtually any desired wavelength-dependent transmission or reflection characteristic with low insertion loss and high reliability—critical characteristics for optical components. Because of the wide range of options available and the incredibly rapid development time, fiber gratings are expected to play a crucial role in the future of fiberoptic systems. Of particular note is the value of fiber grating technology in the research and development phase which by necessity moves even more quickly.

References

[1] K. O. Hill, B. Malo, F. Bilodeau, and D. C. Johnson, "Photosensitivity in optical fibers," *Ann. Rev. Mater. Sci.* **23**, 125–157, 1993.

[2] W. W. Morey, G. A. Ball, and G. Meltz, "Photoinduced Bragg gratings in optical fibers," *Opt. and Photon. News* **5**, 8–14, 1994.

[3] R. J. Campbell and R. Kashyap, "The properties and applications of photosensitive germanosilicate fibre," *Int. J. Optoelectron.* **9**, 33–57, 1994.

[4] P. St. J. Russell, J.-L. Archambault, and L. Reekie, "Fibre gratings," *Physics World*, 41–46, October 1993.

[5] I. Bennion, J. A. R. Williams, L. Zhang, K. Sugden, and N. J. Doran, "UV-written in-fibre Bragg gratings," *Opt. and Quant. Electron.* **28**, 93–135, 1996.

[6] See the Invited Papers in the special issue of *J. Lightwave Tech.* **15**, 1997.

[7] A. Othonos and K. Kalli, *Fiber Bragg Gratings, Fundamentals and Applications in Telecommunications and Sensing.* Boston: Artech House, 1999.

[8] R. Kashyap, *Fiber Bragg Gratings.* San Diego: Academic Press, 1999.

[9] K. O. Hill, Y. Fujii, D. C. Johnson, and B. S. Kawasaki, "Photosensitivity in optical fiber waveguides: application to reflection filter fabrication," *Appl. Phys. Lett.* **32**, 647–649, 1978.

[10] D. K. W. Lam and B. K. Garside, "Characterization of single-mode optical fiber filters," *Appl. Opt.* **20**, 440–445, 1981.

[11] P. J. Lemaire, R. M. Atkins, V. Mizrahi, and W. A. Reed, "High pressure H_2 loading as a technique for achieving ultrahigh UV photosensitivity and thermal sensitivity in GeO_2 doped optical fibres," *Electron Lett.* **29**, 1191–1193, 1993.

[12] R. M. Atkins, P. J. Lemaire, T. Erdogan, and V. Mizrahi, "Mechanisms for enhanced UV photosensitivity via hydrogen loading in germanosilicate glasses," *Electron Lett.* **29**, 1234–1235, 1993.

[13] V. Mizrahi, P. J. Lemaire, T. Erdogan, W. A. Reed, D. J. DiGiovanni, and R. M. Atkins, "Ultraviolet laser fabrication of ultrastrong optical fiber gratings and of germania-doped channel waveguides," *Appl. Phys. Lett.* **63**, 1727–1729, 1993.

[14] K. O. Hill, F. Malo, F. Bilodeau, D. C. Johnson, and J. Albert, "Bragg gratings fabricated in monomode photosensitive optical fiber by UV exposure through a phase mask," *Appl. Phys. Lett.* **62**, 1035–1037, 1993.

[15] D. Z. Anderson, V. Mizrahi, T. Erdogan, and A. E. White, "Production of in-fibre gratings using a diffractive optical element," *Electron Lett.* **31**, 566–568, 1993.

[16] A. E. White and S. G. Grubb, Optical Fiber Components and Devices, Ch. 7, *Optical Fiber Telecommunications IIIB*, ed. by I. Kaminow and T. Koch. Academic Press, Chestnut Hill, Mass., 1997.

[17] M. Douay, W. X. Xie, T. Taunay, P. Bernage, P. Niay, P. Cordier, B. Poumellec, L. Dong, J. F. Bayon, H. Poignant, and E. Delevaque, "Densification involved in the UV-based photosensitivity of silica glasses and optical fibers," *J. Lightwave Tech.* **15**, 1329, 1997.

[18] D. P. Hand and P. St. J. Russell, "Photoinduced refractive-index changes in germanosilicate fibers," *Opt. Lett.* **15**, 102, 1990.

[19] R. M. Atkins, V. Mizrahi, and T. Erdogan, "248 nm induced vacuum UV spectral changes in germanosilicate fibers," *Elect. Lett.* **29**, 385, 1993.

[20] R. J. Campbell and R. Kashyap, "The properties and applications of photosensitive germanosilicate fibre," *Int. J. Optoelectron.* **9**, 33, 1994.

[21] J.-L. Archambault, L. Reekie, and P. St. J. Russell, "100% reflectivity Bragg reflectors produced in optical fibres by single excimer laser pulses," *Electron. Lett.* **29**, 453, 1993.

[22] W. X. Xie, P. Niay, P. Bernage, M. Douay, J. F. Bayon, T. Georges, M. Monerie, and B. Poumellec, "Experimental evidence of two types of photorefractive index effects occurring during photoinscriptions of Bragg gratings within germanosilicate fibers," *Opt. Commun.* **104**, 185, 1993.

[23] J. Albert, B. Malo, K. O. Hill, F. Bilodeau, and D. C. Johnson, "Comparison of one-photon and two-photon effects in the photosensitivity of germanium-doped silica optical fibers exposed to intense ArF excimer pulses," *Appl. Phys. Lett.* **67**, 3529, 1995.

[24] D. S. Starodubov, V. Grubsky, and J. Feinberg, "Bragg grating fabrication of fibers by near-UV light," *Proc. OSA Topical Meeting on Bragg Gratings, Photosensitivity, and Poling in Glass Fibers and Waveguide: Application and Fundamentals* **17**, 151, 1997.

[25] V. Grubsky, D. S. Starodubov, and J. Feinberg, "Wide range and linearity of near-UV induced index change in hydrogen-loaded fibers: Applications for Bragg Grating Fabrication," *Proc. OSA Topical Meeting on Bragg Gratings, Photosensitivity, and Poling in Glass Fibers and Waveguide: Application and Fundamentals* **17**, 156, 1997.

[26] F. Bilodeau, B. Malo, J. Albert, D. C. Johnson, K. O. Hill, Y. Hibino, M. Abe, and M. Kawachi, "Photosensitization of optical fiber and silica-on-silicon/silica waveguides," *Opt. Lett.* **18**, 953, 1993.

[27] H. Hosono and J. Kawazoe, "Preferred concentration enhancement of photobleachable defects responsible for 5 eV optical absorption band in $SiO_2 : GeO_2$ glass perform by heating in H_2 atmosphere," *Appl. Phys. Lett.* **63**, 479–481, 1993.

[28] J. T. Kringlebotn, J. L. Archambault, and D. N. Payne, "Er/sup 3+/:Yb/sup 3+/-codoped fiber distributed-feedback laser," *Opt. Lett.* **19**, 2101–2103, 1994.

[29] B. Malo, J. Albert, F. Bilodeau, T. Kitagawa, D. C. Johnson, K. O. Hill, K. Hattori, Y. Hibino, and S. Gujrathi, "Photosensitivity in phosphorus-doped silica glass and optical waveguides," *Appl. Phys. Lett.* **65**, 394, 1994.

[30] E. M. Dianov, M. V. Grekov, I. A. Bufetov, S. A. Vasiliev, O. I. Medvedkov, V. G. Plotnichenko, V. V. Koltashev, A. V. Belov, M. M. Bubnov, S. L. Semjonov, and A. M. Prokhorov, "CW high power 1.24 μm and 1.48 μm Raman lasers based on low loss phosphosilicate fibre," *Elect. Lett.* **33**, 1542, 1997.

[31] T. A. Strasser, A. E. White, M. F. Yan, P. J. Lemaire, and T. Erdogan, "Strong Bragg phase gratings in phosphorus-doped fiber induced by ArF excimer radiation," Conf. Opt. Fiber Comm., OFC '95, *OSA Tech. Dig.* **6**, 159, 1995.

[32] T. A. Strasser, "Photosensitivity in phosphorus-doped fibers," Conf. Opt. Fiber Comm., OFC '96, *OSA Tech. Dig.* **2**, 81, 1996.

[33] D. L. Williams, B. J. Ainslie, J. R. Armitage, R. Kashyap, and R. Campbell, "Enhanced UV photosensitivity in boron codoped germano silicate fibers," *Electron. Lett.* **29**, 45, 1993.

[34] L. Dong, J. L. Cruz, L. Reekie, M. G. Xu, and D. N. Payne, "Large Photo-induced index changes in Sn-codoped germanosilicate fibers," 1995 OSA Topical Conf. on Bragg Gratings, Photosensitivity, and Poling in Glass Waveguides, Portland, OR, 70, 1995.

[35] L. Dong, P. J. Wells, D. P. Hand, and D. N. Payne, "Photosensitivity in Ce^{3+}-doped optical fibers," *JOSA B* **10**, 89, 1993.

[36] K. O. Hill, F. Bilodeau, B. Malo, and D. C. Johnson, "Birefringent photosensitivity in monomode optical fiber: application to the external writing of rocking filters," *Electron Lett.* **27**, 1548–1550, 1991.

[37] D. C. Psaila, F. Oullette, and C. M. de Sterke, "Characterization of photoinduced birefringence change in optical fiber rocking filters," *Appl. Phys. Lett.* **68**, 900–902, 1996.

[38] M. Parent, J. Bures, S. Lacroix, and J. Lapierre, "Proprietes de polarisation des reflecteurs de Bragg induits par photosensibilite dans les fibres optiques monomodes," *Appl. Opt.* **24**, 354–357, 1985.

[39] P. St. J. Russell and D. P. Hand, "Rocking filter formation in photosensitive high birefringence optical fibres," *Electron Lett.* **26**, 1846–1848, 1990.

[40] F. Ouellette, D. Gagnon, and M. Poirier, "Permanent photoinduced birefringence in a Ge-doped fiber," *Appl. Phys. Lett.* **58**, 1813–1815, 1991.

[41] J. Lauzon, D. Gagnon, S. LaRochelle, A. Blouin, and F. Ouellette, "Dynamic polarization coupling in elliptical-core photosensitive optical fiber," *Opt. Lett.* **17**, 1664–1666, 1992.

[42] V. Mizrahi, D. J. D. DiGiovanni, R. M. Atkins, S. G. Grubb, Y. K. Park, and J. M. P. Delavaux, "Stable single-mode erbium fiber-grating laser for digital communications," *J. Lightwave Tech.* **11**, 2021–2025, 1993.

[43] D. Inniss, et al., "Atomic force microscopy study of UV-induced anisotropy in hydrogen-loaded germanosilicate fibers," *Appl. Phys. Lett.* **65**, 1528–1530, 1994.

[44] T. Meyer, et al., "Reversibility of photoinduced birefringence in ultralow-birefringence fibers," *Opt. Lett.* **21**, 1661–1663, 1996.

[45] T. Erdogan and V. Mizrahi, "Characterization of UV-induced birefringence in photosensitive Ge-doped silica optical fibers," *J. Opt. Soc. Am. B* **11**, 2100–2105, 1994.

[46] T. Erdogan, V. Mizrahi, P. J. Lemaire, and D. Monroe, "Decay of ultraviolet-induced fiber Bragg gratings," *J. Appl. Phys.* **76**, 73–80, 1994.

[47] S. Kannan, J. Z. Y. Guo, and P. J. Lemaire, "Thermal stability analysis of UV-induced fiber Bragg gratings," *IEEE J. Lightwave Tech.* **15**, 1478–1483, 1997.

[48] H. Patrick, et al., "Annealing of Bragg gratings in hydrogen-loaded optical fiber," *J. Appl. Phys.* **78**, 2940–2945, 1995.

[49] A. Asseh, H. Storoy, B. E. Sahlgren, S. Sandgren, and R. A. H. Stubbe, "A writing technique for long fiber Bragg gratings with complex reflectivity profiles," *J. Lightwave Tech.* **15**, 1419–1423, 1997.

[50] W. H. Loh, M. J. Cole, M. N. Zervas, S. Barcelos, and R. I. Laming, "Complex grating structures with uniform phase masks based on the moving fiber-scanning beam technique," *Opt. Lett.* **20**, 2051–2053, 1995.

[51] M. Ibsen, M. K. Durkin, M. J. Cole, and R. I. Laming, "Optimised square pass-band fibre Bragg grating filter with in-band flat group delay response," *Electron Lett.* **34**, 800–802, 1998.

[52] J. F. Brennan III, E. Hernandez, J. A. Valenti, P. G. Sinha, M. R. Matthews, D. E. Elder, G. A. Aeauchesne, and C. H. Byrd, "Dispersion and dispersion-slope correction with a fiber Bragg grating over the full C-band," Paper PD12, Optical Fiber Communications Conf., Anaheim, CA, March 17–22, 2001.

[53] C. M. de Sterke, N. G. R. Broderick, B. J. Eggleton, and M. J. Steel, "Nonlinear optics in fiber gratings," *Opt. and Fiber Tech.* **2**, 253–268, 1996.

[54] T. Erdogan, "Fiber Grating Spectra," *J. Light. Tech.* **15**, 1277–1294, 1997.

[55] D. Marcuse, *Theory of Dielectric Optical Waveguides*, Boston: Academic, 1991, Ch. 2.

[56] M. Born and E. Wolf, *Principles of Optics*, Oxford: Pergamon, 1987, Sec. 8.6.1, Eq. 8.

[57] A. Yariv, "Coupled-Mode Theory for Guided-Wave Optics," *IEEE J. Quantum Electron* **QE-9**, 919–933, 1973.

[58] H. Kogelnik, "Theory of Optical Waveguides," in *Guided-Wave Optoelectronics*, T. Tamir, ed., Berlin: Springer-Verlag, 1990.

[59] J. E. Sipe, L. Poladian, and C. M. de Sterke, "Propagation through nonuniform grating structures," *J. Opt. Soc. Am. A* **11**, 1307–1320, 1994.

[60] H. Kogelnik, "Filter response of nonuniform almost-periodic structures," *Bell Sys. Tech. J.* **55**, 109–126, 1976.

[61] K. O. Hill, "Aperiodic distributed-parameter waveguides for integrated optics," *Appl. Opt.* **13**, 1853–1856, 1974.

[62] B. Malo, S. Theriault, D. C. Johnson, F. Bilodeau, J. Albert, and K. O. Hill, "Apodised in-fibre Bragg grating reflectors photoimprinted using a phase mask," *Electron Lett.* **31**, 223–225, 1995.

[63] F. Ouellette, "Dispersion cancellation using linearly chirped Bragg grating filters in optical waveguides," *Opt. Lett.* **12**, 847–849, 1987.

[64] M. E. Fermann, K. Sugden, and I. Bennion, "High-power soliton fiber laser based on pulse width control with chirped fiber Bragg gratings," *Opt. Lett.* **20**, 172–174, 1995.

[65] P. A. Morton, V. Mizrahi, T. Tanbun-Ek, R. A. Logan, P. J. Lemaire, H. M. Presby, T. Erdogan, S. L. Woodward, J. E. Sipe, M. R. Phillips, A. M. Sergent, and K. W. Wecht, "Stable single mode hybrid laser with high power and narrow linewidth," *Appl. Phys. Lett.* **64**, 2634–2636, 1994.

[66] P. A. Morton, V. Mizrahi, P. A. Andrekson, T. Tanbun-Ek, R. A. Logan, P. Lemaire, D. L. Coblentz, A. M. Sergent, K. W. Wecht, and P. F. Sciortino, Jr., "Mode-locked hybrid soliton pulse source with extremely wide operating frequency range," *IEEE. Photon. Tech. Lett.* **5**, 28–31, 1993.

[67] H. A. Haus and C. V. Shank, "Antisymmetric taper of distributed-feedback lasers," *IEEE J. Quantum Electron* **QE-12**, 532–539, 1976.

[68] W. H. Loh and R. I. Laming, "1.55 μm phase-shifted distributed feedback fibre laser," *Electron Lett.* **31**, 1440–1442, 1995.

[69] B. J. Eggleton, P. A. Krug, L. Poladian, and F. Ouellette, "Long periodic superstructure Bragg gratings in optical fibres," *Electron Lett.* **30**, 1620–1622, 1994.

[70] J. Martin, M. Tetu, C. Latrasse, A. Bellemare, and M. A. Duguay, "Use of a sampled Bragg grating as an in-fiber optical resonator for the realization of a referencing optical frequency scale for WDM communications," presented at the Optical Fiber Communications Conf., Dallas, Texas, Feb. 16–21, 1997, paper ThJ5.

[71] R. Kashyap, H.-G. Froehlich, A. Swanton, and D. J. Armes, "1.3 m long super-step-chirped fibre Bragg grating with a continuous delay of 13.5 ns and bandwidth 10 nm for broadband dispersion compensation," *Electron Lett.* **32**, 1807–1809, 1996.

[72] L. Dong, M. J. Cole, M. Durkin, M. Ibsen, and R. I. Laming, "40 Gbit/s 1.55 μm transmission over 109 km of non-dispersion shifted fibre with long continuously chirped fibre gratings," presented at the Optical Fiber Communications Conf., Dallas, Texas, Feb. 16–21, 1997, paper PD6.

[73] L. A. Weller-Brophy, "Analysis of waveguide gratings: application of Rouard's method," *J. Opt. Soc. Am. A* **2**, 863–871, 1985.

[74] M. Yamada and K. Sakuda, "Analysis of almost-periodic distributed feedback slab waveguides via a fundamental matrix approach," *Appl. Opt.* **26**, 3474–3478, 1987.

[75] V. Mizrahi and J. E. Sipe, "Optical properties of photosensitive fiber phase gratings," *J. Lightwave Tech.* **11**, 1513–1517, 1993.

[76] T. Erdogan and J. E. Sipe, "Tilted fiber phase gratings," *J. Opt. Soc. Am. A* **13**, 296–313, 1996.

[77] R. Kashyap, R. Wyatt, and R. J. Campbell, "Wideband gain flattened erbium fibre amplifier using a photosensitive fibre blazed grating," *Electron Lett.* **29**, 154–156, 1993.

[78] T. A. Strasser, J. R. Pedrazzani, and M. J. Andrejco, "Reflective mode conversion with UV-induced phase gratings in two-mode fiber," presented at the Optical Fiber Communications Conf., Dallas, Texas, Feb. 16–21, 1997, paper FB3.

[79] K. O. Hill, B. Malo, K. A. Vineberg, F. Bilodeau, D. C. Johnson, and I. Skinner, "Efficient mode conversion in telecommunication fibre using externally written gratings," *Electron Lett.* **26**, 1270–1272, 1990.

[80] T. Erdogan, "Cladding-mode resonances in short and long period fiber grating filters," *J. Opt. Soc. Am. A* **14**, 1760–1763, 1997.

[81] A. M. Vengsarkar, P. J. Lemaire, J. B. Judkins, V. Bhatia, T. Erdogan, and J. E. Sipe, "Long-period fiber gratings as band-rejection filters," *J. Lightwave Tech.* **14**, 58–65, 1996.

[82] L. Dong, G. Qi, M. Marro, V. Bhatia, L. L. Hepburn, M. Swan, A. Collier, and D. L. Weidman, "Suppression of cladding mode coupling loss in fiber Bragg gratings," *J. Lightwave Tech.* **18**, 1583–1590, 2000.

[83] Y. Li, M. Froggatt, and T. Erdogan, "Volume current method for analysis of tilted fiber gratings," *J. Lightwave Tech.* **19**, 1580–1591, 2001.

[84] K. A. Winick and J. E. Roman, "Design of corrugated waveguide filters by Fourier transform techniques," *IEEE J. Quantum Electron* **26**, 1918–1929, 1990.

[85] E. Brinkmeyer, "Simple algorithm for reconstructing fiber gratings from reflectometric data," *Opt. Lett.* **20**, 810–812, 1995.

[86] D. W. Huang and C. C. Yang, "Reconstruction of fiber grating refractive index profiles from complex Bragg reflection spectra," *Appl. Opt.* **38**, 4494–4499, 1999.

[87] G. H. Song and S. Y. Shin, "Inverse scattering problem for the coupled-wave equations when the reflection coefficient is a rational function," *Proc. IEEE* **71**, 266–268, 1983.

[88] E. Peral, J. Capmany, and J. Marti, "Iterative solution to the GelFand-Levitan-Marchenko coupled equations and application to synthesis of fiber gratings," *IEEE J. Quantum Electron* **32**, 2078–2084, 1996.

[89] R. Feced, M. N. Zervas, and M. A. Muriel, "An efficient inverse scattering algorithm for the design of nonuniform fiber Bragg gratings," *IEEE J. Quantum Electron* **35**, 1105–1115, 1999.

[90] L. Poladian, "Simple grating synthesis algorithm," *Opt. Lett.* **25**, 787–789, 2000.

[91] J. Skaar, L. Wang, and T. Erdogan, "On the synthesis of fiber Bragg gratings by Layer Peeling," *IEEE J. Quantum Electron* **37**, 165–173, 2001.

[92] R. Feced and M. N. Zervas, "Efficient inverse scattering algorithm for the design of grating-assisted codirectional mode couplers," *J. Opt. Soc. Am. A* **17**, 1573 1582, 2000.

[93] L. Wang and T. Erdogan, "Layer peeling algorithm for reconstruction of long-period fibre gratings," *Electron Lett.* **37**, 154–156, 2001.

[94] K. O. Hill and G. Meltz, "Fiber Grating Technology Overview," *J. Light. Tech.* **15**, 1263, 1997.

[95] V. Bhatia and A. M. Vengsarkar, "Optical fiber long-period grating sensors," *Opt. Lett.* **21**, 693, 1996.

[96] D. L. Weidman, G. H. Beall, K. C. Chyung, G. L. Francis, R. A. Modavis, and R. M. Morena, "A novel negative expansion substrate material for athermalizing fiber Bragg gratings," *Proc. ECOC '96. European Conference on Optical Communication*, Oslo, Norway, 61–64, 1996.

[97] G. W. Yoffe, P. A. Krug, F. Ouellette, and D. A. Thorncraft, "Passive temperature-compensating package for optical fiber gratings," *Appl. Opt.* **34**, 6859, 1995.

[98] G. A. Ball and W. W. Morey, "Compression-tuned single-frequency Bragg grating fiber laser," *Opt. Lett.* **19**, 1979, 1994.

[99] J. Lauzon, S. Thibault, J. Martin, and F. Ouellette, "Implementation and characterization of fiber Bragg gratings linearly chirped by a temperature gradient," *Opt. Lett.* **19**, 2027–2029, 1994.

[100] T. Eftimov, M. C. Farries, S. Huang, N. Duricic, D. Gorbnic, B. Keyworth, and J. S. Obhi, "8-channel tunable drop device with thermal tuning for 100 GHz channel spacing," *European Conference Opt. Comm., ECOC '98* **1**, Madrid, Spain, 127, 1998.

[101] H. G. Limberger, H. K. Nguyen, D. M. Costantini, R. P. Salathe, C. A. P. Muller, and G. R. Fox, "Efficient miniature fiber-optic tunable filter based on intracore Bragg grating and electrically resistive coating," *IEEE Phot. Tech. Lett.* **10**, 361, 1998.

[102] B. J. Eggleton, et al., Conf. on Lasers and Electro-Optics, CLEO '99, paper CMC3, 1999.

[103] L. Quetel, L. Rivoallan, E. Delevaque, H. Poignant, M. Monerie, and T. Georges, "Programmable fiber grating based wavelength demultiplexer," Conf. Opt. Fiber Comm., OFC '96, *OSA Tech. Dig.* **2**, 120, 1996.

[104] J. L. Cruz, A. Diez, M. V. Andres, A. Segura, B. Ortega, and L. Dong, "Fibre Bragg gratings tuned and chirped using magnetic fields," *Elect. Lett.* **33**, 235, 1997.

[105] S. Jin, R. P. Espindola, H. Mavoori, T. A. Strasser, and J. J. DeMarco, "Magnetically programmable fibre Bragg gratings," *Elect. Lett.* **34**, 2158, 1998.

[106] J. B. Judkins, J. R. Pedrazzani, D. J. DiGiovanni, and A. M. Vengsarkar, "Temperature-insensitive long-period fiber gratings," Conf. Opt. Fiber Comm., OFC '96, San Jose, CA, *OSA Tech. Dig.* **2**, 331, 1996.

[107] S. Ramachandran, M. Yan, L. Cowsar, A. Carra, P. Wisk, and R. Huff, "Large bandwidth, highly efficient mode coupling using long-period grating in dispersion tailored fibers," Conf. Opt. Fiber Comm., OFC '01, Anaheim, CA, *OSA Tech. Dig.*, paper MC2, 2001.

[108] P. J. Bennett, T. M. Monro, N. G. R. Broderick, and D. J. Richardson, "Towards a practical holey fibre technology: fabrication, splicing and characterization," European Conference Opt. Comm., ECOC '99 **1**, 20, 1998.

[109] R. S. Windeler, J. L. Wagener, and D. J. DiGiovanni, "Silica-air microstructured fibers: properties and applications," Conf. Opt. Fiber Comm., OFC '99, San Diego, *OSA Tech. Dig.*, paper FG1, 1999.

[110] R. P. Espindola, R. S. Windeler, A. A. Abramov, B. J. Eggleton, T. A. Strasser, and D. J. DiGiovanni, "External refractive index insensitive air-clad long period fibre grating," *Elect. Lett.* **35**, 1–2, 1999.

[111] A. A. Abramov, A. Hale, R. S. Windeler, and T. A. Strasser, "Widely tunable long-period gratings," *Elect. Lett.* **35**, 1, 1999.

[112] A. Abramov, A. Hale, and A. M. Vengsarkar, "Recoated temperature-insensitive long period fiber gratings," *Proc. OSA Topical Meeting on Bragg Gratings, Photosensitivity, and Poling in Glass Fibers and Waveguide: Application and Fundamentals* **17**, paper PDP3, 1997.

[113] P. S. Westbrook, B. J. Eggleton, R. S. Windeler, A. Hale, and T. A. Strasser, "Control of waveguide properties in hybrid polymer-silica microstructured optical fiber gratings," Conf. Opt. Fiber Comm., OFC '00, Baltimore, *OSA Tech. Dig.*, paper ThI3, 2000.

[114] A. M. Vengsarkar, J. R. Pedrazzani, J. Judkins, P. Lemaire, N. Bergano, and C. R. Davidson, "Long-period fiber-grating-based gain equalizers," *Opt. Lett.* **21**, 336–338, 1996.

[115] P. F. Wysocki, J. B. Judkins, R. P. Espindola, M. Andrejco, and A. M. Vengsarkar, "Broad-band erbium-doped fiber amplifier flattened beyond 40 nm using long-period grating filter," *IEEE Phot. Tech. Lett.* **9**, 1343, 1997.

[116] Y. Sun, J. W. Sulhoff, A. K. Srivastava, J. L. Zyskind, T. A. Strasser, J. R. Pedrazzani, C. Wolf, J. Zhou, J. B. Judkins, R. P. Espindola, and A. M. Vengsarkar, "80 nm ultra-wideband erbium-doped silica fibre amplifier," *Elect. Lett.* **33**, 1965, 1997.

[117] T. N. Nielsen, A. J. Stentz, K. Rottwitt, D. S. Vensarkar, L. Hsu, P. B. Hansen, J. H. Park, K. S. Feder, T. A. Strasser, S. Cabot, S. Stulz, C. K. Kan, A. F. Judy, J. Sulhoff, S. Y. Park, L. E. Nelson, and L. Gruner-Nielsen, "3.28 Tb/s (82 × 40 Gb/s) transmission over 3 × 100 km nonzero-dispersion fiber using dual C-band and L-band hybrid Raman/Erbium-doped inline amplifiers," Conf. Opt. Fiber Comm., OFC '00, Baltimore, MD, *OSA Tech. Dig.*, paper PD23, 2000.

[118] M. Rochette, M. Guy, S. LaRochelle, J. Lauzon, et al., "Gain equalization of EDFA's with Bragg gratings," *IEEE Phot. Tech. Lett.* **11**, 536–538, 1999.

[119] S. H. Yun, B. W. Lee, H. K. Kim, and B. Y. Kim, "Dynamic Erbium-doped fiber amplifier with automatic gain flattening," Conf. Opt. Fiber Comm., OFC '99, San Diego, CA; *OSA Tech. Dig.*, paper PD28, 1999.

[120] H. S. Kim, S. H. Yun, I. K. Hwang, and B. Y. Kim, "Single-mode-fiber acousto-optic tunable notch filter with variable spectral profile," Conf. Opt. Fiber Comm., OFC '96, *OSA Tech. Dig.*, paper PD7, 1996.

[121] M. Suyama, T. Terahara, S. Kinoshita, T. Chikama, and M. Takahashi, "Suppression of gain bandwidth narrowing in a 4 channel WDM system using unsaturated EDFAs and a 1.53 um ASE rejection filter," *IEICE Trans. on Comm.* **E77-B**, 449–453, 1994.

[122] R. P. Espindola, P. S. Westbrood, T. A. Strasser, and P. F. Wysocki, "Low noise figure, high gain, high conversion efficiency L-band EDFA," *1999 OSA Conf. Opt. Amp. and Appl.* 132–135, 1999.

[123] S. G. Evangelides, B. M. Nyman, G. T. Harvey, L. F. Mollenauer, P. V. Mamyshev, M. L. Saylors, S. K. Korotky, U. Koren, T. A. Strasser, J. J. Veselka, J. D. Evankow, A. Lucero, J. Nagel, J. Sulhoff, P. C. Corbett, M. A. Mills, and G. A. Ferguson, "Soliton WDM transmission with and without guiding filters," *IEEE Phot. Tech. Lett.* **8**, 1409, 1996.

[124] L. F. Mollenauer, P. V. Mamyshev, and T. A. Strasser, "Guiding filters for massive wavelength division multiplexing in soliton transmission," *Opt. Lett.* **22**, 1621, 1997.

[125] C. R. Giles and V. Mizrahi, "Low-loss add/drop multiplexers for WDM lightwave networks," Proc. IOOC '95, Hong Kong, ThC2-1 1995.

[126] T. A. Strasser, P. J. Chandonnet, J. J. DeMarco, C. E. Soccolich, J. R. Pedrazzani, D. J. DiGiovanni, M. J. Andrejco, and D. S. Shenk, "UV-induced fiber grating OADM devices for efficient bandwidth utilization," Conf. Opt. Fiber Comm., OFC '96, San Jose, CA, *Tech. Dig. Series* **2**, 360, 1996.

[127] G. Lenz, B. J. Eggleton, C. R. Giles, C. K. Madsen, and R. E. Slusher, "Dispersive properties of optical filters for WDM systems," *IEEE J. Quantum Elect.* **34**, 1390, 1998.

[128] G. Nykolak, B. J. Eggleton, G. Lenz, and T. A. Strasser, "Dispersion penalty measurements of narrow fiber Bragg gratings at 10 Gb/s," *IEEE Phot. Tech. Lett.* **10**, 1319, 1998.

[129] M. Ibsen and M. N. Zervas, "99.9% reflectivity dispersion-less square-filter fibre Bragg gratings for high speed DWDM networks," Conf. Opt. Fiber Comm., OFC '00, Baltimore, MD, *OSA Tech. Dig.*, paper PD21, 2000.

[130] M. Ibsen, P. Petropoulos, M. Zervas, and R. Feced, "Dispersion-free fibre Bragg gratings," Conf. Opt. Fiber Comm., OFC '01, Anaheim, CA, *OSA Tech. Dig.*, paper MC1, 2001.

[131] A. H. Gnauck and R. M. Jopson, Dipsersion Compensation for Optical Fiber Systems, Ch. 7, *Optical Fiber Telecommunications IIIA*, ed. by I. Kaminow and T. Koch, Academic Press, Chestnut Hill, Mass., 1997.

[132] L. Garrett, Dispersion Compensation Ch. 7, *Optical Fiber Telecommunications IVB*, ed. by I. Kaminow and T. Li., Academic Press, Boston, 2002.

[133] M. Durkin, M. Ibsen, M. J. Cole, and R. I. Laming, "1 m long continuously-written fibre Bragg gratings for combined second- and third-order dispersion compensation," *Elect. Lett.* **33**, 1891, 1997.

[134] H. Rourke, B. Pugh, S. Kanellopoulos, V. Baker, B. Napier, D. Greene, D. Goodchild, R. Epworth, A. Collar, and C. Rogers, "Fabrication of extremely long fibre gratings by phase matched concatenation of multiple short sections," 1999 OSA Topical Conf. on Bragg Gratings, Photosensitivity, and Poling in Glass Waveguides, Stuart, FL, paper ThD1, 1999.

[135] J. F. Brennan III and D. L. LaBrake, "Realization of >10 m long chirped fiber Bragg Gratings," 1999 OSA Topical Conf. on Bragg Gratings, Photosensitivity, and Poling in Glass Waveguides, Stuart, FL, paper ThD2, 1999.

[136] L. Garrett, A. H. Gnauck, F. Forghieri, V. Gusmeroli, and D. Scarano, "16 × 10 Gb/s WDM transmission over 840 km SMF using eleven broadband chirped fiber gratings," European Conference Opt. Comm., ECOC '98 **1**, 267, 1998.

[137] A. H. Gnauck, et al., "8 × 20 Gb/s, 315 km, 8 × 10 Gb/s, 480 km WDM transmission over conventional fiber using multiple broad-band fiber gratings," *IEEE Photon. Tech. Lett.* **10**, 1495–1497, 1998.

[138] T. A. Strasser, T. Nielson, and J. R. Pedrazzani, "Limitations of dispersion compensation including impact on system performance," *Proc. National Fiber Optic Engineers Conf.* **2** (NFOEC '98), 79, 1998.

[139] C. Bungarzeanu, "System limitations due to group delay and reflectivity ripples of dispersion compensating fibre gratings," Proc. Workshop for Fiber Optic Passive Components (WFOPC '98), 116, 1998.

[140] K.-M. Feng, V. Grubsky, D. S. Starodubov, J.-X. Cai, A. Willnet, and J. Feinberg, "Tunable nonlinearly-chirped fiber Bragg grating for use as a dispersion compensator with a voltage-controlled dispersion," Conf. Opt. Fiber Comm., OFC '98, San Jose, CA, *OSA Tech. Dig.*, paper TuM3, 1998.

[141] M. Ohn, "Tunable fiber gratings and their applications," 1999 OSA Topical Conf. on Bragg Gratings, Photosensitivity, and Poling in Glass Waveguides, paper ThB2, 1999.

[142] B. J. Eggleton, J. A. Rogers, P. S. Westbrook, and T. A. Strasser, "Electrically tunable power efficient dispersion compensating fiber Bragg grating," *IEEE Phot. Tech. Lett.* **11**, 854–856, 1999.

[143] T. N. Nielsen, B. J. Eggleton, J. A. Rogers, P. S. Westbrook, P. B. Hansen, and T. A. Strasser, "Fiber Bragg grating tunable dispersion compensator for dynamic post dispersion optimization at 40 Gbit/s," European Conference Opt. Comm., ECOC '00 **1**, 34–35, 1999.

[144] B. J. Eggleton, B. Mikkelsen, G. Raybon, A. Ahuja, J. A. Rogers, P. S. Westbrook, T. N. Nielsen, S. Stulz, and K. Dreyer, "Tunable dispersion compensation in a 160 Gb/s TDM system by voltage-controlled fiber grating," European Conference Opt. Comm., ECOC '00, Munich, Germany 2000.

[145] J. L. Wagener, T. A. Strasser, J. R. Pedrazzani, J. J. DeMarco, and D. J. DiGiovanni, "Fiber grating optical spectrum analyzer tap," European Conference Opt. Comm., ECOC /IOOC '97 **5**, 65, 1997.

[146] C. Koeppen, J. L. Wagener, T. A. Strasser, and J. J. DeMarco, "High resolution fiber grating optical network monitor," *Proc. National Fiber Optic Engineers Conf.* **2**, 13, 1998.

[147] M. Froggatt and T. Erdogan, "All-fiber wavemeter and Fourier-transform spectrometer," *Opt. Lett.* **24**, 942–944, 1999.

[148] F. Heismann, P. B. Hansen, S. K. Korotky, et al., "Automatic polarization demultiplexer for polarization-multiplexed transmission systems," *Electron Lett.* **29**, 1965–1966, 1993.

[149] P. S. Westbrook, T. A. Strasser, and T. Erdogan, "In-line polarimeter using blazed fiber gratings," *IEEE Photon. Tech. Lett.* **12**, 1352–1354, 2000.

[150] A. Bouzid, et al., "Fiber-optic four-detector polarimeter," *Opt. Comm.* **118**, 329–334, 1995.

[151] B. Scholl, et al., "In-line fiber optic polarimeter with a 99% coupler," *Opt. Engin.* **34**, 1669–1672, 1996.

[152] P. S. Westbrook, L. E. Nelson, T. Erdogan, and T. A. Strasser, "In-fiber polarimeter demonstrated in active polarization demultiplexing at 10 Gbit/s," European Conference Opt. Comm., ECOC '00, Munich, Germany 2000.

[153] D. O. Culverhouse, S. H. Yun, D. J. Richardson, T. A. Birks, S. G. Farwell, and P. S. J. Russell, "Low-loss all-fiber acousto-optic tunable filter, *Opt. Lett.* **22**, 96, 1997.

[154] A. S. Kewitsch, G. A. Rakuljic, P. A. Willems, and A. Yariv, "All-fiber zero-insertion-loss add-drop filter for wavelength-division multiplexing," *Opt. Lett.* **23**, 106, 1998.

[155] C. D. Poole, J. M. Wiesenfeld, D. J. DiGiovanni, and A. M. Vengsarkar, "Optical fiber-based dispersion compensation using higher order modes near cutoff," *IEEE J. Light. Tech.* **12**, 1746–1758, 1994.

[156] A. H. Gnauck, L. D. Garrett, Y. Danzinger, U. Levy, and M. Tur, "Dispersion and dispersion slope compensation of NZDSF for 40-Gb/s operation over the entire C-band," Conf. Opt. Fiber Comm., OFC '00, Baltimore, MD, *OSA Tech. Dig.*, paper PD8, 2000.

[157] S. Ramachandran, B. Mikkelsen, L. C. Cowsar, M. F. Yan, G. Raybon, L. Boivin, M. Fishteyn, W. A. Reed, P. Wisk, and D. Brownlow, "All-fiber, grating-based, higher order mode dispersion compensator for broadband compensation and 1000 km transmission at 40 Gb/s," Postdeadline paper 2.5, European Conference on Optical Communications, Munich, September 2000.

[158] V. Mizrahi, D. J. DiGiovanni, R. M. Atkins, S. G. Grubb, Y. K. Park, and J.-M. P Delavaux, "Stable single-mode erbium fiber-grating laser for digital communication," *J. Light. Tech.* **11**, 2021, 1993; G. A. Ball, C. E. Holton, and G. Hull-Allen, "60 mW 1.5 μm single-frequency low-noise fiber laser MOPA," *IEEE Phot. Tech. Lett.* **6**, 192, 1994.

[159] A. Stentz and K. Rottwitt, Raman Amplifiers, Ch. 5, *Optical Fiber Telecommunications IVA*, ed. by I. Kaminow and T. Li, Academic Press, Boston, Mass., 2002.

[160] S. G. Grubb, T. A. Strasser, W. Y. Cheung, W. A. Reed, V. Mizrahi, T. Erdogan, P. J. Lemaire, A. M. Vengsarkar, and D. J. DiGiovanni, "High-power 1.48 mm cascaded Raman laser in germanosilicate fibers," *Proc. OSA Topical Meeting Optical Amplifiers and their Applications* (OAA '95), 197, 1995.

[161] G. R. Jacobovitz-Veselka, R. P. Espindola, C. Headley, A. J. Stentz, S. Kosinski, D. Inniss, D. Tipton, D. J. DiGiovanni, M. J. Andrejco, J. J. DeMarco, C. Soccolich, S. Cabot, N. Conti, J. J. Veselka, L. Gruner-Nielsen, T. A. Strasser, J. R. Pedrazzani, A. Hale, K. S. Kranz, R. G. Huff, G. Nykolak, P. Hansen, D. Boggavarapu, X. He, D. Caffey, S. Gupta, S. Srinivasan, R. Pleak, K. McEven, and A. Patel, "A 5.5-W single-stage single-pumped erbium doped fiber amplifier at 1550 nm," *Proc. OSA Topical Meeting Optical Amplifiers and their Applications* **16**, 148, 1997.

Chapter 11 | Pump Laser Diodes

Berthold E. Schmidt, Stefan Mohrdiek, and Christoph S. Harder

Nortel Networks Optical Components, Zürich, Switzerland

History and Background

The bedrock of today's commercially available erbium-doped optical fiber amplifiers is its pump laser diodes, which deliver the raw power to regenerate optical signals along the network. This chapter will concentrate on pump laser diodes for coherent amplification of signals within a wavelength range around 1550 nm through conversion of either 980 nm or 1480 nm single lateral mode laser light. Pump laser diodes in optical communication systems must have the characteristics of a generic power supply. They should have a high conversion efficiency to generate a minimum amount of heat and be highly reliable, even though the power and current densities are pushed to their physical limits.

An important milestone in optical networking occurred in 1993 when MCI announced a successful fiber link between Sacramento, California and Chicago, Illinois using erbium-doped fiber amplifiers (EDFA) powered by 980 nm pump lasers from Zürich. Previously, only 1480 nm pump lasers based on the mature InGaAsP material system were in service, but these systems were limited in both bandwidth and range. Other service providers quickly followed suit and 980 nm pump demand burgeoned, soon eclipsing 1480 nm in market presence. The availability of reliable, high-power 980 nm pump lasers has played a major role in enabling the commercialization of high-bandwidth dense wavelength division multiplexing (DWDM) systems.

The shift to DWDM systems capable of transmitting many hundreds of Gb/s over sixty-four or more 2.5 Gb/s channels or up to forty 10 Gb/s channels has radically affected EDFA design and the specifications of the pump laser diodes. DWDM has also changed the status quo between the two pump laser wavelength regimes, causing a resurgence of 1480 nm pump laser diode demand, mainly for C-band power boosters and L-band power sources.

Resurrected in 1999 as a viable commercial technology, Raman amplification is seen to be the key to solve system reach and capacity demands in long-haul optical networks. Raman amplification through already installed fibers, which form the gain medium, acts preferentially as a low-noise preamplifier in addition to the standard EDFAs. Enablers of that technology are single-mode pump sources between 1400 nm and 1500 nm delivering at least 200 mW of optical power in the fiber. Initially, lasers based on ytterbium-doped

OPTICAL FIBER TELECOMMUNICATIONS.
VOLUME IVA

ISBN: 0-12-395172-0

fibers pumped by broad area laser diodes around 975 nm with subsequent Raman cascades [1] have been used. Such fiber laser can output powers up to many watts, but they are very bulky and dissipate lots of heat and are thus limited to a niche in the telecom market. Among other high-power devices are wide-flared-stripe semiconductor laser structures [2, 3], but they need complex lens coupling schemes, and it is difficult to predict beam stability over the full life. Recent developments of narrow-stripe single-mode laser diodes with optical light output power exceeding 500 mW [4–6], excellent reliability, small space requirements and relatively small heat loads have made such narrow-stripe lasers the most cost-effective devices for telecom applications. Raman amplification became viable for widespread deployment due to these recent developments in semiconductor pump laser diodes.

As optical networking architectures continue to evolve, further impact in pump laser development is foreseeable. The quest is continuing for pump laser diodes with ever-higher optical pump powers as well as increased overall power conversion efficiency. Recently, pump lasers that are stable enough to operate in an uncooled low-cost package have opened up the opportunity for truly low-cost erbium-doped optical amplifiers [7].

980 nm TECHNOLOGY

Most of today's high-power semiconductor laser diodes with light emission in the 980 nm wavelength regime are based on the AlGaAs/InGaAs material system [8, 9]. The excellent lattice match, refractive index contrast, dopability, and thermal conductivity of AlGaAs alloys give designers freedom to optimize the vertical structure. A strained pseudomorphic InGaAs/GaAs single- or multiquantum well (QW) active region produces high gain, good electrical confinement and therefore low threshold current and high quantum efficiency. Molecular beam epitaxy (MBE) and metal-organic vapor phase epitaxy (MOVPE) are both mature production techniques for the growth of high-quality AlGaAs laser diodes.

A key part of any reliable AlGaAs-based laser technology is the mirror passivation. If the cleaved mirror is not properly protected, degradation related to the oxidation of the active indium gallium arsenide layer eventually leads to catastrophic optical mirror damage (COMD). The presence of aluminum in the cladding can aggravate the issue, but it is not the true cause for COMD in 980 nm pump lasers. Today we understand the COMD process to be the succession of a slow facet degradation by a chemical oxidation process followed by a very rapid thermal optoelectronic runaway. During the oxidation phase, nonradiative recombination centers are formed in the active layer at the facet due to oxygen-induced decomposition of GaInAs. Electrons and holes which are created via absorption of the laser beam, as well as carriers which are injected into the active layer, recombine via these nonradiative recombination centers, thus heating up the facet, further accelerating the oxidation process.

As soon as the facet temperature exceeds a critical temperature, it enters a phase of very sudden thermal optoelectronic runaway that drives the facet up to melting point, destroying the integrity of the device. The process for the facet to degrade enough to reach the critical temperature can take from minutes to years, depending on the optical power level, as the COMD level depends on the exposure time [10–12]. Today there are two known ways to avoid COMD in InGaAs/AlGaAs lasers: the oxidation is suppressed by chemically passivating the surface; or the bandgap at the facet is increased enough to avoid injection of electrons and holes as well as to suppress the absorption of the laser beam in the oxidized and decomposed surface layer. During the last few years, aluminum-free InGaAsP alloys grown on GaAs substrates that cover almost the same wavelength ranges as AlGaAs have been extensively investigated. A significant number of aging data presented by Pessa, et al. [13] on such lasers without facet passivation did not reveal sudden fails up to quite high power levels. In addition, the hybrid form InGaAsP/AlGaAs/InGaAs has also been tested successfully [14]. However, bandgap engineering with these material systems is complex or less flexible than with AlGaAs, thus limiting the freedom of choice for laser diode optimization somewhat. In consequence, these lasers are still awaiting the commercial breakthrough on a large scale.

Beam Characteristics

The variety of 980 nm single-mode pump laser diodes incorporates the whole family of narrow-stripe-geometry waveguide structures. However, the commonly used expression "single-mode" is misleading, since most pump laser diodes in the field of telecommunication are Fabry-Perot laser diodes with fundamental mode operation only for the lateral and vertical beam. Compared to single-mode distributed feedback (DFB) laser diodes, the longitudinal mode spectra of Fabry-Perot pump lasers are multimode spectra which can be externally stabilized by a fiber Bragg grating (FBG). Nevertheless, such FBG lasers still emit into a multiple of longitudinal modes. Thus, pump laser diodes are only single-mode with respect to their spatial characteristics. The maximum extension of the near field pattern is of the order of the core diameter of single-mode fibers and amounts to 0.2–1 μm in the vertical direction and 3–6 μm in the lateral direction. The resulting far field pattern of the device can be basically derived by Fourier transformation of the optical near field [15]. A schematic sketch of the beam divergence is given in Fig. 1.

To achieve maximum coupling efficiency into a single-mode fiber, a laser beam is desirable that can be matched to the fiber with a simple, low-cost, efficient lensing scheme. It is more complex and therefore expensive to match a highly divergent beam. Thus, the FWHM of the vertical far field has been reduced from around 45° to values below 20° during the last few years [16, 17], while lateral far field angles are around 5–9° under operating conditions. It is advantageous to have a round beam for use of simple low-cost symmetrical

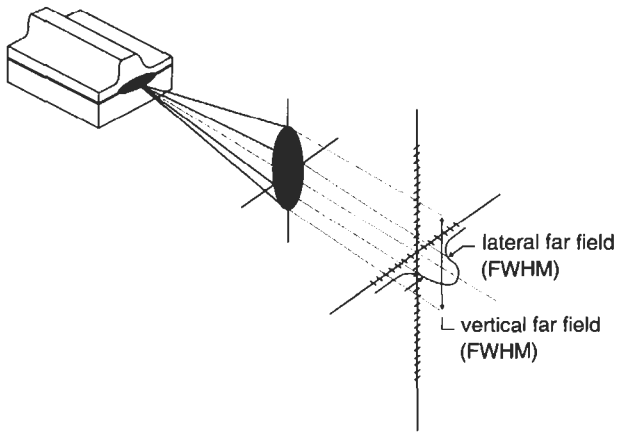


Fig. 1 Schematic far field emission pattern of a ridge waveguide laser diode. The full width at half maximum (FWHM) angles characterize the beam divergence in vertical and lateral direction (after Casey and Panish [15]).

coupling lenses (discrete or fiber tip lenses). In the last few years we have witnessed the success of a special low-cost cylindrical coupling lens, the wedged fiber tip lens. This lens accommodates beams with lateral far fields of around 8° (depending on fiber) and wide-range vertical far fields below 30°. Today laser beams are especially designed to match this wedged fiber tip lens. A large variety of possible lensing schemes for coupling the laser to the fiber are additionally available; the cost (including alignment tolerance) and the coupling efficiency ultimately decide which will be used for high-volume pump laser manufacturing.

The commonest waveguide pattern for AlGaAs-based quasi-index guided high-power laser diodes is the ridge waveguide design, as shown in Fig. 2 [18]. This successful design principle has been proven before for GaAs as well as for InP-based material systems [19, 20].

Alternative approaches to achieve single-mode beam divergence and similar lateral optical and electrical confinement are planar design structures such as buried-stripe or self-aligned waveguide structures [8, 21], channeled-substrate planar-stripe geometry lasers [22] and disordered-stripe-geometry laser diodes [23–25]. Two examples for planar approaches are shown in Fig. 3.

So-called "junction side down" or "p-side down" mounted planar design structures, where the active region is located closer to the heatsink, are used because of improved thermal properties. Due to oxidation of Al-containing material the overgrowth of AlGaAs layers faces significant challenges. Ion implantation technology is a more complex design approach, due to the risk of diffusion of doping material and a thermally induced structural alteration of the QW region during the adjacent annealing step.

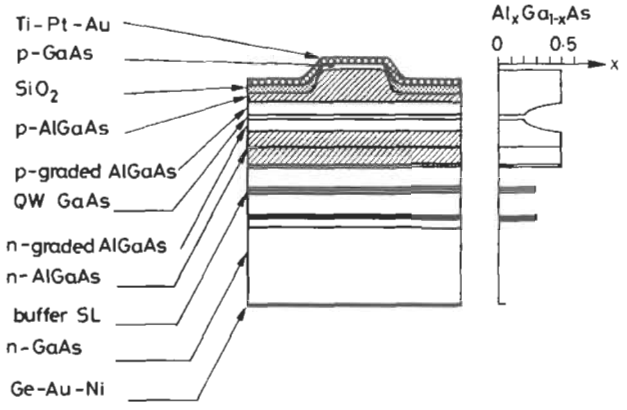


Fig. 2 Schematic cross-section of a ridge waveguide laser diode together with the vertical profile of aluminum concentration (after Harder et al. [18]). (Reprinted with permission.)

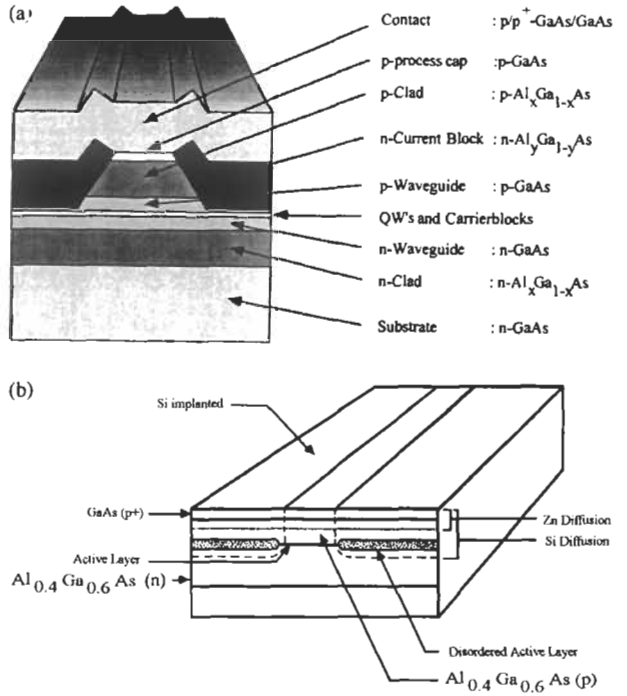


Fig. 3 (a) Cross-sectional view of an InGaAs/AlGaAs buried-ridge waveguide laser structure with regrown cladding layers [8], reprinted with permission from International Society for Optical Engineering; and (b) device schematics of a planar buried heterostructure laser diode fabricated by silicon-implanted impurity-induced disordering [25]. (Reprinted with permission.)

Planar structures additionally can yield increased leakage current around the active region, as compared to ridge waveguide laser diodes, further reducing power conversion efficiency as injection currents become very high. Nevertheless, technological development is in strong progress within the field of pump laser diodes. Thus, the question concerning the optimum design for maximum, highly reliable light output power cannot yet be answered.

Laser Design and Waveguide Characteristics

Since most of the presented design approaches have similar optical properties, the quasi-index-guided ridge waveguide structure will be used to describe the main characteristics of stripe-geometry laser diodes in the 980 nm regime.

The active region of a vertical waveguide structure is normally an intentionally lattice-mismatched InGaAs layer sandwiched between AlGaAs or GaAs barrier material. The $In_xGa_{1-x}As$ layer has a larger native lattice constant with respect to GaAs and is therefore compressively strained, which results in a splitting of light- and heavy-hole bands combined with a reduction of the effective masses due to bending of the QW subbands [26, 27]. In addition to a reduced threshold current, other electrooptical properties (e.g., power conversion efficiency, rollover power) at increased injection currents and temperatures improve due to the strained-QW active region [27]. Within the so-called critical thickness calculated by Matthews's law [28], the QW layer can be grown with high crystalline quality on a GaAs substrate. The epitaxial structure can be further optimized using graded-injection and carrier confinement (GRICC) layers in order to achieve high internal quantum efficiency and high characteristic temperature T_0 [29]. A controlled AlGaAs cladding design together with strained-layer quantum well region and the balanced use of doping material are important to provide low threshold current and operation with high power conversion efficiency.

The trend toward longer semiconductor laser diodes with cavity length above 1 mm is obvious within the field of pump lasers. The reduction of electrical and thermal resistance with increasing chip length leads to a reduced temperature rise of the active layer at high injection currents, enabling reliable high light output power operation. The length of the cavity is mainly limited by the internal cavity loss a_i, which has to be kept small with an optimized vertical profile.

The lateral waveguiding provided by the ridge structure is weak as compared to the vertical, where the index step is defined by growth of epitaxial layers with high refractive index contrast. In the lateral direction, the waveguide is mainly determined by selectively removing the semiconductor material down to a defined level. After additional processing steps, the result is an effective refractive index step of around 2 to 5×10^{-3}, with the waveguide embedded in an insulating dielectric (e.g., SiN_x). Thermal effects modify this waveguiding, i.e., by local Joule heating as well as by local absorption of the beam. As

described by Amann et al. [30], real and imaginary parts of the refractive index pattern are also influenced by injection of free carriers into the QW region via plasma effect, band-filling, and spatial hole burning.

The mirrors of Fabry-Perot-type laser diodes are formed by cleaving semiconductor material along parallel crystal planes perpendicular to the cavity axis; thus, an equal amount of light is emitted from both sides of the resonator. For maximum fiber-coupled light output power, asymmetric facet coatings composed of dielectric mirrors are used. To optimize the light output power conversion efficiency from the front mirror, a low-reflective coating with a reflectivity of only a few percent is deposited on the front, while the rear mirror is coated with a high-reflectivity (90–95%) dielectric stack. The light emitted from the rear side is used via a backmonitor photodiode to control the laser diode front facet output. The influence of an asymmetric coating on the laser diode is shown in Fig. 4, where the optical intensity increases towards the front facet of the laser cavity. As a result, the gain profile within the active region is modified due to inhomogeneous spatial hole burning [31]. Therefore, the lateral waveguide pattern exhibits varying strength in the longitudinal direction with increasing injection currents. Thus, a careful waveguide design is important to guarantee a continuous single-mode light output power performance within the whole operating regime.

As mentioned above, an optimized near and far field design is important for achieving high coupling efficiencies into single-mode fibers. For the fabrication of high-power modules, a narrow and almost circular output beam would be favorable for efficient fiber coupling through circular lenses. Due to the strong vertical index guiding, the pump laser beam has higher divergence perpendicular to the layers of the epitaxial structure, which can be controlled

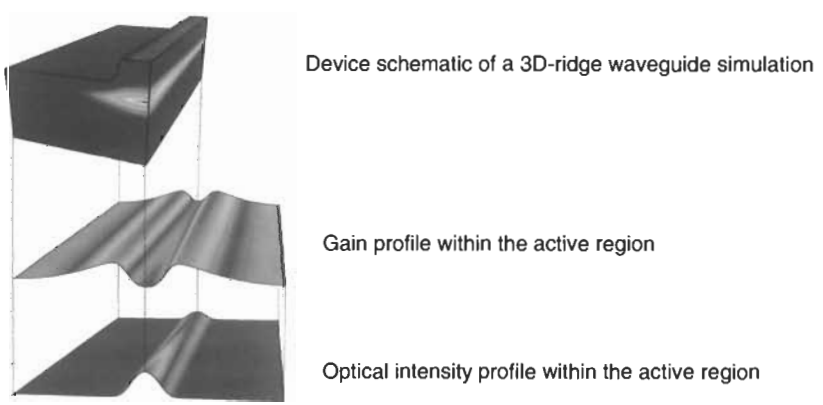


Fig. 4 Three-dimensional simulation of a ridge waveguide laser structure incorporating an asymmetric facet coating. Together with the intensity profile along the active region, the gain profile influenced by lateral and longitudinal spatial hole burning is shown. See also Plate 5.

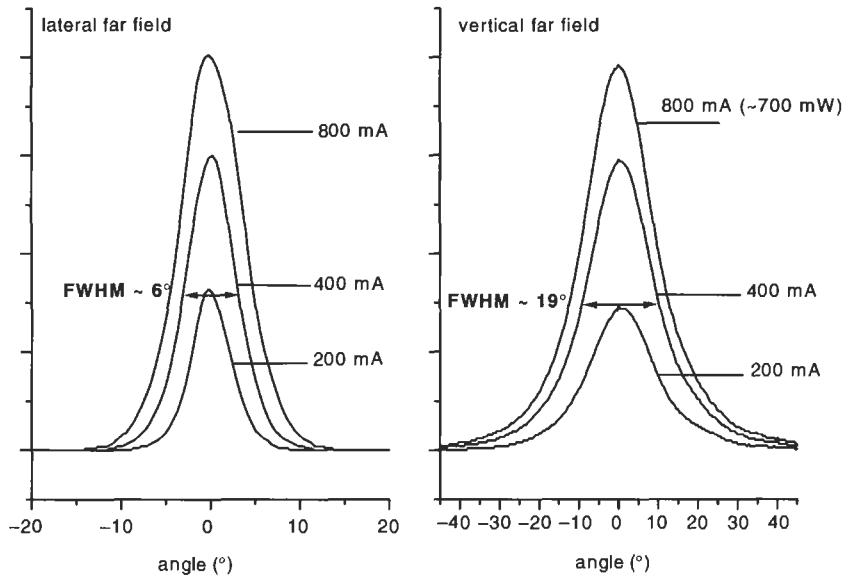


Fig. 5 Lateral and vertical far field pattern of a G06 laser diode (Nortel, Switzerland) for various injection currents at 25°C. The FWHM was measured at an injection current of $I = 400\,\text{mA}$ injection current or approximately 360 mW light-output power. (Reprinted with permission.)

by the composition of the AlGaAs layer and the thickness of GRICC-single quantum well (SQW) active region. It was recently demonstrated that the use of AlGaAs/InGaAs-based GRICC-SQW structures allows the design of 980 nm pump laser diodes with a reduced vertical far field still with low threshold current and high quantum efficiency operation [9]. In Fig. 5 the beam characteristics at various injection currents are shown. The lateral far field spread at 400 mA amounts to 6°, while the vertical far field angel is about 19°.

Light Output Power Characteristics

Light output power as a function of injection current is the most important characteristics of the pump laser diode. Depending on the external differential quantum efficiency η_d, the light output power increases continuously with growing injection current according to

$$\mathrm{d}P_{Opt} = \eta_d \frac{h\nu}{q} \mathrm{d}(I - I_{th})$$

where $\mathrm{d}P_{Opt}$ describes the variation in optical light output power, h is Planck's constant, ν the frequency of the light, q the elementary charge, I the injected current, and I_{th} the laser current at threshold. The relation is valid only for small injection currents. The maximum light output power is limited due to

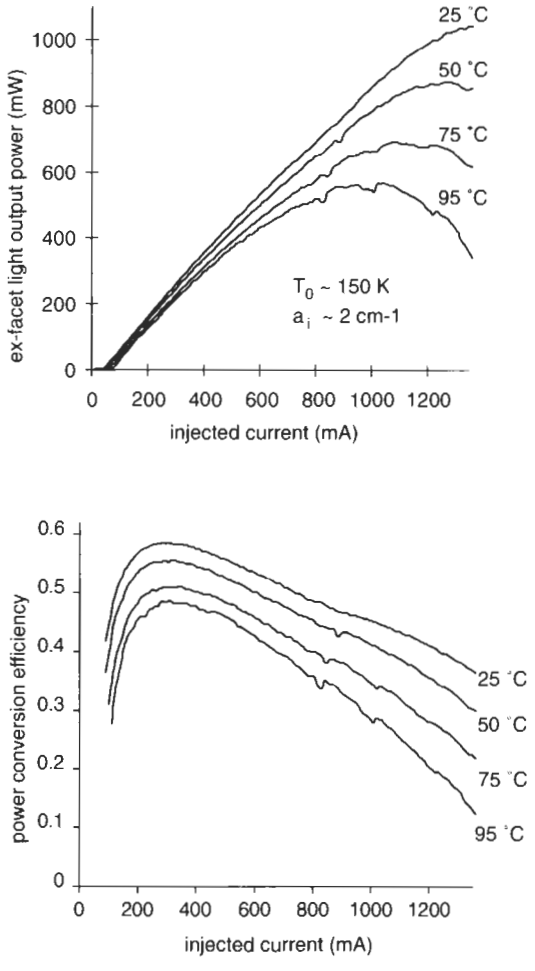


Fig. 6 Ex-facet light output power characteristics and power conversion efficiency versus cw-injection current of a G06 laser diode for various heat sink temperatures. (Courtesy of Nortel Networks Optical Components, Zürich, Switzerland.)

thermal rollover (Fig. 6) which is common to all semiconductor laser diodes. In this case the temperature of the junction increases mostly due to Joule heating, leading to an increased leakage of carriers from the active region into the cladding. Further heating occurs at high currents due to a decreasing power conversion efficiency. For buried heterostructure-like structures additional lateral carrier leakage arises via the surrounding regions. The thermal rollover has no impact on the device, in contrast to the COMD at high power levels (as described above) which causes an irreversible change in the output power-current characteristics of the laser diode.

Another important parameter for laser diodes is the power conversion efficiency, as shown in Fig. 6, describing the ratio between emitted optical light output power and the electrical input power. In the high injection current regime the power conversion efficiency slowly decreases due to increasing current leakage in the device. A good measure for the characterization of leakage currents in laser diodes is the temperature dependence of the laser threshold current at various operating conditions. Normally, the characteristic temperature T_0 is introduced to describe the change of the threshold current with increasing temperature, according to the empirical model

$$I_{th}(T + \Delta T) = I_{th}(T)e^{(\Delta T/T_0)}$$

where $I_{th}(T+\Delta T)$ is the threshold current at elevated temperature in relation to the threshold current $I_{th}(T)$ at temperature T [32]. Increasing the barrier height around the active quantum well by increasing the aluminum concentration will increase the T_0. While a high T_0 is desirable, it is also desirable not to exceed the aluminum concentration in the laser and thus the T_0 of practical lasers is always a compromise between these factors. An important limiting factor for the maximum light output power, especially for long cavity lasers, is the internal cavity loss a_1. It can be estimated by plotting the inverse of the differential quantum efficiency η_d as a function of the inverse cavity length of normally uncoated laser diodes with equal front and back facet reflectivities [15]. The vertical profile of a pump laser is carefully optimized with respect to aluminum and doping profile to reduce this internal cavity loss.

The most relevant nonlinearity in the PI characteristic is caused by excitation of higher-order modes and coherent coupling from the zero order mode to higher-order lateral modes within the waveguide. This effect, known as "kink," causes a change of the near field patterns affecting the gain profile and therefore the light output power characteristics. Most often the kink in the PI curve is accompanied by a shift of the far field maximum in the lateral direction. Since slight deviations in beam divergence lead to a strong decrease of the optical power coupled into the fiber, the fiber-coupled laser diode can only be usefully operated below the kink.

1480 nm TECHNOLOGY

Long-wavelength (1200–1600 nm) telecommunications laser diodes are based on quaternary and ternary alloys of InGaAsP material system. Laser heterostructures can be deposited by various techniques, such as MOVPE (commonly used today), MBE, and liquid phase epitaxy (LPE) onto InP substrates. Multiple quantum wells (MQW) and barrier made form InGaAsP form the active region, which is surrounded in the vertical direction by optical and electrical confinement layers. All layers must be grown either lattice-matched or pseudomorphically strained in relation to the InP lattice constant. The

different atomic radii of In/Ga and As/P restrict the device design to some extent.

Similarly to the laser design described for 980 nm pumps, optimization of the vertical structure of 1480 nm laser diodes follows in principle the same improvement rules. Compressive strain is applied in the active region to inhibit Auger losses and suppress intervalence band absorption [33]. The characteristics of a graded-index separate confinement heterostructure (GRIN-SCH) have been also investigated in detail [34]. A careful choice of barrier composition can reduce losses due to carrier overflow into the SCH layer from the QW. Lateral optical and electrical confinement can be achieved by a wet chemical ridge etch and subsequent regrowth embedding the ridge by either liquid phase epitaxy (LPE) or MOVPE (Fig. 7a). Lateral current flow bypassing the active region can be blocked by either a sandwich structure of p- and n-doped layers forming reverse-biased junctions (MOVPE and LPE) or insulating InP:Fe layers (MOVPE).

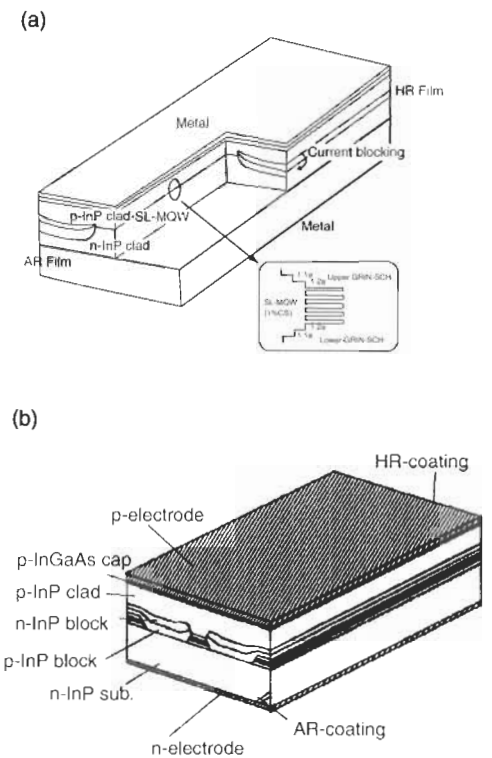


Fig. 7 Two buried heterostructure 1480 nm laser diode structures with different ridge forming techniques (a) regrowth of wet-etched ridge [36] and (b) all selective MOVPE (ASM) growth [5]. (Reprinted with permission.)

Alternatively, a selective MOVPE technique without any semiconductor etching process has been presented (Fig. 7b) to grow the ridge directly and, in a next step, the current blocking layers [35]. With this, the dimensions of the buried heterostructure can be controlled more accurately.

Optimization of laser structures for high light output power with respect to heat dissipation yields relatively long cavity lengths of up to and beyond 2 mm with facet reflectivities of about 5% and 95% for front and rear facets, respectively. The relatively high drive current and corresponding power consumption can pose limits for practical applications. Therefore, shorter cavity lengths to be used in standard-size Butterfly-type packages are also under consideration.

Figure 8a demonstrates the light output power versus injection current characteristics for state-of-the-art 1480 nm laser diodes showing a maximum of 500 mW light output power from the front facet for a 1.5 mm long device at 25°C under constant current (CW) operation. Figure 8b reveals the far-field pattern of these devices, perpendicular and parallel to the junction plane at light output power of 100 mW, with FWHM angles of 25° and 20° respectively [37]. The buried heterostructure design allows for almost symmetric far fields. High light output power has also been demonstrated with laser diodes using very large cavity lengths of 2–3.5 mm [4, 5].

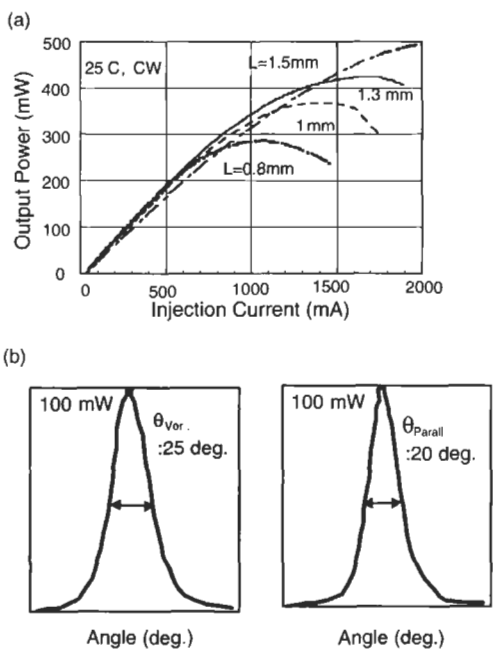


Fig. 8 1480 nm laser: (a) Light output power versus injection current for various cavity lengths and (b) vertical and lateral far field pattern at 100 mW [37, 38]. (Reprinted with permission.)

Performance Comparison Between 980 nm and 1480 nm Technology for EDFA Pumping

The question of whether ridge waveguide or buried-heterostructure-like design approaches would be preferred for a certain wavelength range is complicated by various design aspects and material properties. The two systems, InAlGaAs and InGaAsP, have significant differences, e.g., the mobility of free carriers, possible span of bandgap energies, and the thermal conductivity. In addition, the processing properties deviate from each other, e.g., InGaAsP structures are superior concerning wet-etch selectivity, epitaxial regrowth, and the realization of strain compensated multiquantum well regions. Furthermore, the wavelength dependence on internal absorption loss due to plasma effect and intervalence band absorption plays another important role for the design decision.

The design and fabrication process for most of the commercial pump laser diodes depends on the technological skill and experience developed over the years. Both 980 nm and 1480 nm pump laser diodes require advanced processes, but in different stages of device fabrication. Processing of 980 nm laser diodes does not require complex epitaxial regrowth steps, but the AlGaAs mirror passivation adds significant cost to the device. The implementation of a telecom-qualified manufacturing process for 980 nm pump laser diodes aiming at the highest reliability standards remains a challenge.

The buried-heterostructure 1480 nm laser design involves sophisticated epitaxial growth (for low absorption), regrowth, and excellent contacts (for low power dissipation). The high thermal load together with temperature sensitivity and reduced InGaAsP thermal conductivity requires junction-side-down mounting for proper device operation. InGaAsP laser mirrors require no processing beyond standard reflectivity modification, because the COMD at the facet is not critical. In both 1480 nm and 980 nm, front and back facets of laser diodes have low and high reflectivity coatings, respectively, to allow efficient extraction of power.

The performance of the amplifiers pumped by the two wavelengths varies; 1480 nm introduces more amplification noise than three-level 980 nm pumping, since it is a two-level pumping process and lies close to the 1550 nm signal band. Amplification noise is especially harmful to DWDM systems having narrow 100 GHz (or 50 GHz) channel spacing, and also for long-haul systems which cascade many EDFAs. So 980 nm lasers are strongly preferred for amplification of low-power signals over long distances. Due to the better noise performance and better electrical-to-optical power conversion in the device, almost all new WDM submerged systems are pumped solely by 980 nm laser diodes.

On the other hand, 1480 nm lasers provide higher optical conversion efficiency. Converting a 980 nm photon (1.265 eV) to 1550 nm is ~50% less power-efficient than using a 1480 nm photon (0.838 eV). The lower optical

efficiency of 980 nm pumping is offset at the systems level by better device efficiency due to the superior physical properties of AlGaInAs over InGaAsP for high-power laser operation.

Reliability of Pump Lasers

In addition to optoelectronic performance, the device reliability at operating power is of high interest. The general perception has been that 1480 nm lasers are more reliable than 980 nm devices. Early on, AlGaAs lasers failed frequently, suddenly, and unpredictably, due to COMD and dark-line defect propagation. However, the suppression of dark lines through optimized crystal growth, and the development of mirror passivation procedures have eliminated the major sudden fail mechanisms [39]. InGaAsP lasers tend not to fail suddenly, but degrade slowly and predictably over time. They show higher resistance to dark line defect growth and propagation, along with a higher COMD limit due to reduced surface recombination centers at the laser facets [40].

The reliability standard is the failure rate, measured in FIT (failure-in-time) units. One FIT equals to one single failure per 10^9 device hours. Thus, 1000 FIT is equivalent to about 1% of a device population failing per year. EDFA designers adopt and integrate the laser chip reliability numbers into their own reliability budgets (typically 1000 FIT for the whole EDFA module) to determine how hard the pump laser diode can be driven.

The principal problem concerning the reliability models of semiconductor laser diodes is the difficulty of generating an accurate acceleration model for failures at various power levels. In order to obtain good reliability data, laser diodes are tested under stress conditions, elevated in temperature and injection current as compared to the desired operating condition. In the case of 980 nm laser diodes, an insignificant wearout is seldom observed, and an average lifetime of more than 100 years under operating conditions is predicted. Usually the failure rate is assumed to be accelerated exponentially with temperature according to the Arrhenius law

$$\text{failure rate} \sim e^{-E_a/kT}$$

where E_a is the activation energy in semiconductor devices. To adapt the formula for high-power laser diodes, the current density J and power density P are included in the reliability model. The failure rate at a certain junction temperature (T_j), light output power, and injection current is assumed to have the following functional dependence:

$$\text{failure rate} = A \cdot J^x \cdot P^y \cdot e^{-E_a/kT_j}$$

The relevant parameters A, x, y, and E_a are deduced by the maximum likelihood fit [41] to a matrix of lifetest data obtained by running a few sets of

lasers at different accelerating conditions, i.e., at different injection currents and temperatures. Within the qualification of every new device generation, these parameters have to be determined by a complex experiment to test the stringent reliability requirements for telecom systems. For example, a typical qualification of 980 nm pump laser devices includes stress tests of up to 1000 laser diodes for more than 2 million cumulative device-hours. Lot and individual screening of the devices is used to improve the quality of the shipped parts, which is expressed by a lower prefactor A. However, care must be taken when using the reliability model to predict the failure rate at various operating conditions. Different laser design approaches cause different interpretations; the validity is restricted to a certain operating regime and the error margins are usually quite large, complicating a direct product comparison.

Figure 9 shows the operating current over 10 years' continuous operation for the first generation of 980 nm lasers with E2-facet passivation [9, 39]. The current is adjusted to maintain a constant diode light output power, also known as automatic power control (APC) mode. The noise in the data is not due to laser instability but rather tester instability caused by power outages, temperature variations, and moves into new locations during the last ten years. More than 0.7 million device hours have been accumulated at various stress conditions. Based on the standard reliability model and data for these 9 first-generation devices, a single fail enables a 100 FIT failure rate prediction at operating conditions (130 mW, 25°C). Nowadays, 980 nm pump laser diodes show lower failure rates at much higher power levels [9].

Wearout failure rate of 1480 nm devices is obtained by similar stress tests with subsequent extrapolation to the mean time to failure (MTTF) at operating

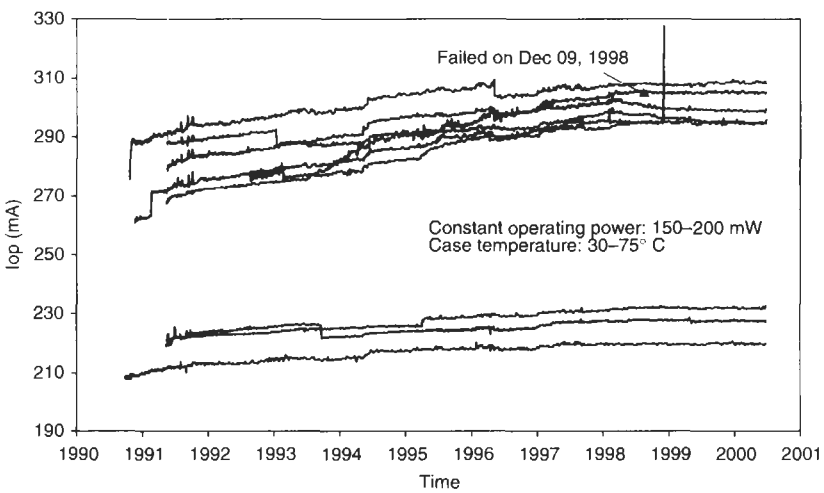


Fig. 9 Nine first-generation E2-lasers under test for 10 years at various stress conditions.

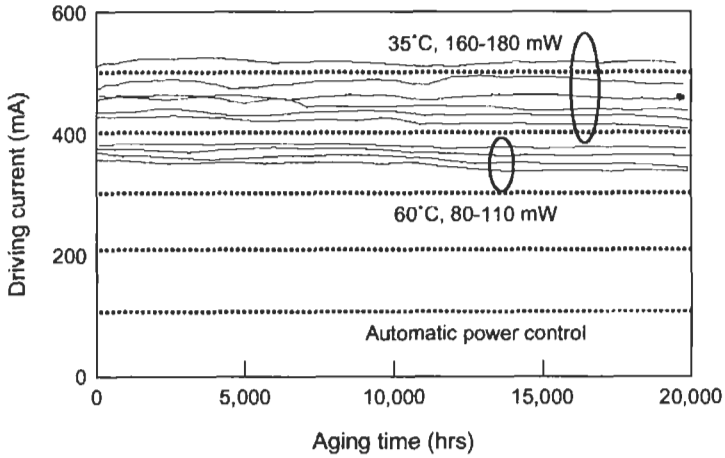


Fig. 10 Reliability test results of 1480 nm lasers [37]. (Reprinted with permisssion.)

conditions. An example for reliability tests of commercially available 1480 nm laser diodes is given in Fig. 10. Aging tests at elevated temperatures of 35°C and 60°C have been performed on a number of devices. The output power is set to be 80% of the maximum output power at a given temperature, determined by the thermal rollover, which is 180 mW at 35°C and 120 mW at 60°C for this chip generation. A MTTF of 100 million hours for laser output powers of 150 mW at 25°C is obtained, by derating with activation energies derived from various aging conditions [37].

Wavelength and Power Stabilization

At the 980 nm pump wavelength, the EDFA absorption spectrum is sharply peaked. Wavelength shifts by a few nanometers of the pump laser diode can result in a large change in EDFA gain. Moreover, optical power fluctuations of the laser diode can also cause variations in EDFA gain. Fluctuations arise from mode-hopping between several longitudinal modes of Fabry-Perot pump laser diodes.

Typical temperature conditions in undersea cables range from 0 through 40°C. This causes an unacceptable wavelength shift of 12 nm due to the natural bandgap narrowing with a shift rate of 0.3 nm/°C of a 980 nm pump laser diode. In addition, varying the laser output power over several hundred mW causes a wavelength shift on the order of a couple of nanometers, depending on the thermal resistance between laser chip and heatsink. Increasing functionality of EDFAs also imposes tighter specifications on pump laser diode wavelength stability. For example, recent systems drive laser diodes depending on data traffic. In undersea systems it is necessary to eliminate the extra power needed for laser cooling. Further, wavelength-stabilized sources enable EDFA

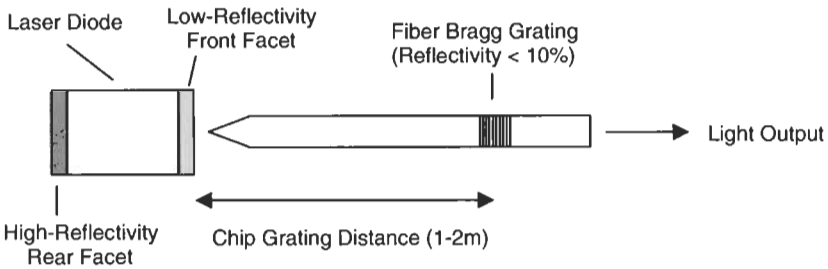


Fig. 11 Schematic of pump laser diode stabilization by FBGs.

designers to multiplex several pump sources within a single EDFA to provide more pump power. This has led to the widespread introduction of fiber Bragg gratings (FBG), which lock the pump laser wavelength over a wide operating range [42, 43]. The FBG, usually written into the fiber pigtail outside the laser package, locks the laser wavelength by reflecting light within a narrow wavelength band back into the chip (Fig. 11).

A small amount of back-reflected power on the order of a few percent is enough to sufficiently reduce the threshold gain at the spectral peak of the FBG to pull the laser toward operating at the FBG wavelength. The center wavelength of the fiber grating has to be sufficiently close to the laser gain peak. If the wavelength offset (detuning) between the free-running laser and FBG is too large, then the laser diode will unlock from the grating and emit at its free-running wavelength. A higher grating reflectivity increases the wavelength range in which the laser is locked to the grating (locking range), at the expense of decreased light output power.

Emission spectra are presented in Fig. 12 for a 980 nm laser diode as an example of operation over a wide locking range [7]. The laser diode in a Butterfly package is fiber-coupled and wavelength-stabilized by a FBG. The ambient temperature for the FBG and laser diode is varied over a large temperature range. Most of the power is concentrated within the vicinity of the wavelength range defined by the FBG. The out-of-band power suppression exceeds 20 dB for the two extreme operating conditions at 10°C and 100°C. A characteristic feature is the broad side peak in the spectrum that corresponds roughly to the spectral position of the gain maximum of the laser. Its magnitude increases with larger detuning. The wavelength offset between the operation at the two temperatures is as large as 35 nm. The inset in Fig. 12 shows the shift of the light output wavelength of only 0.007 nm/K, determined by the shift of the FBG's peak wavelength with temperature.

Important for pump laser diode stabilization by fiber Bragg grating is positioning the grating far enough from the laser's front facet, beyond the laser's coherence length, at a distance of 1–2 m. Incoherent reflection from the FBG is usually preferred, since coherent phase matching would impose tight requirements for chip-to-fiber coupling tolerances. Due to the external cavity of 1–2 m

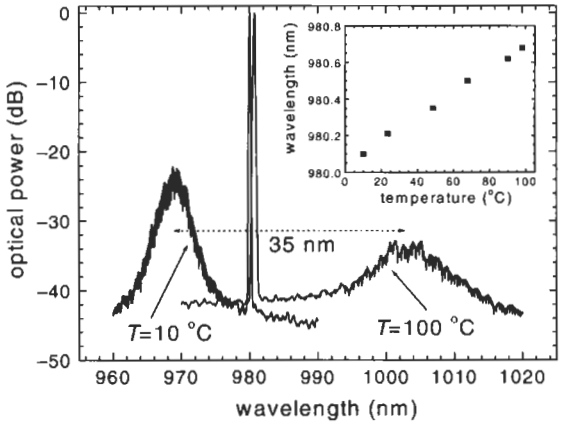


Fig. 12 Emission spectra of the FBG-stabilized laser at the temperatures of 10 and 100°C and output power of 250 mW and 95 mW, respectively. Inset: peak emission wavelength as function of temperature. (Reprinted with permission.)

length, noise is introduced at the corresponding roundtrip frequencies of 50–100 MHz. This is of no concern, because of the slow response time of erbium ions in glass. In order to assure incoherent operation, the FBG bandwidth is chosen to be sufficiently broad, so that several longitudinal laser modes oscillate simultaneously within the spectral width defined by the FBG [43, 44]. In this case, mode-hopping noise is effectively eliminated. It has been observed that modules at 980 nm can display power fluctuations as a function of current, attributable to switching between a coherent single-mode and an incoherent multimode state of operation, and that these effects can be avoided with a proper match of laser diode and FBG parameters [45].

This phenomenon is less critical for pump laser diodes at 1480 nm, since here the broader gain spectrum and higher losses favor multimode operation. Wavelength shift in 1480 nm pump diodes is of less concern, because the erbium absorption band is much broader and increases monotonically at this wavelength. The broader absorption band also enables multiplexing of more 1480 nm pump lasers into a single fiber. In case of Raman amplification, wavelength stabilization is needed for 14*xx* nm laser diodes, because of the strong dependence of gain spectrum on pump wavelength. Multiplexing of FBG-stabilized pump lasers enables system designers to tailor Raman amplifiers for a flatter and broader overall gain spectrum.

With respect to noise, the situation is different for Raman systems, where amplification is an extremely fast process. Here, a pump laser operating in the aforementioned state of incoherent operation can be used only in a counterpropagating scheme, where the fluctuations are averaged out. Copropagating Raman applications require pump lasers with wavelength stabilizations that have low relative intensity noise (RIN) figures comparable to the signal to be

amplified. Low RIN can be provided by DFB or external cavity lasers. DFB lasers stabilize the wavelength by incorporating an on-chip Bragg grating. Such lasers are manufactured with a multistep growth process. The wavelength shift for DFB pump lasers diodes amounts to 0.08 nm/K [46].

Specific Aspects of Pump Laser Packaging

Among the technologies for laser packaging, the butterfly-type package is usually used for pump laser diodes (Fig. 13). Stringent telecom reliability requirements can be met, while at the same time allowing for high chip-to-fiber coupling efficiencies. The cost and effort associated with the development and qualification of a telecom package is often underestimated.

A failure mode related to laser modules is the package-induced failure (PIF) or hermetic enclosure phenomenon (HEP) [47, 48]. The laser facet may become contaminated with hydrocarbon by the omnipresent organic contamination inside the package. Pump laser modules addressing highest reliability avoid the use of epoxy and use fluxless soldering, with parts baked at high temperatures. Residual trace amounts of organics can cause PIF, but a trace of oxygen gas prevents the deposition of hydrocarbons by removing carbon through oxidation.

In order to maintain stable operating conditions, Peltier elements inside the package stabilize the heat sink temperature against ambient temperature changes. The size of a module and the Peltier cooler therein limits the

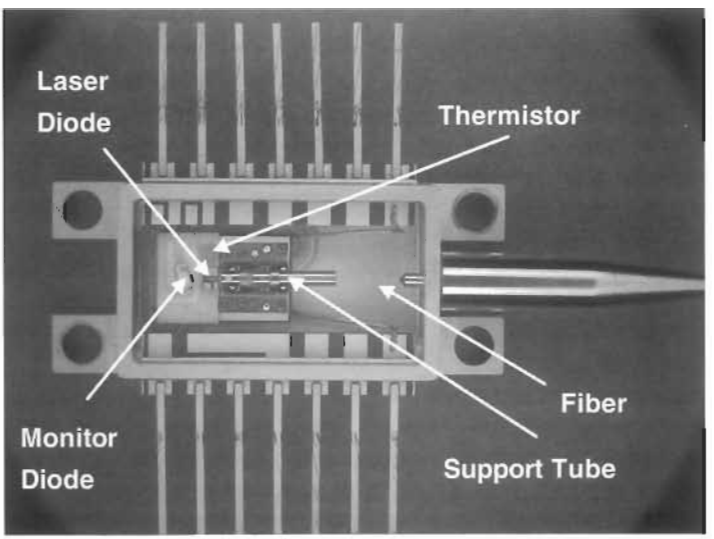


Fig. 13 Butterfly package with housing area dimensions (without pins) of $13 \times 30\,mm^2$. (Courtesy of Nortel Networks Optical Components, Zürich, Switzerland.) See also Plate 6.

maximum power dissipation of the laser diode. Inside the module the entire electrical power P_{El} that is not turned into optical power P_{Opt} is turned into heat $P_{El} - P_{Opt}$.

$$\text{Electrical input power} \quad P_{El} = VI = I(V_g + R_s I)$$

$$\text{Heating} \quad \Delta T = R_{th}(P_{El} - P_{Opt})$$

The light output power at a given current drops with increasing temperature, thus increasing even more the heatload at elevated temperatures. In general, and especially for high injection currents, most power is dissipated as Joule heating caused by the laser series resistance R_s, followed by heating due to lower electrical-to-optical power conversion efficiencies.

The 980 nm pump laser module illustrated in Fig. 13 [49, 50] consists of an optical subassembly (OSA baseplate with AlO_x circuit board) mounted onto a thermoelectric cooler in a 14-pin butterfly housing. On the OSA, a thermistor, photodiode, and the 980 nm laser diode on a AlN-submount are soldered onto the circuit board. A single-mode cylindrical lensed fiber with support tube is laser welded to a clip, which is fixed onto the free side of the OSA. Typical coupling efficiencies for volume production are as high as 75–80% for a laser diode with a far field angle of 29° and 6° in the vertical and the lateral direction, respectively. The coupler alignment tolerances for the full

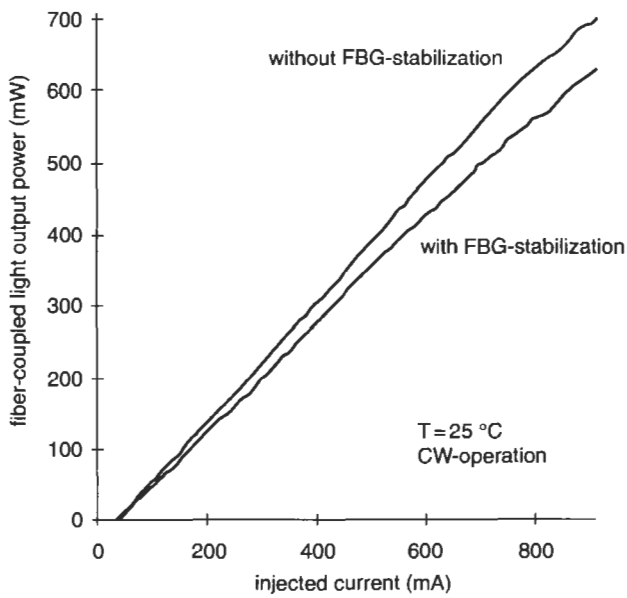


Fig. 14 Light output power coupled into a single mode fiber with and without FBG-stabilization using a ridge waveguide 980 nm pump laser diode. (Courtesy of Nortel Networks Optical Components, Zürich, Switzerland.)

widths of 95% in the lateral, vertical, and laser length axes are 1.6, 0.4, and 1.9 μm, respectively [49, 50]. Generally, laser welding secures the stability of fiber fixing in high-performance packages; otherwise care has to be taken to avoid creep of solder interfaces due to thermal and mechanical stress during operation.

With a wedged fiber lens that matches the ellipticity of the laser beam, coupling efficiencies up to 84% into single-mode fibers have been achieved for narrow-stripe 980 nm laser diodes with the vertical far field of the device reduced to 19° [9]. A fiber coupled light output power at 980 nm in excess of 600 mW is demonstrated in Fig. 14.

Two lens-coupling schemes are predominantly used for 1480 nm pump lasers. In the case of 20° × 25° far field angles [37], 80% coupling efficiency is obtained. The use of discrete lenses enables the positioning of an optical isolator into the light path inside the module, to avoid reflections at the close signal wavelength >1500 nm.

Conclusions

Today, there is a wide variety of optical amplifiers available, optimized for high power, low noise, low power dissipation, wide bandwidth, small size, and low cost. It is apparent that this variety of optical amplifiers will have a further impact on pump laser diode technology. The pump laser diodes will be more differentiated, and we will not only distinguish between 980 nm and 14*xx* nm pump lasers. There will be different pump lasers optimized for each individual specific application, e.g., special pump lasers for low noise, for counterpropagating Raman, for copropagating Raman, for low-cost and high-power amplifiers, all across the full S-, C- and L-wavelength band.

Acknowledgments

We like to thank N. Lichtenstein, N. Matuschek, T. Pliska, R. Baettig, R. Badii, H.-U. Pfeiffer, I. Jung, S. Pawlik, and B. Sverdlov for their support and the interesting and fruitful discussions during the realization of this chapter.

References

1. Grubb, S.G., T. Strasser, W.Y. Cheung, W.A. Reed, V. Mizrahi, T. Erdogan, P.J. Lemaire, A.M. Vengsarkar, D.J. DiGiovanni, D.W. Peckham, B.H. Rockney. 1995. High-power 1.48 μm cascaded Raman laser in germanosilicate fibers. Proc. Optical Amplifiers and Their Applications, Davos, Switzerland, paper SaA4: 197–199.
2. Gerard, F., S. Delepine, H. Bissessur, D. Locatelli, T. Fillion, N. Bouche, P. Salet. 2000. Single transverse-mode filtering utilizing ion implantation: application

to 1.48 μm unstable-cavity lasers. IEEE Photonics Technology Letters, Vol. 12, no. 11: 1447–1449.

3. Mathur, A., M. Ziari, M. Fisher, J. Neff. 2000. Over 1 W fibre-coupled power from single semiconductor source for pumping of Raman fibre amplifiers and erbium-doped fibre amplifiers. Electronics Letters, Vol. 36, pp. 410–411.
4. Garbuzov, D., R. Menna, A. Komissarov, M. Maiorov, V. Khalfin, A. Tsekoun, S. Todorov, J. Conolly. 2001. 1400–1480 nm ridge-waveguide pump lasers with 1 Watt CW output power for EDFA and RAMAN amplification. Optical Fiber Communication Conference 2001. Anaheim. California, paper PD18.
5. Yamazaki, H., T. Hosoda, Y. Sasaki, K. Komatsu. 2000. Over half-watt output power 1.48-μm wavelength EDFA pumping ASM LD's. Optical Fiber Communication Conference 2000, Baltimore, Maryland, paper ThK4.
6. Mathur, A., V. Dominic, S. DeMars, V. Agrawal, E. Vail, T. Kim, A. Kanjamala, M. Ziari. 2000. 800 mW chip power and half watt fiber-coupled power from narrow stripe 1450 nm laser diode. Optical Amplifiers and Their Applications 2000, Quebec, Canada, paper OMC5: 36–38.
7. Pliska, T., S. Mohrdiek, C. Harder. 2001. Power stabilisation of uncooled 980 nm pump laser modules from 10 to 100°C. Electronics Letters, Vol. 37, no. 1: 33–34.
8. Fujimoto, T., Y. Yamada, A. Okubo, Y. Oeda, K. Muro. 1999. High Power InGaAs/AlGaAs laser diodes with decoupled confinement heterostructure. SPIE Vol. 3628: 38–45.
9. Schmidt, B., S. Pawlik, B. Mayer, S. Mohrdiek, I. Jung, B. Sverdlov, N. Lichtenstein, N. Matuschek, C. Harder. 2001. Highly efficient 980 nm single mode modules with over 0.5 Watt pump power. Optical Fiber Communication Conference 2001. Anaheim. California: WC1–1–3.
10. Brugger, H., P.W. Epperlein. 1990. Mapping of local temperatures on mirrors of GaAs/AlGaAs laser diodes. Appl. Phys. Lett., Vol. 56 (11): 1049–1051.
11. Moser, A. 1991. Thermodynamics of facet damage in cleaved AlGaAs lasers. Appl. Phys. Lett., Vol. 59(5): 522–524.
12. Moser, A., E.E. Latta. 1992. Arrhenius parameters for the rate process leading to catastrophic damage of AlGaAs-GaAs laser facets. J. Appl. Phys., Vol. 71(10): 4848–4853.
13. Pessa, M., J. Näppi, P. Savolainen, M. Toivonen, R. Murison, A. Ovtchinnikov, H. Asonen. 1996. State-of-the-Art Aluminum-Free 980-nm Laser Diodes. Journal of Lightwave Technol., Vol. 14, issue 10: 2356.
14. Yang, G., Z. Xu, J. Xu, X. Ma, J. Zhang, L. Chen. 1998. Growth and fabrication of high performance 980 nm strained InGaAs quantum well lasers using novel hybrid material system of InGaAsP and AlGaAs. SPIE Vol. 3547, Sep 1998: 64–70.
15. Casey, H.C., Jr., M.B. Panish. 1978. Heterostructure Lasers. Academic Press, New York.
16. Nagarajan, R., V.V. Rossin, H. Ransom, N. Morozova, A. Kanjamala, R. Parke, D. Dawson, T.J. Kim, H. Clarke, K. Uppal, D. Coblentz, J.S. Major, Jr. 2000. Grating Stabilized 0.5 W, 980 nm pump modules. IEEE Semicon. Laser Conference, Catalog# OOCH37092: 27–28.

17. Schmidt, B., S. Pawlik, H. Rothfritz, A. Thies, S. Mohrdiek, C. Harder. 2000. 980 nm-module with very high power conversion efficiency. IEEE Semicon. Laser Conference, Catalog# OOCH37092: 29–30.
18. Harder, C., P. Buchmann, H. Meier. 1986. High-Power ridge-waveguide AlGaAs GRIN-SCH laser diode. Electronics Letters, Vol. 22: 1081–1082.
19. Amann, M.-C. 1979. New stripe-geometry laser with simplified fabrication process. Electronics Letters, Vol. 15, no. 14: 441–442.
20. Kaminow, I.P., R.E. Nahory, M.A. Pollack, L.W. Schulz, J.C. de Winter. 1979. Single-mode c.w. ridge-waveguide laser emitting at 1.55 μm. Electronics Letters, Vol. 15, no. 23: 763–765.
21. Tsukada, T. 1974. GaAs-Ga/sub 1-x/Al/sub x/As buried-heterostructure injection lasers. J. Appl. Phys., Vol. 45, no. 11: 421–426.
22. Aiki, K., M. Nakamura, T. Kuroda, J. Umeda. 1977. Channeled-substrate planar structure (AlGa)As injection laser, Appl. Phys. Lett., Vol. 30, no. 12: 649–651.
23. Takusagawa, M., S. Ohsaka, N. Takagi, H. Ishikawa, H. Takanashi. 1973. An internally striped planar laser with 3-μm stripe width oscillating in transverse single mode, Proc. IEEE, Vol. 61, no. 12: 1758–1759.
24. Burnham, R.D., D.R. Scifres, J.C. Tramontana, A.S. Alimonda. 1975. Striped-substrate double-heterostructure lasers, IEEE J. Quantum Electron, QE-11(7): 418–420.
25. Welch, D.F., D.R. Scifres, P.S. Cross, W. Streifer. 1987. Buried heterostructure lasers by silicon implanted, impurity induced disordering, Appl. Phys. Let., Vol. 51, no. 18: 1401–1403.
26. Kane, E.O. 1966. Semiconductors and Semimetals, Vol. 1. Academic Press, New York.
27. Coldren, L.A., S.W. Corzine. 1995. Diode lasers and photonic integrated circuits. John Wiley & Sons, New York.
28. Matthews, J.W., A.E. Blakeslee. 1976. Defects in epitaxial multilayers III. Preparation of almost perfect multilayers. J. Cryst. Growth, Vol. 32: no. 2: 265–273.
29. Tsang, W.T. 1982. Extremely low threshold (AlGa)As graded-index waveguide separate-confinement heterostructure lasers grown by molecular beam epitaxy. Appl. Phys. Lett., Vol. 40: 217–219.
30. Amann, M.-C., J. Buus. 1998. Tunable laser diodes. Artech House, Boston.
31. Witzigmann, B., A. Witzig, W. Fichtner. 2000. A multidimensional laser simulator for edge-emitters including quantum carrier capture. IEEE Transactions of Electron Devices, Vol. 47: 1926–1934.
32. Thompson, G.H.B. 1980. Physics of semiconductor devices. John Wiley & Sons, New York.
33. Thijs, P.J.A., T. van Dongen. 1989. High Quantum Efficiency, High Power, Modulation doped GaInAs strained-layer quantum well laser diodes emitting at 1.5 μm. Electronics Letters, Vol. 26: 1735–1737.
34. Kasukawa, A., T. Namegaya, N. Iwai, N. Yamanaka, Y. Ikegami, N. Tsukiji. 1994. Extremely High Power 1.48 μm GaInAsP/InP GRIN-SCH Strained MWQ Lasers. IEEE Photonics Technology Letters, Vol. 6, no. 1: 4–6.

35. Sasaki, Y., T. Hosoda, Y. Sakata, K. Komatsu, H. Hasumi. 1999. 1.48-μm-Wavelength ASM-DC-PBH LD's with Extremely Uniform Lasing Characteristics. IEEE Photonics Technology Letters Vol. 11: 1211–1213.
36. Kasukawa, A. 1999. Recent progress of high power laser diodes for EDFA pumping. Optical Fiber Communication Conference 1999, San Diego, California, paper TuC1: 20–22.
37. Kasukawa, A., T. Mukaihara, T. Yamaguchi, J. Kikawa. 2000. Recent progress of high power semiconductor lasers for EDFA pumping. Furukawa Review, No. 19: 23–28.
38. Kimura, T., N. Tsukiji, J. Yoshida, N. Kimura, T. Aikiyo, T. Ijichi, Y. Ikegami. 2000. 1480-nm Laser diode module with 250-mW output for optical amplifiers. Furukawa Review, No. 19: 29–33.
39. Oosenbrug, A. 1998. Reliability Aspects of 980-nm Pump Lasers in EDFA Applications. Proc. of SPIE, San Jose, California, 1998: 20–27.
40. Fukuda, M. 1991. Reliability and degradation of semiconductor lasers and LEDs. Artech House, Boston/London.
41. Tobias, P.A., D.C. Trindade. 1995. Applied Reliability. Chapman & Hall/CRC, Boca Raton.
42. Giles, C.R., T. Erdogan, V. Mizrahi. 1994. Simultaneous wavelength-stabilization of 980-nm pump lasers. IEEE Photonics Technology Letters, Vol. 6, no. 8: 907–909.
43. Ventrudo, B.F., G.A. Rogers, G.S. Lick, D. Hargreaves, T.N. Demayo. 1994. Wavelength and intensity stabilisation of 980 nm diode lasers coupled to fibre Bragg gratings. Electronics Letters, Vol. 30, no. 25: 2147–2149.
44. Achtenhagen, M., S. Mohrdiek, T. Pliska, N. Matuschek, C.S. Harder, A. Hardy. 2001. L-I Characteristics of Fiber Bragg Grating Stabilized 980-nm Pump Lasers. IEEE Photonics Technology Letters, Vol. 13, no. 5: 415–417.
45. Mohrdiek, S., M. Achtenhagen, C.S. Harder, A. Hardy. 2000. Detuning characteristics of fiber Bragg grating stabilized 980 nm pump lasers. Optical Fiber Communication Conference 2000, Baltimore, Maryland, paper ThK5.
46. Lammert, R.M., J.E. Ungar, S.W. Oh, H. Qi, J.S. Chen. 1998. High-power InGaAs-GaAs-AlGaAs distributed feedback lasers with nonabsorbing mirrors. Electronics Letters, Vol. 34, no. 9: 886–887.
47. Sharps, J.A. 1994. Reliability of hermetically packaged 980 nm diode lasers. IEEE Lasers and Electro-Optics Society 1994, Boston, paper DL1: 35–37.
48. Oosenbrug, A., A. Jakubowicz. 1997. Operational stability of 980-nm pump lasers at 200 mW and above. Proc. SPIE Vol. 3004, Fabrication, Testing, and Reliability of Semiconductor Lasers II: 62–70.
49. Müller, P., B. Valk. 2000. Automated fiber attachment for 980 nm pump modules. Electronic Components & Technology Conference 2000, Las Vegas, Nevada: 5–9.
50. Valk, B., P. Müller, R. Bättig. 2000. Fiber attachment for 980 nm pump lasers. IEEE Lasers and Electro-Optics Society 2000, Rio Grande, Puerto Rico, paper ThV2: 880–881.

Chapter 12 | Telecommunication Lasers

D. A. Ackerman, J. E. Johnson, and L. J. P. Ketelsen

Agere Systems, Murray Hill, New Jersey

L. E. Eng, P. A. Kiely, and T. G. B. Mason

Agere Systems, Breinigsville, Pennsylvania

Explosive growth in the communications industry, spurred by the Internet and resulting demand for communications bandwidth, has stimulated research in the area of semiconductor lasers. As the light sources in high-performance optical communications systems, specialized semiconductor lasers must meet demanding requirements on performance, reliability, and cost. Telecom lasers, as they are known, operate at high power and high speed at wavelengths adapted to fiberoptic systems. They are the subjects of the nine sections comprising this chapter. The first four sections provide an overview of telecom laser history, basic elements of device design and function, how applications shape laser design, and essential fabrication techniques. In-depth treatment of four major classes of telecom lasers, analog, directly modulated, electroabsorption modulated, and wavelength-selectable, are given in the second four sections. Concluding remarks are offered in the final section.

1.0 History of Telecom Lasers

Historical coevolution has shaped the characteristics of both semiconductor lasers and optical fiber for telecommunications systems. Multimode silica fiber carried ~850 nm light from early AlGaAs-based lasers in the late 1970s. By the 1980s, longer-wavelength InGaAsP based lasers took advantage of the 1300 nm chromatic dispersion minimum and the 1550 nm absorption minimum found in single-mode silica fiber. Today, most fiberoptic telecommunication systems operate in the 1550 nm Er-doped fiber amplifier (EDFA) window or at 1300 nm for short spans. However, recent improvements in fiber and amplifier technology have opened a 75 THz transmission window, a continuous band between the wavelengths of 1200 and 1700 nm (Thomas et al., 2000). History suggests that telecom systems will evolve to fill the broad optical fiber spectrum with information-carrying capacity.

Since 1980, optical telecommunication systems have experienced exponential growth in capacity (Chraplyvy et al., 1998) as shown in Fig. 1. Recent

OPTICAL FIBER TELECOMMUNICATIONS, VOLUME IVA

ISBN: 0-12-395172-0

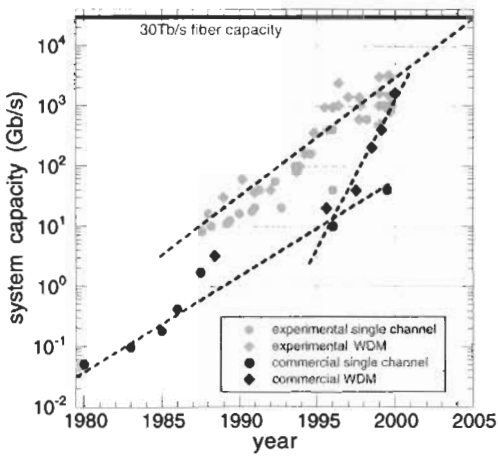


Fig. 1 Information-carrying capacity of telecom systems has increased exponentially with time. Experimental systems can be extrapolated to 30 Gb/s in 2005. Time to deployment of commercial systems is decreasing.

advances in wavelength division multiplexing have only quickened the pace, while market pressures have shortened the interval between lab demonstration and commercialization. Assuming an ambitious information density of 0.4 bits/s-Hz, the 75 THz fiber bandwidth can support a 30 Tb/s system, which, by extrapolation, might be expected in the near future, perhaps as soon as 2005.

Modern optical communication networks serve a variety of purposes from connecting countries on opposite sides of oceans to connecting personal computers in a business setting (Kaminow et al., 1997). They can be categorized by capacity to transmit information, topology, and hardware. A point-to-point network connects a transmitter, containing a telecom laser, through a span of fiber, with a receiver for detecting light. Rings of multicable fiber can support redundant and reliable networks of transmitters and receivers. High-speed transmitters of different wavelengths can operate simultaneously over a single fiber, provided that the power in the fiber is not so high as to excite distorting optical nonlinearities (Chraplyvy, 1990). Specialized components in such wavelength division multiplexed (WDM) systems can selectively inject or extract a single wavelength signal, adding flexibility to network architecture. Telecom lasers in WDM systems produce digitally modulated optical signals at well-defined wavelengths. A modern network architecture (Fig. 2) requires a variety of types of telecom lasers manufactured to diverse specifications for speed, cost, power, tunability, linearity, or temperature operating range. Telecom lasers are currently one of the enabling technologies of this booming and volatile industry as well as a topic of active research. The next section describes common features and differences among the varied component building blocks of modern telecom lasers.

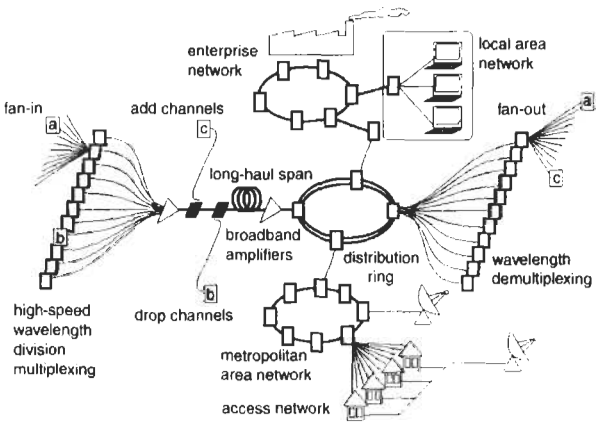


Fig. 2 A complex optical network including a WDM backbone with adds and drops and a distribution ring that feeds metropolitan, enterprise, local area, and access networks.

2.0 Telecom Laser Design and Function

As specialized variants of a broader class of semiconductor lasers, telecom lasers have unique adaptations enabling them to produce powerful, high-speed optical signals that faithfully transmit voice, data, and video signals via optical fiber. The Fabry–Perot laser is the simplest configuration, comprised of a gain medium within a waveguide, set between a pair of crystal facet mirrors. It produces a multimode optical spectrum at wavelengths determined by the Fabry–Perot cavity and the semiconductor bandgap. Single longitudinal mode telecom lasers, such as distributed feedback and distributed Bragg reflector lasers, use frequency-selective elements, which can be electronically or thermally tunable. Additional components that can be monolithically integrated with telecom lasers include optical modulators, amplifiers, spot size converters, and detectors.

Fabry–Perot lasers comprise an active medium that produces light and optical gain (Agrawal et al., 1993). Typically, this medium also serves as an optical waveguide, confining light in a single fundamental mode in two transverse dimensions and guiding light along the third longitudinal direction between two plane mirrors. The active medium can be as simple as a high-refractive-index bulk semiconductor material or can comprise a multiple quantum well stack embedded in a tailored waveguide. Drive current is injected all along the length of the laser via metallized contacts on the p- and n-doped sides of the laser diode. Typically, layers of semiconductor material are fabricated adjacent to the active region for the purpose of confining current to a thin stripe through which carriers are injected into the active region. A Fabry–Perot laser is shown schematically in Fig. 3.

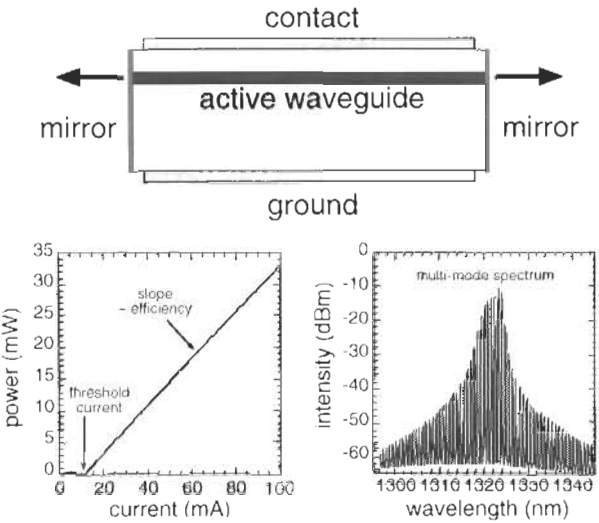


Fig. 3 Schematic of Fabry–Perot laser (top) with power-current (bottom left) and multimode spectral (bottom right) characteristics.

Important static properties of Fabry–Perot lasers include threshold current, differential efficiency (a measure of the rate of conversion of current to useful light), and optical spectrum. Sample characteristics are included in Fig. 3. High-speed performance, often characterized in the small signal regime using response to a modulated current, can be relatively flat up to a carrier-photon relaxation resonance frequency ~10 GHz. Equally important is a laser's large signal response to a rapid turn-on such as experienced at the leading edge of a pulse. Turn-on delay and damped ringing at the relaxation frequency in both output power and wavelength affect the fidelity of electrical-to-optical conversion. Optical fiber chromatic dispersion of the multiple laser frequencies of a Fabry–Perot laser severely limits the distance over which high-speed data can be transmitted due to wavelength-dependent propagation speeds through the fiber medium. A laser with a single longitudinal mode spectrum is better suited to high-speed use over long-haul optical fiber lines.

Distributed feedback (DFB) lasers incorporate a frequency selective diffraction grating in close proximity to the active region in order to favor a single longitudinal mode (Morthier et al., 1997). Light confined to a DFB laser waveguide and evanescently coupled to a first-order grating is resonantly diffracted back into the waveguide in the opposite direction for a narrow range of wavelengths tuned to the grating. Thus, as the name distributed feedback suggests, light within a narrow spectrum bounces back and forth in a DFB laser cavity, without the aid of facet mirrors. A practical DFB laser, including highly reflective (HR) and antireflecting (AR) facet coatings, is pictured in Fig. 4 along with a typical single-mode optical spectrum. The static linewidth

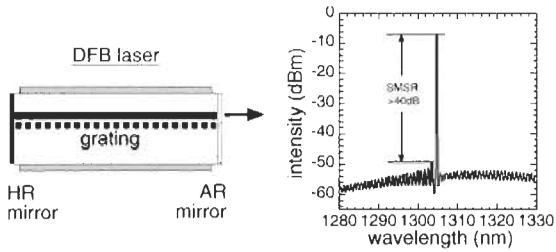


Fig. 4 Schematic side view of DFB laser with active waveguide, grating, HR, and AR mirror facets (top left) and single mode DFB laser spectrum (top right).

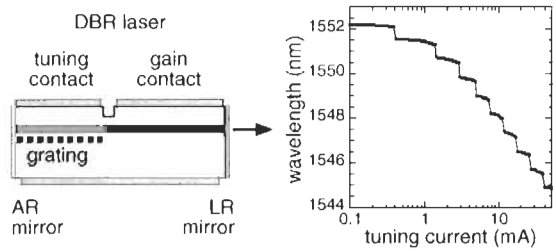


Fig. 5 Schematic of DBR laser with tuning section and gain section, grating, AR, and low-reflectivity (LR) mirrors (bottom left) along with tuning characteristics (bottom right).

of the dominant mode of a DFB laser can be less than 1 MHz or about 1 part in 10^8. The wavelength of a DFB laser can be tuned over a modest range by changing the temperature of the device. Tuning by this means is limited to perhaps 3 nm in a practical device. Another type of single-mode laser is designed expressly for tunability.

The distributed Bragg reflector laser (Amann et al., 1998) is comprised of a gain section, similar to that of a Fabry–Perot laser, coupled end-to-end with a tuning section as pictured in Fig. 5. The tuning section is a waveguide, largely transparent to light from the gain section, with a diffraction grating similar to that found in DFB lasers tuned to light from the gain section. In a DBR laser, the grating effectively forms one mirror and the opposing crystal facet the other. Substantial feedback occurs only over the narrow reflection spectrum of the grating, which typically selects a single longitudinal mode of the cavity formed between mirrors of the gain section. Current to the gain section varies the optical power from a DBR laser. Current injected into the tuning section results in a shift in refractive index of the tuning section waveguide, which in turn shifts the effective pitch of the grating as seen by the light from the gain section. By this means, the grating reflection can be tuned to select one of a range of longitudinal cavity modes as illustrated in Fig. 5. Tunability is desirable for lasers in WDM systems.

The spectral purity of DFB or DBR laser light makes possible the transmission of high-speed optical pulses over long dispersive fiber spans. Large-signal direct modulation of a single-mode laser causes ringing in the optical output power and frequency, however, which significantly broadens the spectrum of the dominant mode beyond its static linewidth. Chromatic dispersion of optical fibers ultimately limits long-haul, high-speed data transmission from directly modulated lasers to ~2.5 Gb/s, even with DFB or DBR lasers.

To achieve the extremely high data rates demanded of modern optical transmission systems, telecom lasers cannot be directly modulated. Instead, separate modulators are used to impose data on an otherwise constant laser signal (Kaminow and Koch, 1997). Signaling at data rates in excess of 40 GB/s is possible using external modulators. Modulators can be monolithically integrated with lasers as will be discussed in Section 7. Electroabsorption (EA) modulators comprising MQW layers embedded in waveguides appear similar in construction to lasers as shown in Fig. 3, but without facet mirrors. Light travelling in a waveguide of an EA modulator is transmitted if the photon energy is less than the energy of the absorption edge of the modulator material. Imposing a reversed bias electric field across the MQW layers of an EA modulator distorts the band structure of the quantum wells in such a way as to decrease the effective bandgap (quantum confined Stark effect). Light that was transmitted under zero or low bias is absorbed and converted to photocurrent under a few volts of reverse bias of the modulator. Rapid modulation of the continuous laser signal creates small but tolerable transient shifts in optical frequency due to phase shifts in the modulator. Losses, even in the on-state of EA modulators, can be overcome by amplifying laser light in monolithically integrated semiconductor gain sections or external optical amplifiers.

3.0 Telecom Laser Applications

As is clear from the previous section, a wide variety of semiconductor optoelectronic devices are used as light sources in optical telecommunications systems. Two broad classes of applications include conversion of electrical to optical signals for subsequent transmission over optical fibers and optical pumping of fiber amplifiers that boost the strength of information carrying optical signals. Within each of these two broad categories exist a number of distinct applications, each requiring a specifically tailored laser design. In this section we describe how telecom applications shape the associated laser design. In-depth design discussions for transmission lasers are the focus of Sections 5 through 8. Lasers for optical pumping are discussed elsewhere in this book.

Digital communication systems consist of four basic entities: a modulated light source, a length of fiber to transmit the modulated light, periodically placed optical amplifiers to compensate for the attenuation of transmission fiber, and a photoreceiver for conversion of optical to electrical signals

(Agrawal, 1997). In such systems, source modulation rate, optical fiber length and type, need for optical amplification, and component cost are the prime forces that shape laser source performance requirements. Specifically, these forces influence laser emission wavelength, modulation rate, wavelength chirp, and temperature sensitivity. In a long-haul system, in which information is transported between major cities, fiber spans are typically 100 to 3000 km and aggregate data rates are as large as 100 Gb/s to 1 Tb/s in a single fiber. Here, laser design is pushed toward high performance and high speed. However, since relatively few laser sources are needed to transmit the traffic of a large number of users in a single long-haul transmission span, laser cost is generally not a major concern. In contrast are short-reach systems designed to transport information within an office building or office complex. Short-reach fiber spans are less than 10 km with data rates of 655 Mb/s–10 Gb/s per fiber, modest in comparison to long-haul applications. The small number of users per transmitter in a typical short-reach system elevates laser cost above performance as a major design focus. Between these two extremes are systems targeted at information transport within a city. Fiber spans in metropolitan networks are of moderate length (2–50 km) with single fiber data rates similar to those found in a short-reach system and a number of users intermediate between short- and long-haul systems. In a metropolitan system, laser cost is still of prime concern but performance requirements can approach those of a long-haul system. In the next several paragraphs we discuss specific laser designs for long-haul and short-reach systems. Sources for metropolitan area systems are not distinctly different but instead represent a complex tug-of-war between high performance and low cost.

The generic long-haul dense WDM (DWDM) optical communications system depicted in Fig. 1 consists of multiple individually modulated sources, each with slightly different emission wavelengths, that are optically multiplexed onto a single fiber prior to transmission over very long fiber spans in which the signal is periodically amplified. Using DWDM, single-fiber information transport capacity approaching 1 Tb/s per fiber is achieved in commercial systems. Source emission wavelength around 1550 nm is currently chosen to match the gain of fiber amplifiers as well as to correspond to a minimum in fiber absorption, enabling signal transport over long distances without regeneration. However, in much of the 300 million miles of installed optical fiber, chromatic dispersion is relatively large at 1550 nm (Saleh et al., 1991). To avoid signal transmission degradation, lasers are designed to produce a single, pure output wavelength with minimal spectral width, also referred to as wavelength chirp (Ogawa, 1982). The need to transport data at high signaling rates requires rapid laser modulation in the range of 2.5 to 10 Gb/s and beyond. Satisfying these demands are three types of sources used in commercial DWDM long-haul systems. For systems in which the longest fiber span is less than 200 km, frequency chirp of a directly modulated 1550 nm DFB laser, typically ~0.3 nm (Kishino et al., 1982), is small enough to support a data

rate of 2.5 Gb/s or lower. However, for systems employing fiber spans from 200 km and longer, an extremely low-frequency chirp of less than 0.02 nm is required and provided by external modulation of a CW-operated laser source. Either an electroabsorptive element (Suzuki et al., 1987) or a Mach-Zehnder interferometer (Pollock, 1995) can produce external modulation of a continuous laser signal. Electroabsorptive (EA) modulation works through voltage control of the semiconductor bandgap through the Franz-Keldysch or quantum confined Stark effect. Mach-Zehnder (MZ) modulators work by splitting and recombining a CW optical signal from a high-power laser source. Modulation is achieved through voltage control of the relative phase shift of the two recombined signals. There are relative advantages to each means of external modulation. EA modulators are smaller and require lower drive voltages than MZ modulators and lend themselves to monolithic integration, which tends to reduce manufacturing costs. However, MZ modulators provide better modulation characteristics than EA components (for example, chirp control). In practice both external modulation schemes are employed to meet the various specific needs of long-haul optical systems.

Short-reach systems are far simpler than the high-capacity DWDM architectures used in long-haul information transmission. A typical short-reach system consists of a modulated laser source, a section of fiber without amplification, and a photoreceiver. Compared to a long-haul system, low data rate and low net fiber dispersion allow considerable relaxation of laser performance. However, the small number of users per transmitter in short-reach systems makes low laser cost essential. Cost reduction is achieved in two ways, through the use of multilongitudinal mode sources and by eliminating costly laser temperature control. Multilongitudinal mode or Fabry–Perot laser sources (Agrawal and Dutta, 1993) have cost advantages derived from fabrication simplicity and higher yields relative to inherently more complicated single-mode DFB lasers. The level of chromatic dispersion-induced signal degradation originating from the broad spectral width of a Fabry–Perot laser (>3 nm) can be tolerated in a short-reach system due to the small total dispersion of short fiber spans. Dispersion is further reduced by aligning the low-cost laser emission wavelength with the 1310 nm dispersion minimum of standard optical fiber (Saleh and Teich, 1991). Remarkably, the relatively simple Peltier cooler is often more expensive than the laser chip itself. However, elimination of the cooler can only be effective at reducing overall cost if the so-called uncooled laser is designed to operate over a broad temperature range, typically $-40°C$ to $+85°C$ for the outside plant, and $0°C$ to $+70°C$ for a central office. Performance degradation in output power and modulation characteristics at high temperature presents great difficulty (Bhat et al., 1994). Laser source heat dissipation and active region quantum well structure must be carefully optimized. Heat dissipation is addressed by minimizing device series resistance and by proper heat sinking (Joyce et al., 1975). The quantum well structure is usually designed to contain a large number of wells to increase optical confinement

factor and reduce laser threshold currents at high temperature (Zory, 1993). The quantum well structure must be simultaneously designed to avoid carrier transport problems through the quantum well stack, as well as to ensure that the quantum well waveguide supports only a single fundamental optical mode. These issues place an upper limit on the number of quantum wells. State-of-the-art uncooled 1310 nm InP-based laser sources can produce more than 20 mW power at 85°C and can be directly modulated at up to 10 Gb/s. It should also be noted that, due to improvements in laser manufacture, uncooled DFB laser cost is now sufficiently low to make them popular light sources for short-reach applications where fiber chromatic dispersion is not negligible.

Laser sources for analog systems are fundamentally different in performance, although not in fabrication, from their digital counterparts. Analog optical systems are primarily used for efficient transmission of analog video signals (Darcie, 1991). When transmitted electrically, cable TV signals require regeneration every few kilometers due to high electrical power dissipation in coaxial transmission lines. Optical transmission, on the other hand, increases transmission span, reducing system cost, and improving system reliability by reducing the overall number of system components. Optical analog systems work by converting a multichannel analog electrical television signal into an optical signal. The television signal consists of an electrical, amplitude modulated, subcarrier multiplexed signal containing well over 100 individual television channels. Electrooptic conversion is accomplished by direct modulation of a semiconductor laser. The resulting optical signal is then transmitted up to 25 km through optical fiber. After reconversion, the electrical signal is distributed to end users. High-power (>30 mW) single-mode operation, low RIN, and low-fiber dispersion requirements favor DFB lasers emitting at 1310 nm for analog lasers. They must also faithfully convert electrical to optical signals to minimize distortion and crosstalk between channels. Demands on linearity in electrical to optical conversion distinguish analog application of telecom lasers from all others. Linearity is achieved through optimized current confining technology aimed at reducing electrical shunt paths that divert terminal current from passing directly through the laser active region. The complicated electrical carrier/photon interaction that is present in DFB lasers must also be carefully tailored to achieve a high degree of linearity (Ketelsen et al., 1991). Electrical predistortion of the laser drive signal is also used to compensate for residual laser nonlinearities.

4.0 Fabrication of Telecom Lasers

Idealized optoelectronic components are discussed in Sections 2 and 3 in terms of perfect materials in arbitrary three-dimensional geometries; the materials are defect-free, doping profiles and heterointerfaces are abrupt, and gratings are perfectly regular. Realizing these ideals depends on the properties

of the materials themselves and on the epitaxy and fabrication tools available. Historically, semiconductor laser development has mirrored that of process technology development. In many instances, the availability of new or improved tools has shaped the evolution of telecom laser designs. An example of this is the replacement of liquid-phase and vapor-phase epitaxy (LPE and VPE) with organometallic vapor phase epitaxy (OMVPE). Improved thickness and composition control offered by OMVPE opened the door for completely new active layer designs incorporating quantum wells, and later, strained-layer quantum wells. In other instances, new or improved tools enabled high-volume manufacture of existing designs with better control, improving yields, reducing costs, and allowing better design optimization. As an example, the replacement of wet-etching of holographic Bragg gratings in DFB lasers with reactive ion etching (RIE) greatly improves the control of the grating coupling, leading to a higher yield of devices operating in a single longitudinal mode (Takemoto et al., 1989). Armed with improved processes, the telecom laser industry has moved into huge, but more cost-sensitive, markets, such as metropolitan, local area, and access networks.

Most of the fabrication technologies that are used to make telecom lasers are the same as those used in the manufacture of silicon integrated circuits (Sze, 1983). Contact photolithography, chemical vapor deposition (CVD) of oxide and nitride films, ion implantation, evaporation of metal and dielectric films, and rapid thermal annealing are examples of laser fabrication technologies that are nearly identical to their silicon counterparts, except for the smaller wafer sizes of InP and GaAs. The major exception is epitaxial growth, which is much more demanding for III–V materials than silicon. Early growth techniques such as LPE and VPE (Malik, 1989; Tsang, 1985; Willardson et al., 1990) have relatively poor control of thickness and composition and are not capable of the abrupt ($\sim$1 monolayer) interfaces necessary to make modern telecom lasers with quantum well active layers. These early crystal growth techniques have been largely supplanted by molecular beam epitaxy (MBE) (Parker, 1985; Panish et al., 1993) and OMVPE (Malik, 1989; Willardson and Beer, 1990; Stringfellow, 1989). In the MBE technique, high-purity solid or gaseous sources are heated to produce beams of group III and group V elements that impinge upon a heated substrate in an ultra-high vacuum (UHV) chamber. The UHV environment makes MBE the preferred technique for the growth of AlGaAs/InGaAs/GaAs lasers, because of the high reactivity of Al with H_2O and O_2. It also allows in situ monitoring of the growth by characterization techniques such as reflection high-energy electron diffraction (RHEED). High quality InGaAsP/InP lasers have been grown using MBE, but the high vapor pressure of P makes it more difficult to maintain UHV conditions. For this reason, OMVPE is the preferred growth technique for phosphide compounds usually used in longer wavelength (1200–1700 nm) devices. In the OMVPE technique, group III metal alkyls such as trimethylgallium and trimethylindium and group V hydrides such as AsH_3 and PH_3

in a H_2 carrier gas are passed over the heated substrate. The gas flow creates a stagnant boundary layer through which the III–V precursors diffuse to the hot wafer surface, where they decompose into elemental species. The chamber can be at atmospheric pressure or as low as ~100 Torr, with the low pressure being preferred for quantum well growth because of the speed at which the gas mixture can be changed. Because OMVPE relies on vapor-phase diffusion, it is well suited for growth over nonplanar surfaces, such as Bragg gratings and etched mesas. Large, automated multiwafer reactors are commercially available for both MBE and OMVPE, enabling high-volume manufacturing.

Since the structure of a semiconductor laser is determined by the specific combination of epitaxy and processing techniques used to fabricate it, variations are almost limitless. Laser structures can be broadly classified by type of optical waveguide as gain-guided, weakly index-guided, and strongly-index guided. Because of the need to couple light into optical fiber, gain-guided lasers are not generally used for telecom applications, except very recently as pump lasers for cladding pumped fiber lasers (Po et al., 1993; Grubb et al., 1996). Of the other two classes, the most common structures used in telecom applications are the ridge waveguide laser and the buried heterostructure laser. Although there are many variations even within these subclasses, we will use them to illustrate how telecom lasers are made.

The cross-sectional structure of a ridge-waveguide laser is illustrated in Fig. 6(a). In this type of laser, the vertical layer structure, which forms the slab waveguide and heterostructure carrier confinement, is grown by MBE or OMVPE. Next, a portion of the upper cladding layers alongside the active stripe is etched away and replaced with a lower index dielectric, creating a small lateral effective refractive index step. The ridge can be etched by wet chemical etchants, which may be either material-selective or nonselective, or by a dry etching technique, such as reactive ion etching (RIE), which is generally nonselective. Wet etching is simple and low-cost, and selective etchants can be used to provide excellent control of the etch depth, but wet etching suffers from poor control of ridge width due to undercutting of the mask. Dry etching is usually very anisotropic, that is, there is little undercut of the mask, but control of the etch depth is not as good as selective wet etching. After etching the ridge, a dielectric, typically SiO_x or Si_xN_y, is deposited over the wafer by standard techniques such as plasma-enhanced chemical vapor deposition (PECVD), sputtering, or evaporation. Next, an opening in the dielectric is made on the top of the ridge with standard wet or dry oxide etching techniques, and a photoresist mask is formed on the wafer for the p-ohmic contact metals. Alternatively, the original ridge etch mask can be used as a liftoff mask for a subsequent dielectric layer to self-align the metal contact to the ridge. The p-ohmic metals are then deposited, typically with electron-beam evaporation. Because the ohmic metals, e.g. (AuBe)TiPtAu for p-ohmic and (AuGe)TiPtAu for n-ohmic, are not homogeneous, a liftoff technique is used to define the patterns in place of etching. The wafer is then thinned using a combination of abrasive and

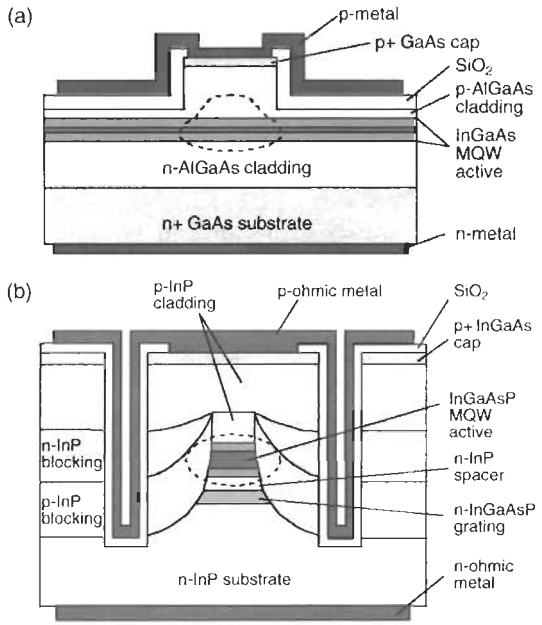


Fig. 6 (a) Illustration of a ridge waveguide laser structure. Materials are given for a 980 nm pump laser. The dashed line shows the location of the optical mode. (b) Cross-sectional illustration of a planar buried heterostructure laser. Materials are given for a 1.55 μm directly-modulated DFB laser. The dashed line shows the location of the optical mode.

chemical–mechanical polishing, and the n-ohmic contact is deposited. The ohmic metals are subsequently alloyed in a reducing atmosphere in a furnace tube or rapid thermal anneal (RTA) system. Finally, the wafer is then scribed and cleaved along the crystal planes into bars of 10–20 lasers and loaded into a fixture for facet coating. Facet coatings are typically evaporated, and consist of one or more dielectric layers designed to passivate the facets and give the desired reflectivity at the lasing wavelength. After coating, individual lasers are cleaved from the bars, ready for testing and packaging.

Figure 6(b) shows the cross-sectional structure of a typical buried heterostructure laser. In order to illustrate the fabrication of Bragg gratings, we will use the example of a 1.55 μm DFB laser. The first step involves the epitaxial growth of a thin n-InGaAsP grating layer and n-InP cap. Optical holography is used to create a short-period grating in a thin photoresist layer on the wafer, which is then used as a mask to transfer the grating into the grating layer. After removing the resist, OMVPE is used to planarize the grating with n-InP and grow the MQW active layers and a portion of the upper p-InP cladding. A stripe mask is then formed on the wafer using standard

techniques. The active mesa is formed by wet or dry etching, or a combination. A third OMVPE step then buries the active mesa in InP. The layers are doped p- or n-type, or made semi-insulating by doping with Fe, in order to form a series of homojunction blocking diodes that confine carrier injection to the active stripe. The oxide mask is then stripped, and a fourth OMVPE step grows the remaining p-InP cladding and a p+ InGaAs contact layer. If high-speed modulation is required, a pair of deep trenches is sometimes etched along the active stripe to reduce the capacitance of the blocking junctions. The remainder of the process starting with the dielectric deposition and p-ohmic metal deposition is similar to the process described above for ridge waveguide lasers.

Photonic integrated circuits (PICs), in which lasers, modulators, passive waveguides, and other optical functions are integrated on the same chip, represent one of the most important trends in telecom laser design today, because they make it possible to greatly increase the functionality of a device while holding down costs. The complexity of these PICs and the conflicts involved in joining devices with dissimilar materials, doping profiles, optical modes, and electronic properties, make their design and fabrication challenging. As an example, consider the externally modulated tunable laser (EA-DBR) shown in Fig. 7. This PIC consists of a two-section tunable DBR laser monolithically integrated with a semiconductor optical amplifier (SOA), a tap for power monitoring, and a high-speed electroabsorption (EA) modulator (Ketelsen et al., 2000). These sections are described in Section II. The process that marries these parts into a single device must provide two different active materials (gain and absorption), a nonabsorbing passive waveguide, and a means of optically coupling them together. In addition, each section needs to be electrically isolated from its neighbors, and the optical interfaces must have low transmission loss and back-reflection.

Many fabrication and epitaxial techniques have been developed for integrating dissimilar materials, including "butt-joint" growth (Soda et al., 1990), quantum well disordering (O'Brien et al., 1991), and selective area growth (SAG) (Tanbun-Ek et al., 1994). SAG is used to grow both the lower passive waveguide and the MQW active layers of the EA-DBR of Fig. 7. In the SAG

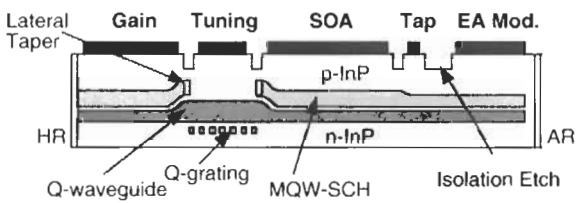


Fig. 7 Illustration of a monolithically integrated PIC, consisting of a 1.55 μm tunable DBR laser, semiconductor optical amplifier, power monitor tap and electroabsorption modulator.

technique, oxide pads with a narrow gap between them are first patterned on the wafer. During subsequent OMVPE growth, the reactants in the boundary layer diffuse laterally over the pads, resulting in an enhanced growth rate in the gap and around the edges of the pads. The composition of the layer is also changed because of the different diffusivities of the group III reactants. Low-loss optical coupling of the active and passive sections of the EA-DBR is achieved by etching a lateral taper in the upper waveguide over the thickest portion of the lower waveguide. Electrical isolation in this device is achieved by etching away the highly conductive upper cladding layers between sections, but high-energy ion implantation of H, He, or O can also be used for this purpose.

In order to be used in a fiberoptic transmission system, the fragile laser chip must be provided with a package that protects it from the outside environment, connects it with the external electronics of the system and couples the light into an optical fiber. The choice of how each of these functions is implemented is determined by the cost and performance requirements of the application (Mickelson et al., 1997). As an example, consider the telecom laser package in Fig. 8. This type of package is typically used for high-end applications such as directly modulated DFB lasers. The case is a 14-pin butterfly lead package with impedance-controlled feedthroughs on the RF inputs and a hermetically sealed window in the package wall to couple the light out. A thermoelectric cooler (TEC) and a thermistor sensor are used to keep the laser at a fixed temperature, so that lasing wavelength and modulated output characteristics can be tightly controlled. The optical subassembly (OSA) consists of a silicon or ceramic submount on which the laser chip, backface monitor detector, and collimating optics are mounted. The OSA is mounted on a ceramic carrier that has bond sites for the OSA and optical isolator, as well as electrical wiring traces. An optical isolator is necessary for single-mode

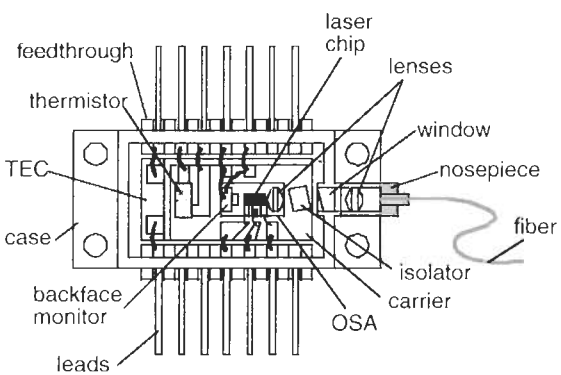


Fig. 8 Illustration of a typical cooled, isolated laser module for high-bitrate telecom applications. The view is from the top, with the lid removed.

lasers in order to prevent any reflected light from fiber connectors or other optical components from being coupled back into the laser and perturbing its operation. After wire bonding the TEC, thermistor, laser, and monitor to the feedthroughs, the case is purged with an inert gas and the lid is welded in place to hermetically seal the package. Alignment tolerances of less than 0.1 μm are typically required to achieve good coupling efficiency to single-mode fiber, so the alignment is performed with the laser on while monitoring the power in the fiber, a procedure known as active alignment. Once the power in the fiber is maximized, the nosepiece is locked in position using laser welding.

Because of the push of optics into enterprise and access applications, there is considerable pressure to reduce the cost of laser packaging. These applications generally use uncooled Fabry–Perot lasers designed for use without a TEC or optical isolator, but the active alignment using a lensed fiber and hermetically sealed package technology remains costly. New technologies are being developed to make lasers with integrated spot-size converters (Moerman et al., 1997), which expand the tightly confined optical mode of the laser to more closely match the mode size of optical fiber. Spot size conversion of lasers allows coupling to cleaved single-mode fiber without the use of lenses with ~2.0 μm alignment tolerances. This makes it possible to mechanically place the fiber in a precision V-groove on the OSA without active alignment. Nonhermetic packaging is also being investigated as a way to reduce costs. In an example of this technology, the OSA (with passively aligned spot-size converted laser) is bonded to a low-cost lead frame. The optical subassembly and fiber are embedded in a blob of silicone gel for protection from moisture and stress. Then a plastic case is molded around the lead frame in much the same way low-cost ICs are packaged (Tatsuno et al., 1997). Vertical cavity surface emitting lasers (Wilmsen et al., 1999) or VCSELs are another potential candidate for low-cost lasers because of their large emitting aperture that is well matched to optical fiber. As these technologies mature, broadband fiber-to-the-home access networks will become a reality for millions of consumers.

5.0 Analog Lasers

5.1 HISTORY OF ANALOG TRANSMISSION

Simultaneous transmission of numerous analog video channels over fiberoptic links has enabled high quality, low-cost, reliable distribution of community antenna television (CATV) signals and information services. Unlike the binary on-off keying of standard digital fiberoptic transmission, fiberoptic analog links transmit through glass fiber an amplitude-modulated analog optical replica of radio frequency (RF) video signals similar to those received by

analog televisions. Semiconductor lasers, which are commonly used in digital fiberoptic signal transmission, also serve as optical sources for fiber-based analog transmitters. Analog signal transmission imposes unique and stringent specifications upon laser transmitter performance, however. As a result, lasers adapted for use in analog transmitters differ from their digital counterparts, particularly in terms of linearity. While a directly modulated digital laser manufactured for a 2.5 Gb/s transmitter typically needs a larger bandwidth than an analog laser, it needs only to produce binary ones and zeros. In contrast, an analog laser is often required to faithfully reproduce analog RF signals, with distortion of less than a part in a thousand (~60 dBc) in each of over 100 channels. In this section, we discuss the specialized breed of semiconductor 1310 nm laser developed for analog transmitters used for CATV signaling. We focus upon nonlinear mechanisms of distortion inherent to the laser.

The history of optical transmission of analog signals dates back to the late 1970s (Michaelis, 1979; Chinone et al., 1979; Ito et al., 1979). Early experiments were plagued by systemic impairments related to multimode fiber (Sate et al., 1981), multitransverse-mode lasers (Chinone et al., 1979) and optical reflections (Ito et al., 1979; Hirota et al., 1979). Single transverse-mode lasers coupled to single-mode fiber solved some of the problems encountered in the early experiments (Chinone et al., 1979). However, chromatic dispersion of optical fiber still limited transmission of analog signals produced by multilongitudinal or Fabry–Perot lasers. Single longitudinal mode distributed feedback (DFB) lasers operating at 1310 nm solved multimodal noise problems (Nakamura et al., 1984). Development of highly linear 1310 nm DFB lasers, tuned to the fiber dispersion minimum and uniquely suited for analog transmission, finally enabled production of analog transmitters in the late 1980s that were capable of sending many subcarrier multiplexed channels over significant spans of fiber with excellent signal-to-noise and distortion characteristics. Wide distribution of optical CATV signals, initially delayed by poor quality and high price of laser transmitters, commenced.

Historically, the use of analog optical links was considered an interim step en route to fully digital video distribution. However, rapid deployment of 1310 nm fiberoptic analog links within conventional RF coaxial cable systems gave birth to cost effective hybrid fiber-coax (HFC) transmission systems (Chiddix et al., 1990). In these systems, fibers linked head-end receivers to hubs serving 500 to 1000 subscribers. Fiber also penetrated deeper into the network to carry optical signals to finer-grained RF coax distribution points (Chiddix et al., 1990, 2000). Favorable economics of HFC systems, a large embedded base of conventional, long-lasting analog televisions, a consumer cost barrier to digital-to-analog set-top converters, protracted standards negotiations, and costs of digital high definition television have all contributed to extending the lifecycle of analog CATV systems well beyond original expectations. Research on analog lasers for CATV transmission has evolved as competition has spurred more ambitious technical specifications on performance. Through

the first half of the 1990s, 1310 nm analog transmitter technology matured. Electronic predistortion circuitry, first used in conjunction with very nonlinear lasers (Straus and Szentesi, 1975), was refined and incorporated into transmitters (Darcie and Bodeep, 1990) to boost performance and/or reduce cost by loosening laser chip specifications.

Attention in the research community shifted in the mid-1990s to analog transmission at 1550 nm to take advantage of lower fiber losses and Erbium-doped fiber amplifiers (EDFAs) and to address long fiber spans (Atlas et al., 1995; Phillips, 1996; Kuo et al., 1996). High fiber chromatic dispersion at 1550 nm, relative to that at 1310 nm, coupled with DFB laser chirp forced transmitter design to incorporate external modulation or alternate transmission formats. External modulators required linearization schemes to reduce distortion (Wilson, 1999). Quadrature-amplitude modulation (QAM) was employed at 1550 nm as a means of loosening linearity specifications while maintaining high information-carrying capacity (Fuse et al., 1996; Chen, 1998). Spectral broadening techniques for reducing stimulated Brillouin scattering were investigated as a means of improving signal quality for 1550 nm systems (Phillips and Sweeney, 1997). The actual 1550 nm DFB lasers used in QAM transmission of CATV signals are similar in nature to standard 1550 nm digital DFB lasers such as those used in OC-48 transmission, discussed elsewhere in this chapter. In the early 2000s, the markets for 1310 and 1550 nm analog transmitters remain competitive (Blauvelt, 2001). In the context of the extremely cost-sensitive CATV market and largely digital world of telecommunications, it is a measure of overall effectiveness of 1310 nm systems that they have not been displaced by digital 1550 nm systems. The remainder of this section deals with the evolution of 1310 nm analog lasers and the specific adaptations that allow them to operate in an open-loop application calling for extreme fidelity in electrical-to-optical conversion.

5.2 ANALOG TRANSMISSION IMPAIRMENTS

Semiconductor lasers are, by nature, linear electrical to optical converters. Simple models of lasers describe light output as a function of injected current in terms of photon and carrier rate equations (Agrawal and Dutta, 1993). From a simplified rate equation model, static, linear light output vs. injected current is predicted for currents beyond a well-defined threshold current. Figure 9 shows a measured light output vs. current (L-I) plot as well as showing the derivative of output power dL/dI vs. current. In principle, such a linear transducer is ideal for optical transmission of analog signals. The rate equation model predicts, for modulation of current about a bias point above the threshold value, a proportional modulation of optical output. A schematic representation of such an analog current signal and resulting optical signal are shown in Fig. 9 as well. Deviation from linearity (Bissessur, 1992) causes distortion in the

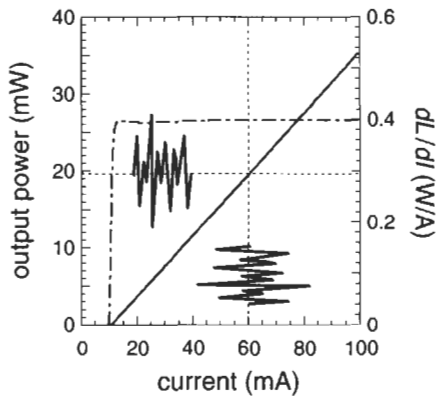


Fig. 9 Measured light output (solid) and derivative dL/dI (dash-dot) vs. current showing static laser linearity. Vertical dashed line shows bias point with schematic analog current input superimposed, while proportional light output response is plotted on horizontal dashed line.

optical signal and the generation of distortion crossproducts, which becomes particularly problematic in multichannel systems spanning a large range of modulation frequencies, as will be discussed. Sources of distortion can be traced to the laser as well as the laser package (Helms, 1991) and interactions of optical signals with the fiber (Blauvelt et al., 1993).

Early experiments using 1310 nm semiconductor lasers to transmit analog vestigial sideband (VSB) signals showed promise (Darcie and Bodeep, 1990; Darcie, 1990) but distortion mechanisms inherent in the semiconductor lasers were recognized to limit transmission fidelity (Darcie et al., 1985). One early-recognized mechanism relates to modulation of a semiconductor laser through the highly nonlinear *L-I* region around lasing threshold resulting in clipping distortion. While care is taken to inject a sufficiently high DC bias current into an analog laser to raise its operating point to well above threshold, statistically, the superposition of signals from individual analog channels results in occasional excursions of drive current below threshold. Clipping of the optical signal occurs as the laser output cuts off and remains zero at currents below threshold. Clipping distortion was investigated analytically (Saleh, 1989) and via simulation (Phillips and Darcie, 1991), from which fundamental limitations on the number of channels and the optical modulation depth (OMD) per channel were derived. By operating with prescribed limits of channel count and OMD, clipping distortion can be held to an acceptable level.

Effects of nonlinearity, such as those due to clipping, become detrimental in transmission by mixing multiple analog channels. Typically, a set of carrier frequencies, such as those used for NTSC (North American) or PAL (European) frequency plans, are used to subcarrier multiplex video channels carrying CATV information. Second harmonic distortion products from i^{th}

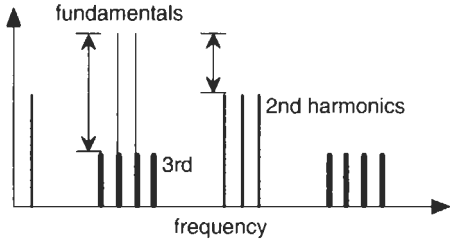


Fig. 10 Schematic representation of a pair of fundamental tones with second and third harmonics.

and j^{th} channels appear at frequencies equal to $(\pm f_i \pm f_j)$ including the case of $i = j$. Since typical analog CATV frequency plans exceed one octave in frequency range, second-order distortion products can appear in-band, adding unwanted spurious signals. Third-order harmonic distortion products of the form $(\pm f_i \pm f_j \pm f_k)$ always appear in-band. Figure 10 depicts a pair of carrier tones (fundamentals) together with second and third order products. The power in the IMD products relative to the power in the carriers is a measure of linearity. Although third-order is typically weaker than second-order distortion, the larger number of third-order products and the fact that they are always in-band makes third-order distortion as much of a problem as second-order distortion. Operational specifications detail the integrated power that may be carried by unwanted harmonic distortion products in a given channel band in units of dBc, decibels relative to the carrier. Such specifications usually call for a given carrier-to-noise ratio (CNR) and optical modulation depth for each channel.

5.3 1310 nm ANALOG LASER DESIGN

Typical fiberoptic CATV systems use 1310 nm DFB laser-based transmitters broadcasting over 100 downstream channels through a fiber plus splitting total attenuation of over 10 dB. Interharmonic modulation products contribute distortion (IMD) to each channel band. Under testing, during which carrier tones are substituted for video signals, aggregate harmonic distortion is distinguishable from the carrier in each channel. Composite second-order (CSO) distortion measures the maximum aggregate second-order IMD product in a given channel relative to the carrier tone in that channel. Composite triple-beat (CTB) distortion measures the corresponding quantity for third-order IMD products. Typical specifications for an 80 to 110 channel system with an 8 to 11 dB link loss budget are CNR $>$ 53 to 54 dB (measured in a 4 MHz band), CSO $<$ -63 to -65 dBc, CTB $<$ -67 to -70 dB (both measured in a 30 kHz band) at a CW fiber-coupled power of 2 to $\sim$30 mW (Blauvelt, 2001). These specifications might also include QAM channels loaded at the high frequency end of the band.

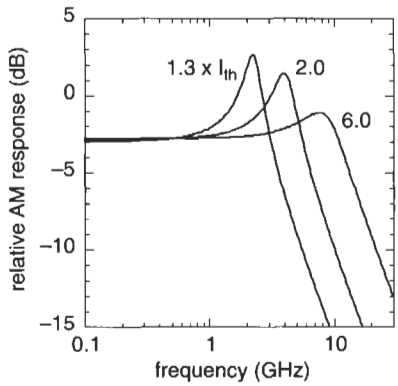


Fig. 11 Calculated response of laser to small signal modulation showing variation of resonance at 1.3, 2.0, and 6.0 threshold current bias.

Another source of distortion relates to the nonlinear relaxation oscillation arising from carrier–photon interaction (Lau and Yariv, 1984). A simple rate equation model for semiconductor lasers (Agrawal and Dutta, 1993) shows that the frequency of natural relaxation in photon and carrier density depends on optical power (or equivalent photon density) as illustrated in Fig. 11 for three different bias currents. As the time averaged photon density in a laser cavity increases, so does the strength of feedback coupling carriers and photons thereby increasing the relaxation frequency. Because of this dynamic nonlinearity inherent in the laser, the photon density or output power of a laser modulated high frequency oscillates not only at the driving frequency, but also at harmonics of the driving frequency (Darcie et al., 1985). Photon density-dependent relaxation results in harmonic distortion similar to that produced by an audio speaker with nonlinear mechanical properties. In addition to its dependence upon photon density, relaxation frequency also increases with differential gain. Early 1310 nm lasers used for analog transmission were typically based on bulk active designs with relatively low differential gain and hence, relatively low relaxation frequencies of ~5 GHz. Although the analog CATV band extends to RF frequencies of only ~0.7 GHz, or less in the early days, nonlinear effects of multi-GHz nonlinear relaxation oscillation were plainly observable. Multiquantum well active lasers tend to have higher relaxation frequencies compared to bulk active devices due to higher differential gain and, hence, tend to suffer less from relaxation related distortion.

Nonlinearity due to current leakage also contributed substantially to nonlinear electrical-to-optical conversion, especially in early analog lasers. Static output power vs. drive current characteristics of ideal lasers are linear due to the extremely high efficiency of stimulated carrier conversion to photons. Thus, ideal semiconductor lasers are inherently linear to a very high degree. However, in real diodes, some fraction of drive current flows around the active

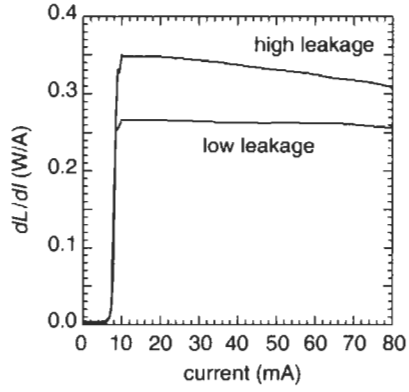


Fig. 12 Slope efficiency dL/dI vs. current for two lasers showing high leakage with drooping efficiency and low leakage. Data obtained using pulsed current to avoid junction heating.

region through homojunctions intended to block current and therefore does not pump the active material. Some injected carriers, energetic enough to pass over heterobarriers and recombine outside the active region, are also lost. Measurement of static output slope efficiency (dL/dI) vs. drive current characteristics of analog lasers, shown in Fig. 12, showed drooping of efficiency with increasing current most notably at high injected current levels. Homojunction and heterojunction leakage bears a nonlinear relationship to current above threshold and therefore introduces nonlinearity into the output power vs. current characteristic of a real laser. Semiquantitative analysis of current leakage is accomplished by spatially imaging radiative recombination at homojunction blocking layers or heterobarriers adjacent to active regions (Kuidersma et al., 1985). Infrared images of homojunction and heterobarrier leakage recombination in capped mesa buried heterostructure lasers appear in Fig. 13. The ~0.96 μm emission is bandpass filtered and appears along homojunction blocking structures as well as above the heterobarrier adjacent to the active stripe. Efficient blocking structures have been introduced to largely eliminate homojunction or shunt leakage with an eye to avoiding highly capacitive blocking junctions that would adversely affect modulation bandwidth. These include combinations of doped and undoped InP layers that form reverse-biased junctions (Kuindersma et al., 1985; Asada et al., 1989; Lee et al., 1998). Doping profile control (Belenky et al., 1995; Flynn and Ackerman, 1997) reduces heterobarrier leakage. Output power vs. drive current characteristics of low-leakage lasers show good linearity, even at high currents and powers. Reduction of current leakage at high bias is essential to producing highly linear analog lasers since operation at high DC bias currents is required to reduce clipping distortion, maintain high relaxation frequency and provide high output power.

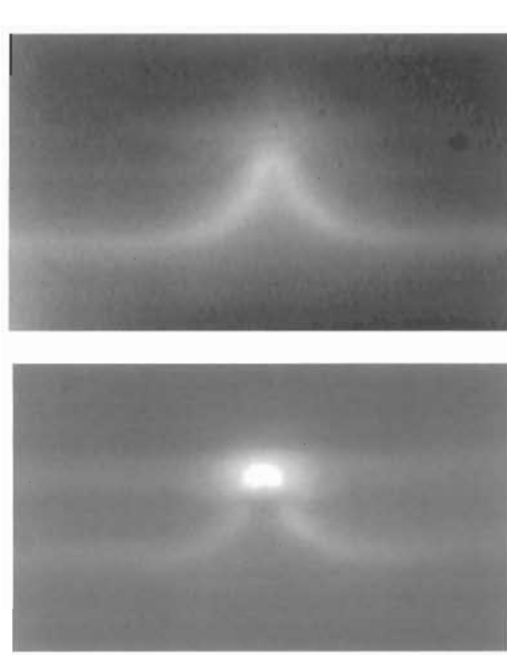


Fig. 13 Infrared images of homojunction and heterobarrier recombination in InP layers of buried heterostructure (CMBH) laser. Radiation is ~0.96 μm.

Another source of distortion inherent in semiconductor lasers derives from gain compression (Majewski, 1990; Girardin et al., 1995; Vinzio et al., 1998; Futakuchi et al., 1998; Yasaka et al., 1991), the phenomenon that reduces gain at high photon densities. Included into simple rate equations, this effect produces a nonlinearity in the static light output vs. injected current characteristics that increases at high bias currents (Bissessur, 1992). Gain compression is associated with very rapid processes such as carrier heating. Over the analog modulation band of <1 GHz, a static estimate of nonlinearity due to gain compression is adequate.

Fiber chromatic dispersion limits span length for multilongitudinal mode FP analog lasers. Distributed feedback lasers emit radiation in a single longitudinal mode, often referred to as single-frequency, by means of a diffraction grating coupled to the active region over the length of the laser cavity. Light within a band of wavelengths defined by the pitch and coupling strength of the grating is reflected and coupled to radiation traveling in the opposite direction within the active waveguide. Nonlasing longitudinal modes are typically suppressed by a factor of 10^4 (40 dB) in DFB lasers. With spectral purity at this level, fiber chromatic dispersion is much improved over that experienced using FP lasers. In addition, noise due to multilongitudinal mode competition is reduced in DFB lasers. Relative intensity noise in DFB lasers is typically

< -160 dB/Hz over the analog frequency band. However, single frequency, low-leakage DFB lasers also suffer impairments to linearity.

In addition to clipping distortion, nonlinear dynamics and current leakage, DFB lasers exhibit unique nonlinearities related to a phenomenon known as spatial hole-burning (SHB) (Plumb et al., 1986; Morton et al., 1989; Rabinovich and Feldman, 1989; Motheir and Baets, 1996). Nonuniformity in photon density along the optical axis of a semiconductor laser cavity can result in nonuniformity in carrier density (Ketelsen et al., 1991). Higher rates of stimulated recombination deplete carriers faster in the brighter regions of the laser cavity, leaving corresponding regions of low carrier density, loosely called holes. Although the round-trip gain in the DFB laser cavity must remain high enough to overcome cavity losses, the nonuniformity in gain resulting from holes burned in the carrier population affects DFB laser operation. SHB was recognized in quarter-wave shifted DFB lasers in which regions of high photon density occur around an intentional phase shift in the diffraction grating (Whiteway et al., 1989). Asymmetrically facet coated DFB lasers (Henry, 1985), of the type used for analog transmission, also exhibit nonuniform photon density as the reflection from a highly reflective (HR) facet combines with the distributed reflection of the grating to create complex longitudinal modes.

An illustrative example is the case of constructive reinforcement of HR facet and grating reflections at the HR facet end of a DFB laser. In this case, the intensity of the field is at a maximum at the HR facet end of the cavity and drops exponentially toward the antireflection (AR) coated facet. The intense field nearest the HR facet burns a hole in carrier population in the vicinity of the HR facet as illustrated in Fig. 14. In response to the resulting drop in gain near the HR facet, carrier density in the dimmer part of the cavity rises, maintaining round-trip gain at a level sufficient to balance cavity losses. Increasing drive

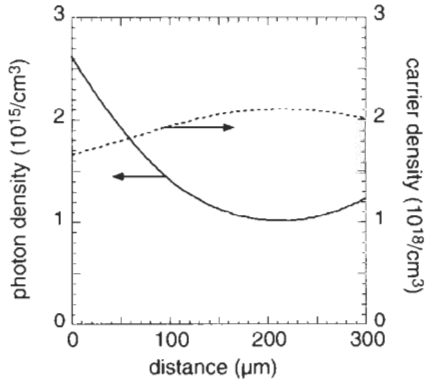


Fig. 14 Calculated photon (solid) and carrier (dashed) densities as a function of distance from the HR facet of a DFB laser with midgap mode illustrating spatial hole-burning.

current to the laser intensifies the degree of spatial hole-burning. With increasing drive current, refractive index, a function of local carrier density, shifts. The frequency of the cavity mode and the envelope of the longitudinal field in the laser cavity respond by shifting as well. The threshold current of the mode can be thought of as deviating from the actual threshold with the onset of SHB. The shape of the longitudinal mode in the presence of SHB is different than that just above threshold when SHB is negligible. The proportion of photons exiting the AR-coated output facet, a measure of output efficiency, shifts with SHB as the mode changes shape with increasing current. It is the changing cavity mode threshold and output efficiency that distorts the otherwise linear output power vs. current characteristic to produce SHB-related nonlinearity (Ackerman et al., 1996). In the example above, slope efficiency, defined as the derivative of AR facet output power with respect to drive current, starts at threshold at a value which quickly increases as SHB in the cavity causes the longitudinal mode to shift from the HR toward the AR-coated facet. SHB saturates as current increases beyond several times threshold. The level of distortion that comes from SHB-related nonlinearity contributes significantly to the overall analog distortion due to all other effects, especially from threshold to several times threshold current. The dynamics of spatial hole-burning have been investigated thoroughly (Phillips et al., 1992; Kuo, 1992; Kito et al., 1994; Schatz, 1995). To a good approximation, the time constants associated with SHB can be considered short, since above threshold, carrier lifetimes are dominated by stimulated emission. Thus, SHB-related distortion of cavity modes occurs in phase (or 180° out of phase) with modulation current (Phillips et al., 1992).

Coupling of photons in a DFB laser cavity to the grating is characterized by grating strength κ. Soon after SHB was recognized in quarter-wave shifted DFB lasers, SHB was also theoretically shown to contribute to distortion in analog lasers (Morthier et al., 1990). A value of normalized coupling of $\kappa L \sim 0.8$ (Susaki, 1991) to 1 (Takemoto et al., 1990) was determined optimal for analog DFB lasers. Subsequent experimental studies (Zhang and Ackerman, 1995) confirmed that SHB adversely affected analog performance of DFB lasers with asymmetric facet coatings. The phase of a reflection from a HR-coated facet, relative to the phase of DFB grating reflection is determined in detail by the position of the difficult-to-control facet cleave relative to the grating. Significant change in SHB-related nonlinearity can be observed for facet phase changes of 10°, where 360° corresponds to a typical first-order grating pitch of 240 nm. As such, HR facet phase for a HR/AR-coated DFB laser is a random variable due to lack of control of the mirror forming process. Dramatically varied effects of SHB upon static linearity are calculated in Fig. 15 for nine DFB lasers from three wafers of various grating strengths. Thus, even for a single wafer, SHB-related nonlinear behaviors associated with random HR-coated facet positions relative to the grating produce diverse effects. For the calculated optimum value of grating coupling, SHB still produces a minimum variation of SHB-related effects. Thus, the random nature of

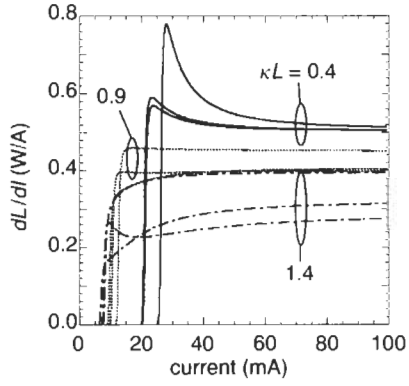


Fig. 15 SHB produces nonlinear static slope efficiency for high (1.4), intermediate (0.9) and low (0.4) κL and for various HR-facet phases, all contributing to analog distortion. Note that the distribution over facet phase is most similar for intermediate value of κL.

SHB-related distortion has a predictable impact on analog yield (Plumb et al., 1986; Ackerman et al., 1996; Zhang and Ackerman, 1995). Improvements, including refined strained layer MQW active structures, more efficient current blocking layers, and increased control of grating coupling (Morthier, 1994; Watanabe et al., 1995; Watanabe Aoyagi et al., 1995) in conventional DFB lasers pushed analog performance to higher levels through the early 1990s.

A solution to the problem of random HR facet phase was devised using a DFB laser with a missing patch of grating adjacent to the HR facet (Yamada et al., 1997). Field envelopes in such a configuration vary less as a function of HR facet phase compared to conventional DFB lasers. Consequently, the variation over a wafer of SHB character is smaller, and analog distortion originating from SHB can be more tightly controlled. Yield from such a partially corrugated configuration is improved over conventional DFB yield (Yamada et al., 1997).

Nonlinearities discussed above such as current leakage and SHB are inherent in semiconductor lasers. Nonlinearities that distort analog transmission also arise from interactions of laser and package or fiber. While these mechanisms are not the focus of this section, it is useful to mention two examples that impact analog laser design. From early days of analog experiments, it was clear that laser cavities must support only one transverse mode (Chinone et al., 1979). Ill effects of multitransverse mode lasers have been documented in many contexts (Guthrie et al., 1994; Schemmann et al., 1995). In multimode cavities in which only a single mode lases, scattering of light from the fundamental to higher-order modes results in coherent spatial beating along the cavity (Peale et al., 1999). Such mode competition adds noise to the signal and tends to steer the output beam, even in buried heterostructure devices such

as 1310 nm analog DFB lasers. Beam steering is instantaneously correlated to injected current. Thus, despite laser linearity, coupled signal linearity is adversely affected as the output beam wanders relative to coupling optics.

An example of laser–fiber interaction that affects transmission of analog signals relates to the magnitude of laser chirp (Blauvelt et al., 1993), a quantity measuring the small signal variation of optical frequency with drive current. Optical frequency modulation (FM) that accompanies analog amplitude modulation (AM) can create distortion if FM is converted to unwanted AM by optical frequency dependent optics such as would be found in packaged lenses, windows, or isolators. To reduce distortion due to FM-to-AM conversion, laser chirp should be designed to be small. A DFB laser forced to lase on the high-energy side of its gain peak, where differential gain is high, tends to exhibit lower chirp, as is the case for 1550 nm digital DFB lasers. However, Rayleigh scattering in the transmission fiber creates a competing effect. Interferometric noise occurs when weak double back-scattering occurs throughout the fiber (Blauvelt et al., 1993). Increasing the magnitude of laser chirp forces the interferometric noise to higher radio frequencies that are out of the analog band. Therefore, analog laser chirp is bounded below by FM-to-AM distortion and above by interferometric intensity noise. Proper design of laser active material and optical confinement factors can center chirp at a value that is tolerable from both considerations. Typically, 1310 nm analog DFB chirp measures ~250 MHz/mA, double or triple that of digital DFB lasers designed for digital applications.

Finally, high output power is key in making a good analog laser due to span loss and power splits in a CATV distribution network (Blauvelt et al., 1992). Given a good laser design, it is essential to couple as large a fraction to the output fiber as possible. This means not only a low-loss optical isolator in the package, but optics that are well matched to the numerical aperture of the laser and are mechanically stable over time. Fiber coupling efficiencies of over 70% can be obtained using aspherical lenses, thereby reducing the laser AR-facet power needed for +15 dBm fiber power to 45 mW. While difficult to achieve, the advantages in terms of span lengths and potential for power splitting make high-power analog DFB lasers advantageous in CATV network architectures.

In summary, 1310 nm DFB lasers are used to transmit high-quality multichannel analog video signals over fiberoptic links. Requirements on open-loop linearity and noise are stringent. Analog DFB lasers suffer a variety of nonlinear mechanisms including current leakage, nonlinear dynamics, and spatial hole-burning, not to mention suffering from a variety of fiber-package and fiber-laser interactions. Each of these nonlinearities has been studied extensively. Despite these mechanisms, analog transmitters based on solidly designed 1310 nm DFB lasers have proliferated and have become part of the video distribution infrastructure.

6.0 Directly Modulated Digital Lasers

Over the past decade, optical telecommunications transmission has evolved from single wavelength transmission at 1310 or 1550 nm and speeds up to 2.5 Gb/s, to more than 80 wavelength-specific channels with speeds up to 10 Gb/s. The choice of operating wavelength is determined by the dispersion and optical loss of the fiber used. At 1310 nm, commercial standard fiber has a zero dispersion but high optical loss of 0.6 dB/km, whereas at 1550 nm the dispersion is 17 ps/nm-km with very low loss of 0.2 dB/km. While advances in fiber technology have made it possible to tailor the dispersion coefficient to a desired wavelength, most commercial components are required to operate over standard embedded fiber or are characterized by their performance over standard fiber as a common yardstick. Long-haul WDM systems are designed around 1550 nm due to the low fiber loss and availability of Erbium-doped optical amplifiers (EDFA). Shorter links in uncontrolled temperature environments are often designed at 1310 nm due to zero fiber dispersion, and the availability of high-speed uncooled DFB lasers at this wavelength.

Directly modulated semiconductor distributed feedback (DFB) lasers are used extensively in today's telecommunications systems. The attractive features of modulating the laser output directly with input current are ease of use, high optical output power capability, ability to operate over a wide temperature range, and low cost. However, associated with current modulation of a semiconductor laser is a carrier density modulation, giving rise to frequency chirp that limits the ultimate transmission distance due to the effects of fiber dispersion. Directly modulated diode lasers can transmit signals from low data rates (Mb/s) to 2.5 Gb/s up to 200 km in the 1550 nm regime, without the need for pulse regeneration, with dispersion penalties less than 2 dB. For longer distances external modulation is required, either integrated with the laser chip, as discussed in Section 7 for an EML, or a separate $LiNbO_3$ or GaAs MZ modulator. The effects of operating wavelength and data rate for a fixed laser chirp are shown in Fig. 16(a) and 16(b). Here we see that for 2.5 Gb/s transmission at 1620 nm (L-Band), we expect a 0.5 dB penalty compared with the 1550 nm band (C-Band). From Fig. 16(b), it becomes apparent that at 10 Gb/s, dispersion becomes a larger issue limiting direct modulation applications to lasers operating in the 1300 nm range.

Even as externally modulated lasers become cost competitive with DFBs, directly modulated lasers will continue to play a role where there are distinct advantages: unamplified links, very-short-reach 10 Gb/s, and power dissipation-sensitive applications where small form factor and uncooled wide temperature performance are required.

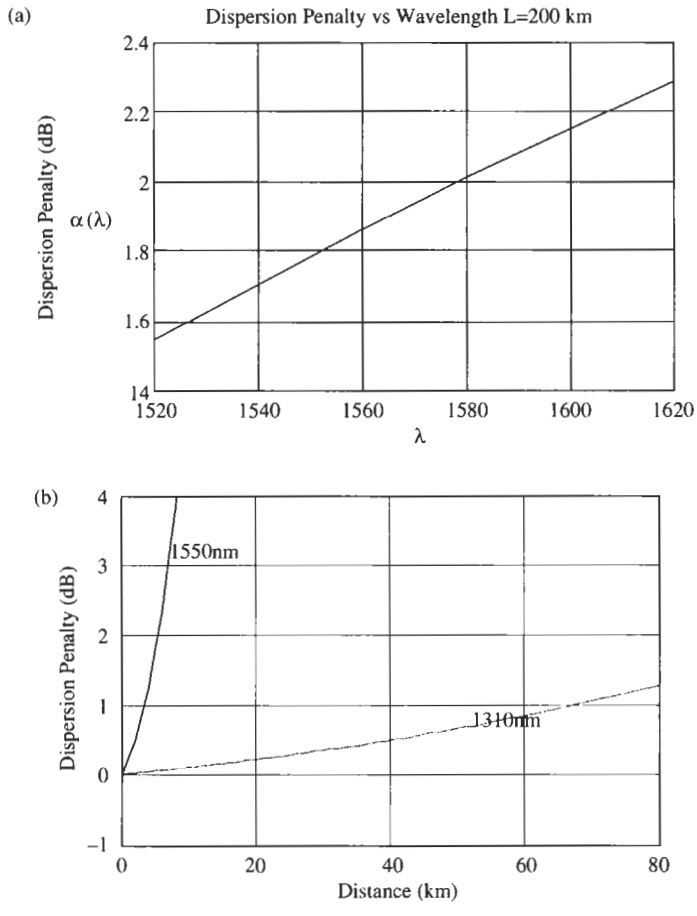


Fig. 16 (a) Fiber chromatic dispersion increases 2.5 Gb/s transmission by 0.5 dB from the center of the C-band to the center of the L-band. (b) At 10 Gb/s fiber chromatics dispersion restricts direct modulation to the 1310 nm range.

6.1 RATE EQUATIONS

To analyze the performance of directly modulated lasers it is necessary to model laser output power, $P(t)$, and wavelength chirp, $\lambda(t)$, as function of time-dependent input current $I(t)$. The quality of transmission, parameterized by bit error rate (BER), will depend on the detailed shape of the pulse, the spectral content, and dispersion in the fiber. When current is applied to a semiconductor laser, the output power overshoots and then reaches steady state through damped relaxation oscillations. The emission wavelength similarly varies under modulation, as fluctuations in power gives rise to fluctuations

in carrier density. Important laser properties such as power overshoot, relaxation oscillations, damping, and wavelength chirp can be described well by the single-mode rate equations:

$$\frac{dN}{dt} = \frac{I}{e \cdot V} - \frac{N}{\tau_n} - G(N) \cdot (1 - \varepsilon \cdot S) \cdot S \tag{1a}$$

$$\frac{dS}{dt} = \Gamma_a \cdot G(N) \cdot (1 - \varepsilon \cdot S) \cdot S - \frac{S}{\tau_p} + \frac{\Gamma_a \cdot \beta_{sp} \cdot N}{\tau_n} \tag{1b}$$

where $N(t)$ is the carrier density, $S(t)$ is the photon density in the laser cavity, e is the electronic charge, V is the active volume, $G(N)$ is the optical gain, Γ_a is the optical confinement factor, β_{sp} is the spontaneous emission factor, and $\tau_{n,p}$ are the carrier and photon lifetimes, respectively. The gain compression factor ε must be included in the analysis in order to model key chirp and time-dependent power behavior, as will be seen shortly. Physically, ε represents a reduction in gain at the lasing wavelength due to the presence of the optical field, i.e., spectral hole burning. By solving the rate equations for a digital input current, we can model the output power $P(t)$ and wavelength chirp, $\Delta\nu(t)$. Key parameters are power overshoot, associated chirp, and the observed wavelength offset between the on and off states.

The rate equations can be solved analytically for chirp and output power waveforms (Corvini and Koch, 1987). Results from this analysis are summarized in the following. The frequency chirp is given by:

$$\Delta\nu(t) = -\frac{\alpha}{4\pi}\left(\frac{1}{P}\frac{dP}{dt} + \kappa \cdot \Delta P(t)\right) \tag{2}$$

where $\alpha = (dn/dN)/(dg/dN)$ is the linewidth enhancement factor, n is the refractive index, and $\kappa = 2\Gamma\varepsilon/V\eta h\nu$. The first term represents the dynamic chirp, which is the wavelength shift associated with on and off. The second term is the adiabatic chirp and is the steady-state emission frequency difference between the on and off states. The adiabatic chirp is linearly proportional to the output power difference between the on and off states. Dynamic chirp dominates when the laser off state is close to threshold, or if dP/dt is large due either to a large output power swing or fast rise time current pulses.

The relaxation oscillations are exponentially damped by the factor

$$\gamma = \frac{\Gamma\varepsilon P_0}{V\eta h\nu} \tag{3}$$

From the above equations, we see that to design a laser for minimal chirp it is necessary to adjust the optical confinement factor to minimize dynamic chirp without excessively increasing the adiabatic chirp. In addition, the overall chirp can be minimized by reducing the α factor; this can be done by increasing the differential gain. In practice this has been achieved with a multiple quantum

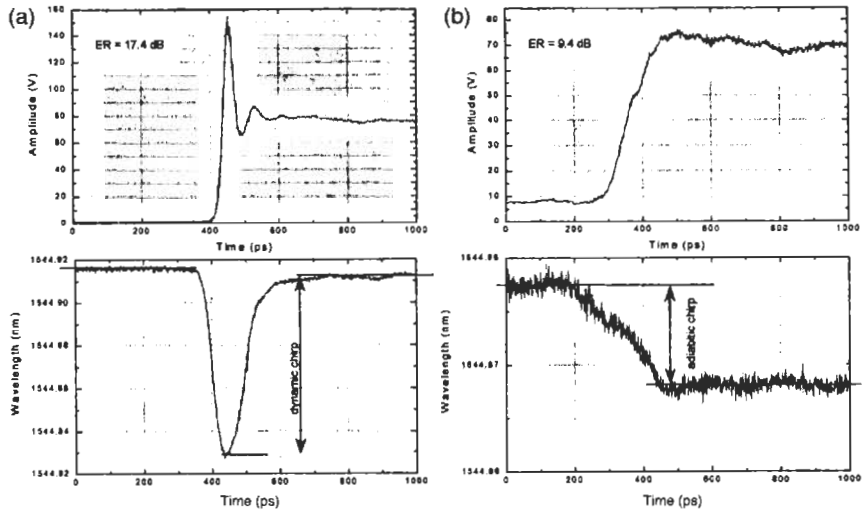


Fig. 17 Directly modulated dynamic wavelength chirp decreases as extinction ratio is decreased from (a) 17.4 dB to (b) 9.4 dB. Adiabatic chirp is significant for both extinction ratios.

well (MQW) active region, p-doping the active region, and increasing the offset, or detuning, between the material gain peak and the DFB emission wavelength. These efforts have resulted in alpha factor decrease from $\alpha \sim 3$ to $\alpha < 1.5$ in the past decade.

Following are some experimental results, which demonstrate the various components of chirp in commercial semiconductor DFB lasers. Figure 17 shows output power and wavelength as a function of time for a commercial 1.55 μm buried heterostructure (BH) DFB laser (Agere Systems, product code D2570). The BH laser is a single-mode laser with high lateral optical confinement. In the figure, the laser is operating at 2.5 Gb/s at a fixed 40 mA modulation current, for two different extinction ratios of 17 dB and 9.4 dB. For the 17 dB extinction case, we see about 1A of dynamic chirp. In Fig. 18, the influence of extinction on dynamic chirp is shown, and we see that for ER < 10 dB the dynamic chirp is negligible. While system signal-to-noise ratio improves with on/off state discrimination, the dynamic laser chirp sets an upper limit to the maximum extinction ratio used in practice. Figure 19 shows the effect of detuning on adiabatic chirp. As detuning $(\Delta\lambda_{DFB} - \Delta\lambda_{gain})$ becomes more negative, the differential gain dG/dN is enhanced with little change in dn/dN resulting in reduced α. The adiabatic chirp (Fig. 20) for a given design can be reduced a factor of 2 by changing the detuning from -20 to -50 nm. The variation within each sample of lasers with a nominal detuning target is due to inter- and intrawafer variation in detuning, doping, optical confinement, and facet phase.

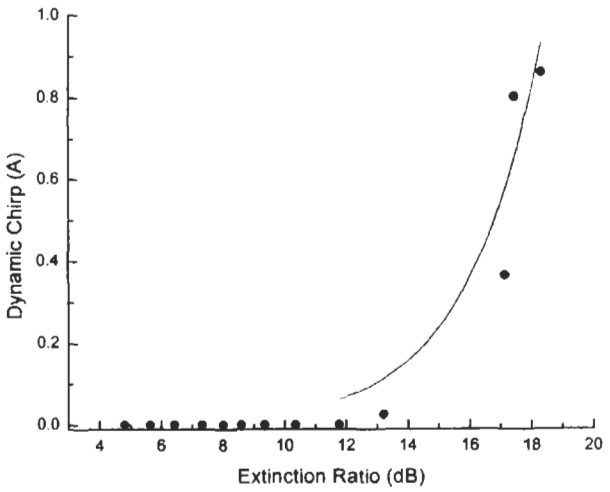


Fig. 18 Directly modulated dynamic chirp is negligible below 10 dB extinction.

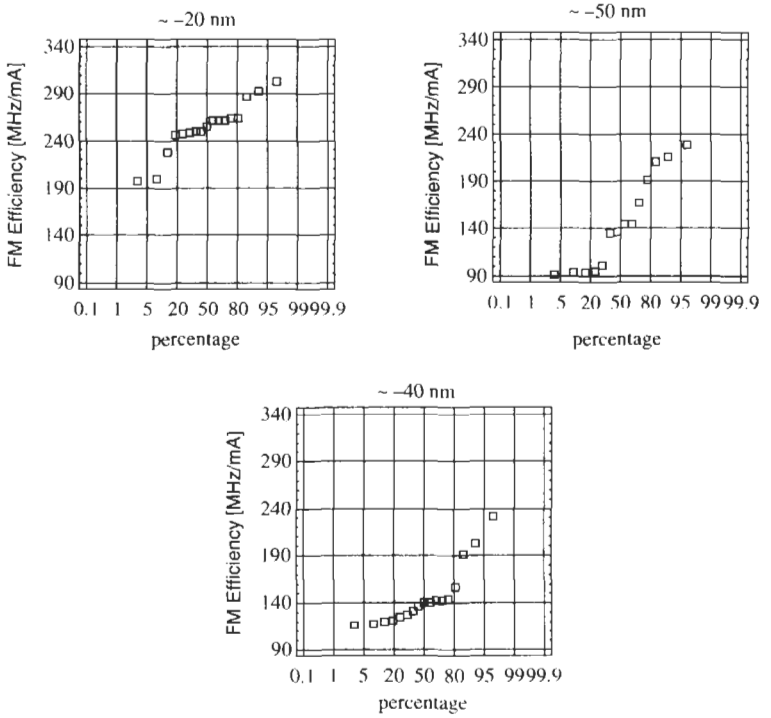


Fig. 19 FM efficiency is reduced as emission/gain wavelength detuning becomes more negative.

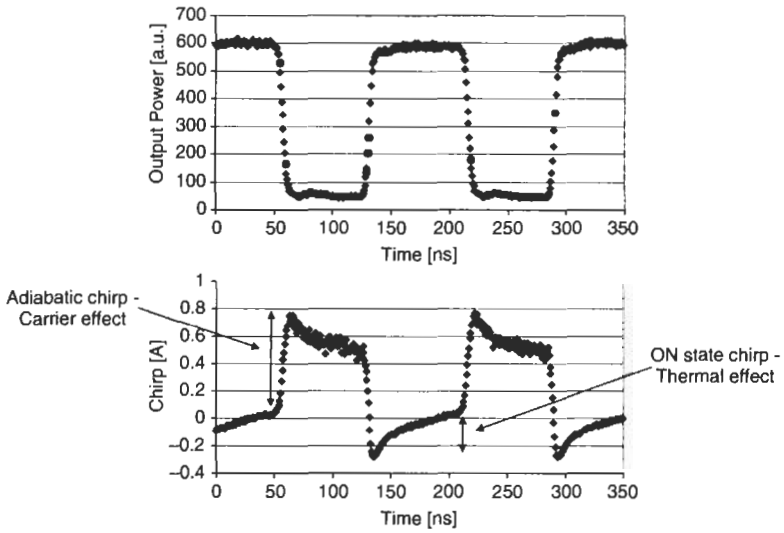


Fig. 20 Adiabatic wavelength chirp resulting from bias current-induced heating is a significant effect for data rates <100 Mb/s.

An additional source of laser chirp is caused by self-heating (Fig. 20). This effect is noticeable at low data rates, i.e., less than 100 Mb/s, and results in a red-shift of the on state wavelength and a blue shift of the off state wavelength. This effect can be minimized through design for low thermal impedance between the active region and the heat sink.

6.2 WAVELENGTH BUDGET FOR LASERS USED IN DWDM SYSTEMS

In DWDM systems, the wavelength budget is of utmost importance. As the channel spacing reduces from 200 to 100 GHz, and then to 50 GHz, it is important to ensure that there is no optical crosstalk between adjacent channels and the on-state power does not degrade due to wavelength drift out of the filter pass band. Note that the effect of filtering the off state is essentially an increase of the extinction ratio and enhanced transmission performance (Morton, 1997; Shtengel, 1998). This will hold as long as the off state power is sufficiently filtered out of other channels.

The wavelength requirements must be met over the life of the system. The wavelength age rate is thermally activated with an activation energy E_a,

$$\Delta\lambda(t) = Ae^{E_a/kT} \tag{4}$$

and wavelength reliability estimates can be made using conventional reliability theory (Nash, 1993). Figure 21 shows a laser wavelength age experiment where a sample of devices have been operated at elevated temperatures

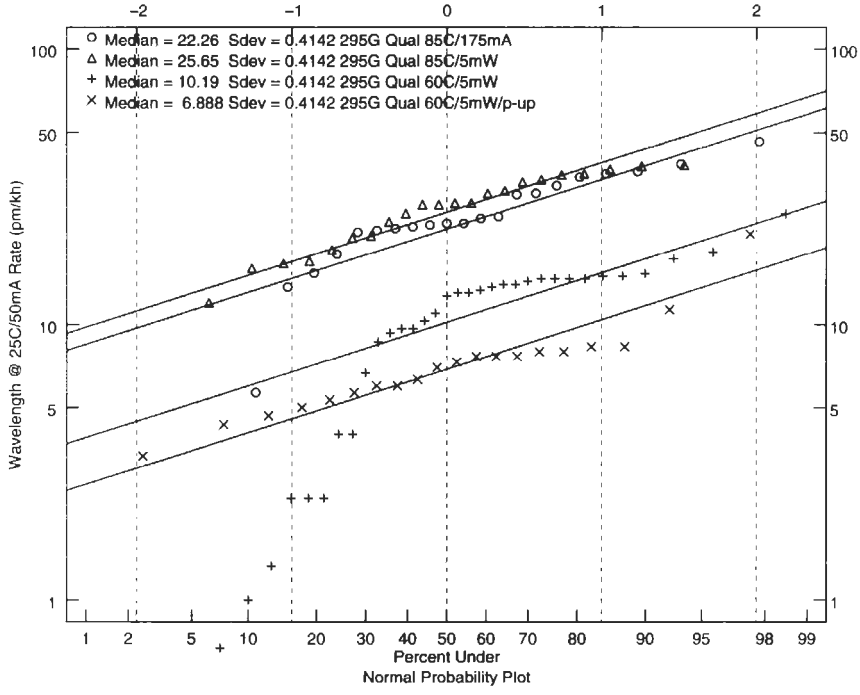


Fig. 21 Wavelength aging under varying accelerated aging conditions (295G qual wavelength aging).

(85°C and 60°C) in order to accelerate aging. The wavelength of each device was measured before and after age at $I = 50\,\text{mA}$ and 25°C. The wavelength age distribution is lognormal with an activation energy of 0.7 eV, and an 85°C median age rate of 22 pm/khr. The positive wavelength shift is due to an increased optical loss with aging, requiring a higher threshold carrier density to achieve lasing. The aging blueshift is countered by a redshift due to additional current required to keep the laser operating in constant power mode. The net effect is still a blueshift, and the calculated hazard rates are acceptable for 100 GHz channel spacing (Wilt, 1998). For tighter spacing, a closed-loop-controlled wavelength locker is required to achieve acceptable wavelength stability.

6.3 DIRECTLY MODULATED LASERS FOR 10 Gb/s TRANSMISSION

Typical directly modulated laser transmission systems use lasers in which the laser relaxation oscillation frequency is a minimum of twice the bit rate, ensuring a minimum of two oscillations per bit period. Due to the intentionally limited bandwidth of the receiver (~0.7 × the bit rate, i.e., ~7 GHz for a

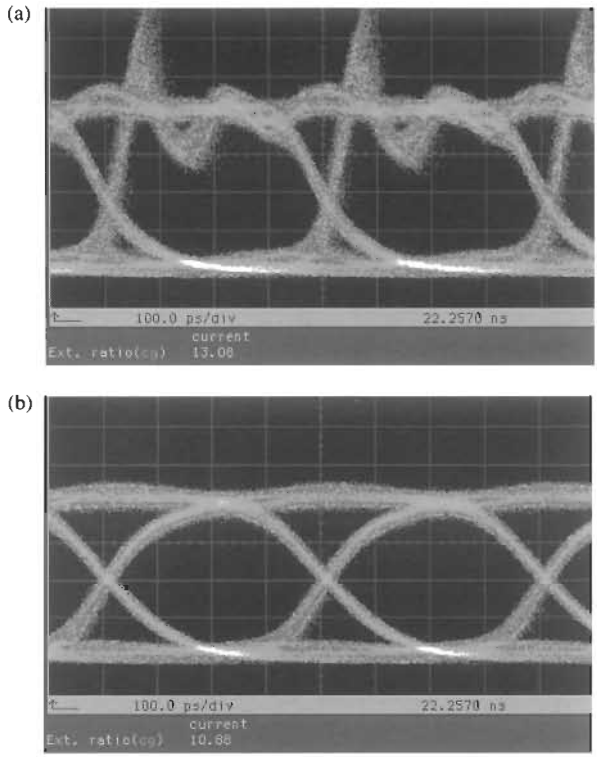


Fig. 22 Directly modulated 2.5 Gb/s eye patterns (a) before and (b) after filtering.

10 GHz system) these oscillations will be completely filtered out in the receiver. Figures 22(a) and (b) show the case for a 2.5 Gbit eye where the laser resonance frequency is in excess of 5 GHz (Raddatz, 1999). The filtered eye is "wide open" and has no difficulty meeting the ITU eye mask requirements.

Using this traditional metric, a laser diode with an oscillation frequency in excess of 20 GHz over the complete system temperature range is required for 10 Gb/s systems (typical minimum central office application temperature range is 0°C to +70°C). Although there have been demonstrations of long-wavelength lasers achieving this at room temperature (Morton et al., 1992) this requires modulation currents of the order of 80 mA or more. This is a typical maximum modulation current obtainable by 10 Gb/s laser drivers; thus with a cooled laser package it may be possible to realize a 20 GHz oscillation. However, for uncooled laser packages operating in constant power mode with a driver which has a maximum current swing of 80 mA, the maximum room temperature modulation current allowed is ~40 mA, which limits the oscillation frequency to substantially below 20 GHz with today's materials. Material systems with higher differential gain may permit this.

6.4 LASER CONSIDERATIONS

From a small-signal analysis of the laser rate equations, the laser oscillation frequency can be calculated to be

$$f_r \propto \sqrt{\frac{\partial G/\partial N \times P}{\tau_p}} \propto D \times \sqrt{I_{\text{mod}}} \tag{5}$$

where f_r is the laser resonance frequency, $\partial G/\partial N$ is the differential gain, P is the optical power, τ_p is the photon lifetime, D is the modulation efficiency, and I_{mod} is the magnitude of the current above threshold. This is true while the laser output is linear ($\sqrt{P} \propto \sqrt{I_{\text{mod}}}$).

It can be shown that D increases with increasing differential gain and decreasing photon lifetime. One way to increase the differential gain of a DFB laser is to increase the magnitude of the detuning between the emission wavelength and material gain peak (Yamanaka et al., 1993). The photon lifetime can be decreased by increasing the doping in the active region (Morton et al., 1992) or by shortening the laser cavity length.

Figure 23 shows the variation of the D-factor with both detuning and active region doping for 1310 nm DFBs. In this experiment the DFB active region comprised 970 Å strained quantum wells. Note that the plot includes data from a Fabry–Perot laser wafer with essentially the same active region (for this, the detuning is obviously 0). The expected trends can be observed—if a linear model is fit to the data, the D-factor (expressed in GHz/$\sqrt{\text{mA}}$) can be calculated to increase ~0.02 per nm increase in detuning and ~0.1 per 10^{17} increase in active doping.

As may be expected this increase in D-factor comes at the expense of increased laser threshold current and decreased quantum efficiency. The threshold current (in mA) increases by ~0.17 per nm increase in detuning and ~0.44 per 10^{17} increase in active doping, while the quantum efficiency (in W/A) decreases by ~0.01 per nm increase in detuning and ~0.026 per 10^{17} increase in active doping. This increase in threshold and decrease in quantum efficiency is especially important in lasers intended for uncooled applications.

Figure 24 shows the effect of decreasing cavity length on 1310 nm Fabry–Perot lasers. As can be seen, there is an increase in the D-factor from ~1.4 GHz/$\sqrt{\text{mA}}$ for a 300 μm cavity to ~1.8 GHz/$\sqrt{\text{mA}}$ for a 150 μm cavity. However, it can also be noted that the rate of increase in relaxation frequency is rolling over much faster for the 150 μm device, due to thermal effects. Figure 24(b) shows the maximum oscillation frequency is ~6 GHz for the 150 μm device (the 300 μm device has a maximum in excess of 10 GHz).

6.5 ESTIMATION OF MINIMUM REQUIREMENTS

An estimation for the minimum laser oscillation frequency required to allow 10 Gb/s transmission can be made by numerically solving the laser rate

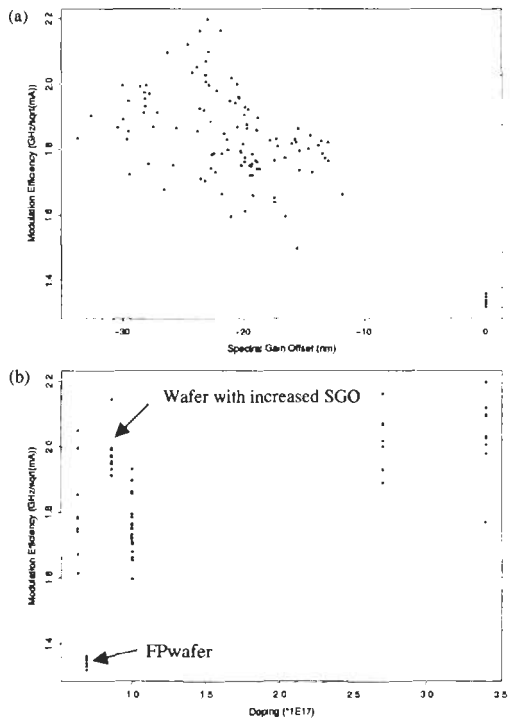


Fig. 23 Experimentally determined D-factor variation vs. emission/gain wavelength detuning (a) and active layer p-doping concentration (b).

equations. It is assumed that there is no requirement on laser overshoot, and the inner eye mask requirement is as stated in ITU G953. This simple model is applicable to single-frequency laser diodes.

The list below contains the values of constants and parameters employed:

A	$8.5 \cdot 10^{-16}\,\mathrm{cm}^2$	parameter	Differential gain coefficient
t_c	1.83 ps	calculated value	Cavity photon lifetime
Γ	0.07	constant	Optical confinement factor
ε	$3 \cdot 10^{-17}\,\mathrm{cm}^3$	parameter	Gain compression factor
N_{tr}	$1.88 \cdot 10^{18}\,\mathrm{cm}^{-3}$	parameter	Transparency carrier density
η	0.8	constant	Internal quantum efficiency
V	$1.89 \cdot 10^{-11}\mathrm{cm}^3$	constant	Volume of the active region
e	$1.6 \cdot 10^{-19}\,\mathrm{C}$	constant	Electron charge
v_{gr}	$8.1 \cdot 10^{9}\,\mathrm{cm/s}$	constant	Group velocity

The differential gain coefficient A, gain compression factor ε, and transparency carrier density N_{tr} were chosen as fitting parameters and were varied

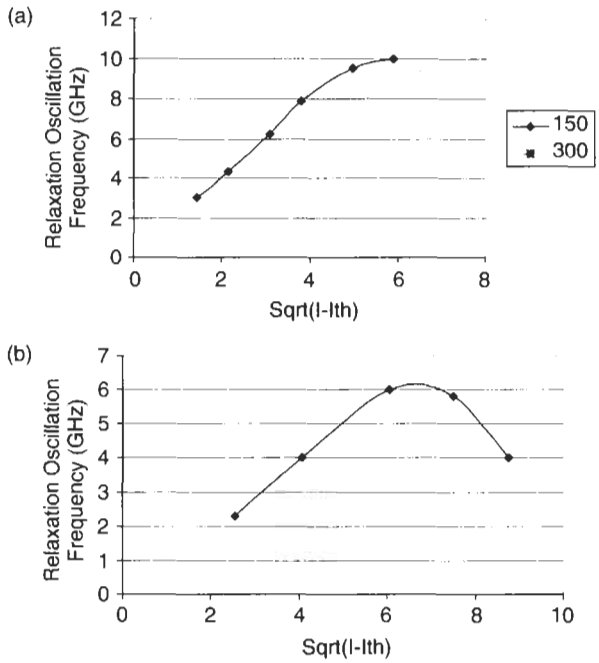


Fig. 24 *D*-factor is increased for shorter cavity length (a) but the rate of increase in the relaxation oscillation frequency decreases at higher currents in the short device due to thermal effects (b).

within 50% of measured values (Shtengel et al., 1998) in order to fit calculated laser characteristics to experimentally measured values. Variations in the differential gain coefficient strongly influence both dynamic and steady-state laser characteristics. This is the primary parameter used to tune the laser relaxation frequency, which is an essential factor determining laser performance at 10 Gb/s. The gain compression factor was used to fit damping of relaxation oscillations. Finally, calculated laser threshold and differential quantum efficiency were tuned to experimentally measured values by varying the transparency carrier density, which has almost no effect on dynamic laser characteristics such as relaxation frequency and damping factor.

The laser response to a large signal is obtained by numerically solving the system of rate equations, Eq. 1, using a standard Runge-Kutta method. As a pump current $J(t)$ a 7-bit-long stream of ones and zeros is applied while the laser is biased just above threshold. The laser output is calculated for 128 different data patterns, accounting for all possible permutations of ones and zeros in the 7-bit-long stream of data.

The effect of the limited bandwidth of an optical receiver is modeled by passing the calculated signal through a fourth-order Bessel-Thomson spectral filter $H(f)$ with a reference frequency $f_r = 7.5\,\text{GHz}$ [8]:

$$H(f) = \frac{105}{105 + 105i\dfrac{2.11f}{f_r} + 45\left(i\dfrac{2.11f}{f_r}\right)^2 + 10\left(i\dfrac{2.11f}{f_r}\right)^3 + \left(i\dfrac{2.11f}{f_r}\right)^4} \tag{6}$$

Figure 25 shows experimentally measured and calculated eye diagrams for a directly modulated Fabry–Perot laser operating at room temperature. In order to suppress excessive oscillations in the off-state in the modeled characteristics a β_{sp} of 0.01 is required—this is approximately two orders of magnitude higher than measured values from Shtengel et al., 1998, and shows the limitations of a simple model. However, the model should be sufficient to allow estimation of the minimum oscillation frequency required for 10 Gb/s operation.

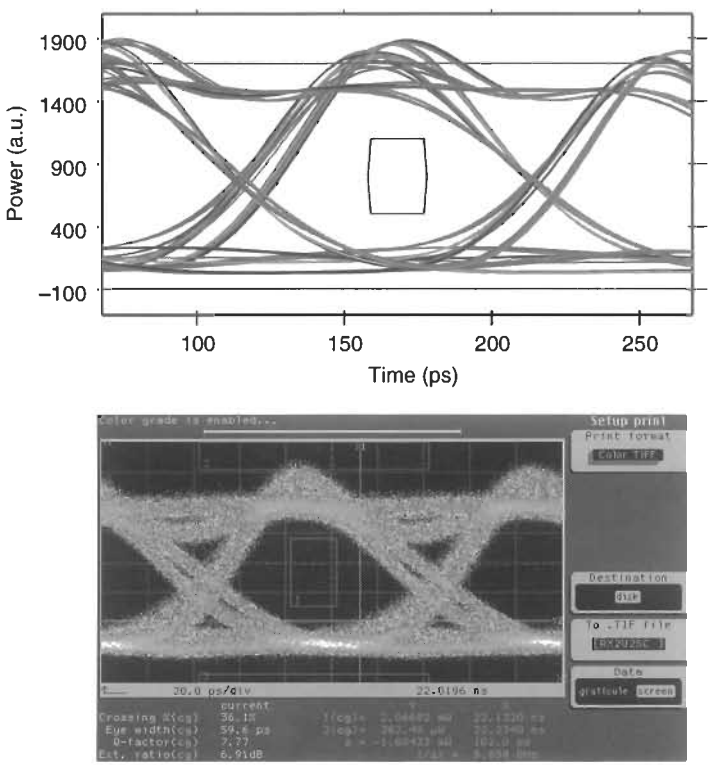


Fig. 25 Calculated (a) and measured (b) eye diagrams from 10 Gb/s directly modulated Fabry–Perot laser.

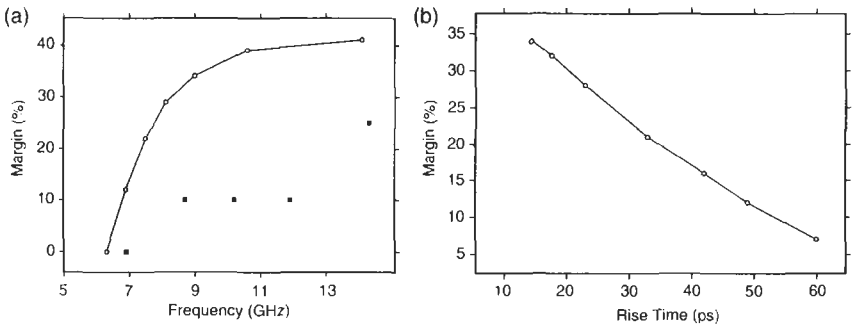


Fig. 26 Calculated (open symbol) and measured (closed symbol) 10 Gb/s eye margin as a function of relaxation oscillation frequency (a) and rise time (b).

Using the above model an estimate of the inner mask margin can be made as the laser oscillation frequency is varied (by varying current in the model). The margin is calculated by growing the inner eye until one point touches one of the modeled traces. Figure 26(a) shows the variation of eye margin with relaxation oscillation frequency. The line shows the modeled value, while the points show values measured for FP laser diodes (measured on submounts, with a 50 Ω matching probe). It can be seen that the calculation overestimates the margin but predicts the correct trend. The overestimation probably lies in a combination of four factors: the drive current is modeled to have a perfect rise/fall; the pattern length for the model is $2^7 - 1$, whereas the actual data is $2^{23} - 1$; the model assumes a perfect match to the laser diode, whereas there is ~7 dB return loss at 10 GHz for the submount; and the model does not take into account any noise/random fluctuations.

Figure 26(b) shows the degradation in eye margin as the laser driver rise/fall time degrades from 15 to >50 ps for a laser with an oscillation frequency ~11 GHz. It can be seen this has a significant impact on the eye margin, which falls from ~35 to ~12%.

From the above plots it can be seen that in order to have an eye with a reasonable margin (say 10% minimum), the laser oscillation frequency must be in excess of 9 GHz over all temperatures, and the driver should have a rise/fall time better than 40 ps.

7.0 Electroabsorption Modulated Lasers

The electroabsorption modulated laser (EML) monolithically integrates a light source and an electrooptic modulator. EMLs represent both an advance in digital laser performance and a milestone in the steadily maturing art and science of telecommunications laser fabrication. EML performance advantages originate from substantial dynamic wavelength chirp reduction afforded

by the use of electroabsorption, rather than direct, light modulation. Historically, EMLs are important in that they are the first mass-produced III–V electrooptic component with more than one optical element monolithically integrated on a single chip.

In this section we describe the basic EML physics and design considerations. We begin with a "black box" overview of the device to define essential operating parameters. We then focus on EA modulator physics, EML fabrication, and design and performance complications arising from monolithic integration of source and modulator. We conclude this section with the present state of the art and comments on future trends.

7.1 EML DEVICE OVERVIEW

The basic elements of an EML are depicted in Fig. 27. The device monolithically integrates two active and one passive waveguide sections (Kawamura et al., 1987). Starting from the left side of Fig. 27, these are a DFB laser source, an electrical isolation region, and an electroabsorption modulator. The laser and modulator have separate p-side contacts, but share a common n-side ground. In operation the DFB laser section is forward biased and continually on, producing a single longitudinal mode with optical power P_{in} emitted from the right hand side of this element. The light from the laser section passes through the passive waveguide in the isolation region. In a well designed device optical losses in the isolation region are low and can be neglected. P_{in} is then injected into the reverse biased modulator experiencing a loss, $\alpha(\lambda, V)$ that is both voltage and wavelength dependent. After traversing a modulator length L_{mod}, light P_{out} emitted from the device is given by

$$P_{out} = P_{in} \cdot \exp(-\alpha(\lambda, V) \cdot L_{mod}) \tag{7}$$

Data encoding is accomplished simply by transitioning the modulator between a transparent state and an opaque state by changing the applied reverse bias

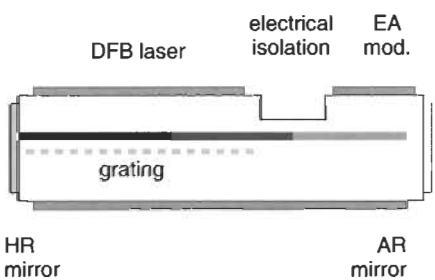


Fig. 27 Lateral view of the three sections of an electroabsorption modulated laser (EML). CW light is provided by the DFB laser on the left side. The center section provides electrical isolation with minimal optical losses. Modulation is achieved through reverse bias of the electroabsorption section.

from V_{on} to V_{off}, respectively. We now define some important EML operational parameters. The optical power lost in the on-state due to residual modulator absorption and scattering losses is referred to as insertion loss:

$$I_l = 10 \cdot \log[P_{on}(V_{on}, \lambda)/P_{in}] \tag{8}$$

The modulation voltage is defined as

$$V_{mod} = V_{off} - V_{on} \tag{9}$$

The optical power difference between the on-state and off-state for a given drive voltage is referred to as the extinction ratio $ER(V_{mod})$ and is expressed in dB.

$$ER(\lambda, V_{mod}) = 10 \cdot \log_{10}(P_{off}/P_{on}) \tag{10}$$

Using Eq. 7, extinction can be expressed using more fundamental modulator parameters, assuming that light scattered around the modulator waveguide and subsequently collected by the coupling optics is negligible:

$$ER(\lambda, V_{mod}) = -L_{mod} \cdot (\alpha(\lambda, V_{off}) - \alpha(\lambda, V_{on})) \cdot \log_{10} e \tag{11}$$

Figure 28 shows extinction characteristics of a typical 1550 nm EML device as a function of modulator bias and wavelength. For a 250 μm modulator some important values to note are: with 0 volts applied to the modulator, insertion loss is 2 to 4 dB and DC extinction ratio of better than 18 dB can be achieved with less than 2.5 V of modulation bias. Insertion loss and extinction ratio are strongly wavelength dependent, giving acceptable values in only a 15 nm to 20 nm range.

EML dynamic wavelength behavior is crucial in determining transmission quality in the presence of optical fiber chromatic dispersion. Emission

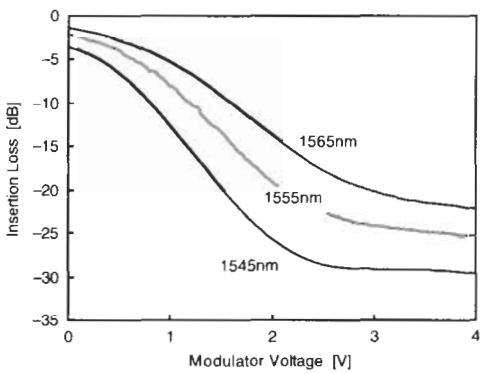


Fig. 28 Optical insertion loss of a typical 1550 nm EA modulator at three wavelengths. Insertion loss and extinction ratio decrease as the probe wavelength moves to larger values than the modulator bandgap (about 1515 nm for this device).

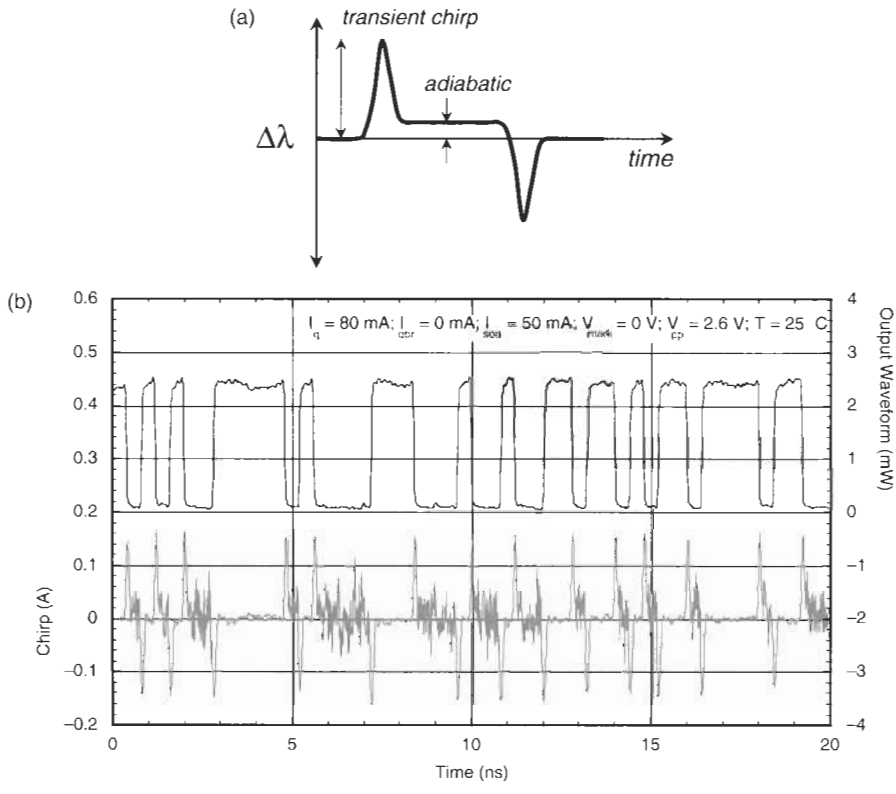


Fig. 29 (a) Schematic of wavelength deviations displayed by an EML under modulation. Transient chirp, due to self-phase modulation, is intrinsic to the modulator and occurs at the rising and falling edges. Adiabatic chirp occurs during the central portion of the pulse and can result from electrical or optical crosstalk between the modulator and DFB laser. (b) Actual 2.5 Gb/s chirp of an EML.

wavelength as a function of time is schematically shown in Fig. 29(a). Two features are noteworthy: emission wavelength shifts rapidly during both the off-to-on and on-to-off transitions, and a wavelength shift exists between the off-state and the steady-state region of the on-state. The first type of wavelength shift, measured from the off-state to the maximum wavelength during the transition, is referred to as transient chirp and is intrinsic to electroabsorption modulators. The second type, adiabatic chirp, is extrinsic resulting from interactions between the modulator and laser source, and possibly external coupling optics. Actual chirp data, measured using time-resolved transient spectroscopy (Linke, 1985) in Fig. 29(b), shows that transient chirp is about 0.1 Å, and adiabatic chirp is <0.05 Å for a 2.5 Gb/s pseudo-random bit stream. This chirp is more that an order of magnitude smaller than directly modulated lasers, which leads to longer span distances for 1550 nm EML sources.

Understanding EML optical amplitude and frequency modulation characteristics, and how to optimize them, requires a detailed explanation of the EA modulator physics and the electrical and optical interactions between the modulator and laser.

7.2 EA MODULATOR PHYSICS AND DESIGN

Electroabsorption modulators take advantage of the electric field-induced wavelength shift and broadening of the semiconductor optical absorption edge (Fig. 30). The device is designed so that in the on-state the emission wavelength is slightly longer than the absorption edge and thus experiences relatively small optical absorption at zero bias. When reverse biased the absorption edge moves to longer wavelength and broadens. The subsequent increase in absorption at the emission wavelength provides amplitude modulation. Absorption edge shift can be achieved through either the bulk layer Franz-Keldysh (FK) (Suzuki et al., 1992) effect, or the quantum confined Stark effect (QCSE) (Miller et al., 1984). In practice the QCSE provides higher extinction for a given reverse bias than the FK effect, and is thus used in most telecommunications-grade modulators. We therefore restrict our discussion to QCSE modulators.

The quantum confined Stark effect is observed in structures where charge carries are confined in narrow potential wells created by sandwiching low-bandgap material between higher-bandgap layers. In Fig. 31(a) a two quantum well modulator structure is illustrated in zero applied field. The effective bandgap energy is given by the sum of the quantum well bulk bandgap and the ground state energies of the electrons and holes,

$$E_{gap} = E_{bulk} + E_e + E_h \tag{12}$$

For the simple case of an infinite square well potential the quantum confined carrier ground state energy is simply expressed in terms of the carrier mass

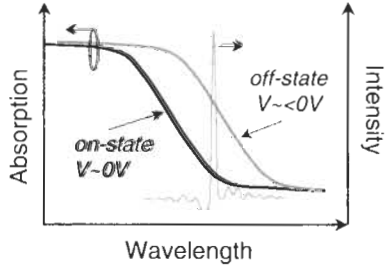


Fig. 30 Schematic of EA modulator bandgap bias inducted changes from on- to off-state. Through the quantum confined Stark effect the bandgap moves to longer wavelength as the reverse bias increases in magnitude.

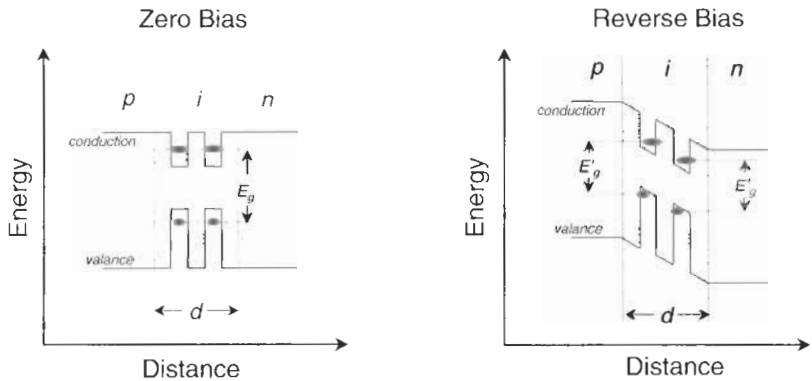


Fig. 31 Schematic view of electron and hole quantum confined states under reverse bias. Coloumb forces push electrons and holes to different sides of the quantum well, effectively reducing the bandgap energy with reverse bias.

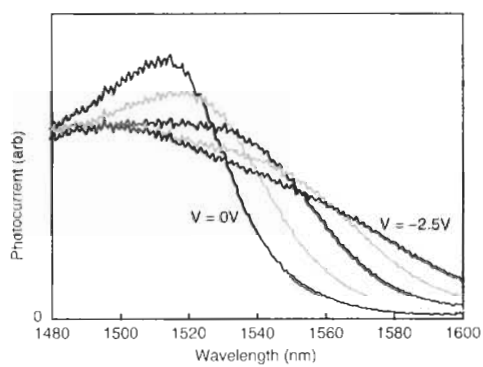


Fig. 32 Photocurrent EA measurements. The strong peak in the zero bias curve is due to exitonic effects. Ionization of the exiton washes out this feature under reverse bias.

$m_{e/h}$ and well width L_w:

$$E_{e/h} = (\pi\hbar)^2/2m_{e/h}L_w^2 \tag{13}$$

When reverse biased, the energy potentials tilt (Fig. 31(b)) causing electrons to be drawn to the right side of the quantum wells (toward the anode) and holes to the left side (toward the cathode). Just from geometrical considerations, it is clear that the new effective bandgap energy $E'_{gap} = E_{gap} - \delta E_{gap}$ is lower than the zero field gap and thus will absorb light in a modest wavelength band. Since absorbed light creates electron-hole pairs, absorption magnitude and wavelength dependence can be inferred from the resulting photocurrent (Bigan et al., 1991) (Fig. 32). Inspection of this data reveals that the absorption (photocurrent) vs. wavelength is nonmonotonic, displaying a distinct peak near the band edge. This feature results from a hydrogenic-like state

formed through Coulomb attraction between the photogenerated electrons and holes. Referred to as an exciton, this state significantly sharpens the band edge allowing DFB emission wavelength to be placed near the band edge without excessive absorption, and also producing high optical extinction with relatively small fields (Miller et al., 1985; Wood, 1988). Excitonic effects are an important advantage of the QCSE modulator, and their presence at room temperature is a direct consequence of carrier quantum confinement.

Inspection of Fig. 31(b) reveals several important aspects of QCSE physics and subsequent modulator design. First, just from geometrical considerations it is clear that the degree of bandgap shrinkage, $\delta E_{gap} = \delta E_e + \delta E_h$, increases strongly with increasing well width. From second-order perturbation theory, the energy shift for the infinite square well is

$$\delta E_{e/h} \propto m_{e/h} \cdot (V/d)^2 \cdot L_w^4 \tag{14}$$

exhibiting a quartic dependence on well width below ~200 Å (Wakita, 1998). However, thickening of the wells is not necessarily as advantageous as it might seem from this simple expression. The first issue is straightforward. Assuming the individual barrier layer thickness is kept constant, increasing the well width increases depletion layer thickness, yielding a smaller electric field for a given applied voltage. For a constant applied voltage, this results in a bandgap shrinkage that is closer to a square, rather than quartic, dependence on well width. In addition, in high fields, wider wells produce a larger physical separation between quantum confined electrons and holes, weakening their Coulomb attraction and hence suppressing excitonic effects. This reduction in exciton binding energy broadens the band edge, degrading extinction. A balance must therefore be struck between the need for wide wells to enhance field-induced bandgap shrinkage, and narrow wells to maintain strong excitonic effects at high fields. Well number is also critical in optimizing extinction characteristics. For an idealized extinction curve, where modulator loss is relatively voltage-insensitive at low bias (near 0 V) and asymptotically approaches a constant value at high reverse bias, it is easily shown that maximum extinction ratio, expressed in dB, increases linearly with quantum well number. However, this improvement is offset by degradation of both on-state insertion loss and modulation voltage, both of which can be shown, in dB, to linearly increase with well number.

Further inspection of Fig. 31(b) highlights the role proper barrier design plays in preserving excitonic effects at high electric fields. If barriers are either too thin or too low in bandgap energy, carrier tunneling from the well into continuum states occurs, ionizing the exciton. This effect, accentuated for electrons because of their low mass and higher ground state energy, is suppressed by widening the barriers and increasing their bandgap energy. However, while these alterations are effective at preserving the exciton, they may adversely effect the modulator extinction ratio and modulation speed. For

a given applied voltage, extinction ratio is degraded when the barrier width is increased due to increased depletion width and subsequent reduced electric field. Increased barrier bandgap energy can reduce modulation response due to difficulty in extracting carriers from a deep potential well. When this happens, carriers accumulate and screen the electric field seen by the quantum wells leading to reduced extinction ratio under high-speed modulation. In lattice-matched InP based quaternary quantum wells, carrier escape is primarily an issue for holes (Wood et al., 1990). Holes are substantially more massive than electrons and thus reside deeper in their respective potential wells. The 60/40 valence band/conduction band discontinuity accommodation in InP-based systems further compounds this problem. The massive holes sit in a deep potential well, while the light electrons ride high in a shallow well. This is opposite from what is ideal for simultaneously preventing exciton ionization and maintaining high-speed performance. In an effort to alleviate this problem, modulator barriers are generally in tensile strain, driving the bandgap discontinuity closer to a 50/50 split (Sahara et al., 1994).

While slow carrier escape can reduce modulation speeds, there is sufficient quantum well and barrier design flexibility to eliminate this as a limiting factor in both 2.5 and 10 Gb/s EML modulators. At these line rates, modulation response is RC limited. The relevant capacitances are the modulator p–n junction capacitance and parasitic capacitances originating from current confining layers and metal contact pads (Devaux et al., 1991). The series resistance arises from both the InP material placed on top of the modulator active waveguide and the p-ohmic contact made to the semiconductor. Parasitic capacitance is a property extrinsic to the modulator design and is strongly influenced by device structure choice (e.g., ridge vs. buried heterostructure). Junction capacitance and series resistance, on the other hand, are intrinsic modulator design issues that are varied through geometrical changes to the active waveguide width, thickness, and length. Waveguide width, however, is typically not a significant design variable, since the need to support a single lateral optical mode, concatenated with practical fabrication requirements, stringently constrain its value. Thus active thickness and length remain as the exercisable design variables. Improvement in EA modulator bandwidth is typically achieved through a decrease in modulator length, providing the dual benefit of reducing both junction and most parasitic capacitances. Along with this, however, comes increased series resistance and reduced extinction. The extinction ratio loss is compensated for by increasing the number of quantum wells, which also further reduces junction capacitance through an increase in depletion width. However, as discussed previously, increasing the number of quantum wells may adversely effect the modulation voltage needed to obtain maximum extinction ratio.

Reduced wavelength chirp during high-speed modulation is the major driving force behind the emergence of EA modulators. At first glance it might seem that the intrinsic wavelength chirp of an EA is zero, since varying the

EA modulator's loss does not inherently change the emission wavelength of the CW source. However, additional scrutiny reveals that while chirp is small, it is not identically zero. The electric field of a lightwave can be expressed as,

$$E(t) = E_o \cdot \exp(-i \cdot (\omega_o t + \phi)) \tag{15}$$

The instantaneous frequency ω_i is given simply by the time derivative of the exponential's argument:

$$\omega_i = \omega_o + d\phi/dt \tag{16}$$

Equation 16 shows the equivalence between a time-varying phase and changes to the emission frequency. Chirp due to this self-phase modulation exists in EA modulators because of the fundamental link between absorption and refractive index changes expressed in the well known Kramers–Kronig relations. When an EA modulator's loss is either increased or decreased, a simultaneous change in the refractive index occurs, altering the phase of the light propagating in the modulator, and therefore shifting its wavelength (Pollack, 1995, p. 200). This wavelength change occurs only during the on/off transients. A more intuitive understanding can be gathered by imagining an observer looking directly through the EA modulator to the DFB light source. During modulation, the refractive index shift accompanying the on/off transition changes the modulator's optical length. The observer therefore sees the DFB laser as either receding or approaching, depending on the sign of the index change, leading to a Doppler shift in the source emission wavelength. For an EA modulator, this shift is simply expressed in terms of amplitude modulation (Koyama and Iga, 1988):

$$\omega_i = \omega_o + (\alpha_H/(2 \cdot P_{out})) \cdot dP_{out}/dt \tag{17}$$

where α_H is Henry's linewidth enhancement factor, relating the real, n (refractive index) and imaginary, g (absorption) parts of the complex index:

$$\alpha_H = \partial n/\partial g \tag{18}$$

A few things are worth noting about Eq. 17. First, for positive α_H, EML emission wavelength increases during the on- to off-state transition and then returns to its on-state value after the transient is over, as is seen experimentally in Fig. 29(b). Second, reducing α_H is clearly important in minimizing EA modulator chirp. For typical EA modulators α_H is between -1.5 and $+2.0$. α_H is positive when the emission wavelength is much longer (>25 nm) than the on-state absorption edge, and as the emission wavelength is moved closer to the on-state absorption edge, α_H passes through zero and becomes negative (Devaux et al., 1993). While moving the emission wavelength closer to the absorption edge reduces chirp, this comes at the expense of higher modulator insertion loss and therefore reduced output power.

While the qualitative relationships between quantum well design and modulator insertion loss, extinction ratio, modulation speed, and wavelength chirp

are fairly straightforward, simultaneous optimization of all these parameters is not. As we have seen from the discussion preceding, varying one design parameter typically improves one aspect of performance while degrading others. In practice, design optimization is done semiempirically. Models and physical insight are used to carefully direct experiments from which design improvements are derived. This is a time-consuming and difficult procedure which is made even more onerous by modulator/laser interactions inherent in the monolithic integration of these two elements.

7.3 EML FABRICATION

EML devices can be realized using either ridge waveguide or etched and regrown buried heterostructure configurations as discussed in Section 4. While the latter structure is significantly more complex to fabricate, it has become the foundation upon which commercially successful EMLs are built owing to its greater design flexibility. Monolithic integration of both DFB laser and modulator active regions, required in any EML realization, can be accomplished using three different techniques: identical layer active, butt-joint technology (Oshiba et al., 1993), and selective area growth (Aoki et al., 1993; Thrush et al., 1994). As the name implies, DFB and modulator share the same active material in the identical layer active approach. This technique greatly simplifies device fabrication but also yields essentially no design space to separately optimize laser and modulator. As a result, performance is poor, and therefore this technique is not widely used. Butt-joint technology employs separate epitaxial growth steps for laser and modulator active formation. After the growth of one either laser or modulator active, the wafer is masked and epitaxial layers are etched away in unprotected areas. The remaining active structure is then regrown in the resulting void. Offering the ultimate in design flexibility, butt-joint allow independent optimization of both DFB and modulator actives. However, the need for extremely accurate vertical alignment of the two actives ($<0.2\,\mu m$) presents significant fabrication complexity and has therefore limited the use of this technology in commercial products. Selective area growth (SAG) represents a midpoint between the oppressive design constraints of the identical layer active approach and the difficulty of butt-joint technology. SAG enables laser and modulator actives to be formed in a single epitaxial growth. Oxide pads, placed on the wafer prior to growth, shift laser bandgap wavelength more than 100 nm beyond that of the modulator. In addition, perfect vertical alignment of the two actives is automatically achieved. The combination of fabrication ease and design flexibility has led most prominent EML vendors to select selective area growth.

7.4 EML LASER/MODULATOR INTERACTIONS

When EML lasers were first considered as practical light sources, the accepted device design philosophy was that a poor DFB laser monolithically integrated

with a high-performance modulator resulted in a high-performance EML. In the absence of laser/modulator interactions this view is largely correct. However, experience has demonstrated that management of these interactions both drives laser performance to approach that of the best discrete DFBs and is the prime factor in distinguishing a poor EML from one with world-class performance.

EML laser/modulator interactions come in two forms: electrical crosstalk between the modulator drive voltage and the laser section, and optical crosstalk resulting from parasitic optical reflections downstream of the modulator. We first describe electrical crosstalk issues, then move on to discuss optical interactions.

In a typical EML, the DFB laser contact is less than 200 μm from the RF terminal of the modulator. Any electrical leakage from the modulator contact to the laser results in direct modulation of the laser current and an ensuing wavelength chirp. In terms of the EA peak-to-peak modulation voltage, V_{mod}, this chirp $\delta\lambda$ is given simply by

$$\delta\lambda = \delta V_{mod} \cdot (\eta / Z_{iso}) \tag{19}$$

Requiring $\delta\lambda < 0.05$ Å, forces the isolation impedance Z_{iso} to exceed 800 Ω for a typical laser FM efficiency of $\eta = 200$ MHz/mA. DC isolation resistance of this magnitude is readily achieved using fabrication techniques described in Section 4. At the high frequencies ω_m used to drive the modulator, however, obtaining this degree of isolation is not straightforward. Careful attention to EML chip isolation region, chip submount, and package design is required. Small degrees of capacitive or inductive coupling between laser and modulator drive voltage can produce significant wavelength chirp. In addition to laser/modulator isolation region design, the effects of electrical crosstalk can be minimized by reducing laser FM efficiency (Agrawal and Dutta, 1986, p. 266),

$$\eta \equiv \delta\lambda / \delta I = (\alpha_H \cdot \tau_p / (4\pi q P_{in})) \cdot (\omega_m^2 + \Gamma_p^2)^{1/2} \tag{20}$$

Since the photon lifetime τ_p and photon decay rate in the resonant cavity Γ_p are not easily varied, reducing α_H becomes the most direct way to minimize FM efficiency. Techniques and limitations of laser section α_H control will be discussed shortly.

While EML electrical crosstalk is governed by both on-chip and off-chip sources, optical crosstalk is largely due to on-chip phenomena when output fiber coupling optics are properly designed. In an idealized EML, optical feedback in the DFB laser section comes from only the rear cleaved facet and the fine-pitch DFB grating. However, in a real device the modulator output facet has a finite reflectivity. From the perspective of the DFB laser, this residual reflectivity is seen when the modulator is in the on-state, and invisible in the off-state due to high modulator absorption loss. Thus the DFB mirror loss,

threshold gain, and hence emission wavelength, are different in the on- and off-states. While this adiabatic chirp is smaller in magnitude than transient chirp, it carries substantial optical energy since it occurs during the entire bit length. Minimizing adiabatic chirp is accomplished by reducing the magnitude of the residual reflection and by increasing the DFB laser's immunity to optical reflections.

Calculations show that modulator output reflections must be $<10^{-4}$ to keep adiabatic chirp in an acceptable range. Antireflection coatings of this quality are difficult to reproducibly achieve in a manufacturing environment. Incorporation of a diffraction window at the modulator output is often used to get around this limitation. With a diffraction window, the modulator waveguide ends 10–30 μm prior to the cleaved, AR coated, facet. This gap allows light emitted from the modulator to diffract within the chip on its way to the output facet and also on the return trip after hitting the AR coating. The added loss reduces the effective reflection tenfold or more.

Increasing the DFB's immunity to the parasitic reflection is also a productive means of suppressing adiabatic chirp $\Delta\lambda_{ad}$. It can readily be shown that

$$\Delta\lambda_{ad} = -\alpha_H \cdot (\Lambda/k_0) \cdot \Delta g_{th} \tag{21}$$

where Δg_{th} is the change in the laser threshold modal gain between the modulator on- and off-states, pitch of the fine period DFB grating is given by Λ, and α_H is the laser section linewidth enhancement factor.

Equations 20 and 21 highlight the importance of reducing laser section α_H to minimize the deleterious effects of both electrical and optical crosstalk. As described in Section 6 low α_H is achieved through increases in laser differential gain. Key design parameters are carrier quantum confinement, laser gain/emission wavelength detuning, active region p-doping, and cavity length increases. The degree to which these parameters can be controlled in an EML is intimately linked to the technology used to integrate the laser and modulator actives. In the case of identical layer active, only the device length can be exploited, leaving insufficient design flexibility to adequately reduce α_H. With butt-joint technology all the listed laser design parameters can be optimized for low α_H without degrading modulator performance. Using selective area growth, however, only some of these parameters can be freely adjusted. Laser active layer p-doping cannot be employed, since the resulting modulator active doping prevents the application of a uniform field across the quantum wells and therefore drastically degrades extinction characteristics. Carrier quantum confinement effects are somewhat compromised in the laser section, since SAG growth enhancement produces thick quantum wells whose differential gain is midway between a true quantum well and bulk active. Thus cavity length and emission/gain detuning remain as the dominant design variables for reducing α_H in a SAG EML. Cavity length optimization is straightforward. Emission wavelength detuning is controlled through the dimensions of laser section

oxide growth pads. In practice laser $\alpha_H \leq 3$ can be achieved in a carefully optimized SAG-based EML. While this is not quite as good as found in discrete DFB lasers, it is adequate to produce EMLs with excellent transmission performance.

7.5 EML PERFORMANCE

EMLs are the dominant light source for 2.5 Gb/s long-haul DWDM systems due to their compact size, low cost, and excellent transmission characteristics. Error-free transmission over more than 1000 km of standard optical fiber has been demonstrated using mass-produced EMLs, Fig. 33. In contrast, directly modulated sources are limited to spans of ~200 km.

At 10 Gb/s line rates, EMLs share the long-haul spotlight with $LiNbO_3$ external modulators. 10 Gb/s EMLs can transmit error-free over more than 80 km of standard fiber, Fig. 34. While this is close to the maximum distance for a transform-limited pulse, the somewhat more ideal and adjustable chirp characteristics of $LiNbO_3$ give it the edge in very long-distance communications. For shorter distances, such as interoffice or metro applications, the size and cost advantages of an EML give it the upper hand.

Advances in EMLs and EA modulators are currently being pursued along three axes: improved chirp performance, higher functionality, and increased line rate. Improving chirp characteristics will come from more complete theoretical models of EA modulators, subsequent improvements in MQW design, and refined epitaxial growth procedures needed for robust implementation of these modifications. Central to advancing EML functionality is the incorporation of emission wavelength selectability. As described in the next

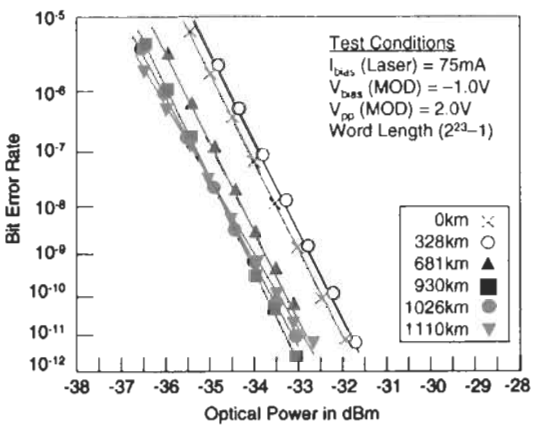


Fig. 33 State of the art EML can transmit over 1000 km of standard fiber without dispersion compensation. The negative transmission penalty results from pulse compression due to self-phase modulation within the fiber.

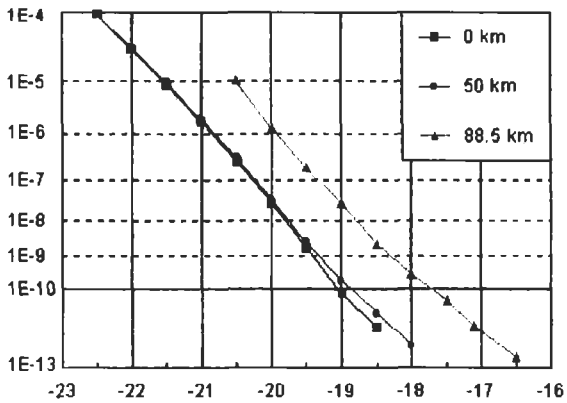


Fig. 34 At 10 Gb/s EMLs can transmit over 80 km of uncompensated standard fiber. This is the transmission limit expected for a purely envelope modulated pulse.

section, industrywide focus has been intense in this area, and great strides have been made toward the realization of a practical wavelength-selectable EML. Increasing EML and EA modulator line rates to 40 Gb/s is already well under way and promises to showcase an interesting competition between EA and $LiNbO_3$ technologies. In addition to cost and size advantages, EA modulators provide good performance at 40 Gb/s with only 2.5 V modulation voltage. $LiNbO_3$ modulators currently require more than 6 V single-ended modulation swings at 40 Gb/s. Given the challenges faced by 40 Gb/s electronics, this is a sizeable difference. In addition, the high spectral content inherent in short 40 Gb/s pulses makes the role of wavelength chirp unclear. Whether the chirp advantages and wavelength insensitivity of $LiNbO_3$ will outweigh the size and low RF drive advantages of EA modulators is an open question that will be answered in the next few years.

8.0 Wavelength-selectable Lasers

This section will cover the design growth and fabrication issues related to the development of wavelength-selectable semiconductor lasers. Wavelength-selectable sources, or tunable lasers as they are also known, are essential components for many current and next-generation fiberoptic communications systems. They can be employed as simple replacement sources in long-haul dense WDM transport systems to reduce cost and inventory complexity. They also can be used to provide advanced network functionality such as wavelength routing and dynamic bandwidth provisioning. There are many different types of tunable lasers; some use external cavity elements for wavelength control, whereas others are thermally or electronically tuned. Each type has its own unique set of characteristics that make it well suited to a particular application.

The most immediate benefit of tunable lasers is the ability to enable service providers to reduce inventory levels. Since different lasers must be kept in inventory for each channel, the problem of maintaining spares increases as the capacity of DWDM systems expands. For systems with 160 channels it becomes extremely difficult to manage. With a tunable laser a single device type could be used for every channel in the system. In the metropolitan environment, lasers that can be remotely programmed to change wavelengths can be deployed to add flexible bandwidth. To route data to a new location it would only be necessary to remotely tune the transmitting laser to a new wavelength. There are a number of companies that are currently pursuing transmission systems that take advantage of this wavelength routing capability that tunable lasers allow. The ultimate goal is the true all-optical network wherein individual packets of information could be routed in the wavelength domain using fast tunable lasers. Commercial products capable of this are still several years off, but the current generation of widely tunable lasers will be capable of supporting them.

Next-generation access networks based on WDM passive optical networks also represent a significant potential application for tunable lasers. The critical issue for current PONs is one of reducing the cost yet providing a scaleable architecture to enable the system installer to upgrade as capacity demands increase. Lucent Bell Laboratories has developed a network architecture that uses a tunable laser and a remote wavelength router to provide a time-shared WDM network (Frigo et al., 1994).

8.1 TYPES OF TUNABLE LASERS

8.1.1 DFB

Perhaps the simplest type of tunable laser is the distributed feedback laser (DFB). This device is designed to be a fixed-wavelength single-frequency source. The operating wavelength for a DFB laser is controlled by the period of the intracavity grating. This period is dependent on the effective index for the lasing mode. This index has a temperature dependence, which is governed by Eq. 47. For a typical DFB laser the thermal tuning coefficient is on the order of 0.1 nm/°C. Thus for a 5 nm tuning range a 50°C temperature excursion is required (Hong et al., 1999). In a conventional nonphase shifted AR/AR coated DFB with an index grating the cavity is antiresonant at the Bragg wavelength; thus the laser is forced to operate in one of two degenerate modes on either side of the stop band. In an asymmetric coated HR/AR DFB the degeneracy of these two modes is broken by the random phase of the facet reflections (Buss, 1987). To improve the single-mode yield for DFB an extra $\lambda/4$ phase shift can be introduced in the center of the AR/AR coated laser cavity. In this case the device is resonant at the Bragg wavelength and the degeneracy between the two modes is removed. Unfortunately, these devices are somewhat more difficult to fabricate, and they require very low facet reflectivity for

good single-mode yield. There are other shortcomings of using a $\lambda/4$ phase shift grating including the spatial hole burning-related unstable single-mode operation and the output power waste from the laser rear facet (Soda et al., 1987). The most promising technology for high single-mode yield in a DFB laser is the use of gain- or loss-coupled gratings. The gain coupling structure breaks the degeneracy between the two side modes and forces the device to lase predominantly on the longer side of the stopband regardless of the random laser facet reflection phases (Li et al., 1993). These gain-coupled lasers are the best suited for thermal tuning, although the maximum tuning range is limited by the temperature range over which acceptable lasing performance can be achieved, typically on the order of 2 to 3 nm though in some cases as much as 5. The main advantage of using a thermally, tuned DFB laser is that it has stable, well-behaved characteristics and requires only rudimentary control for wavelength stabilization. To achieve increased tuning ranges; several groups have investigated combining multiple thermally-tuned DFB lasers in an array (Hong et al., 1999; Ketelsen et al., 2000). Hong et al. have demonstrated 15 nm continuous wavelength tuning by using a laser with three separate inline DFB sections. Each section has a different Bragg grating pitch, which allows it to cover a different wavelength range. Wavelength selection is achieved by pumping the appropriate section for the wavelength range of interest and then thermally tuning to the final value. The other two sections are either left unbiased or are pumped to transparency if they are in front of the operating section. Ketelsen et al. took a different approach, using an array of five $\lambda/4$ DFBs in parallel with monolithically integrated spot size converters (Fig. 35). Phase-shifted DFBs were used to ensure high single-mode yield of the entire array chip. These were combined into a single output using a silica-based planar lightwave circuit and then coupled into an integrated amplifier modulator chip. This hybrid device was capable of tuning over a range of about 9 nm with a 20°C temperature variation. The high degree of complexity, large chip size, and the long time constants associated with thermal tuning have led the drive to more advanced tunable lasers.

8.1.2 DBR

One of the earliest types of electronically tunable lasers is the distributed Bragg reflector laser (Fig. 36). In this device a waveguide Bragg grating is used to provide a narrowband reflector for half of the laser resonator. The cleaved front facet of the device acts as the second mirror. The grating is formed by etching a periodic corrugation into the waveguide layer. This results in a periodic index variation for the guided mode. This periodic index variation provides coupling between the forward and backward traveling waveguide modes at a wavelength corresponding to the Bragg period. The reflection spectrum of a Bragg grating can be determined analytically using coupled mode theory or numerically using a transmission matrix analysis. The Bragg wavelength λ_B

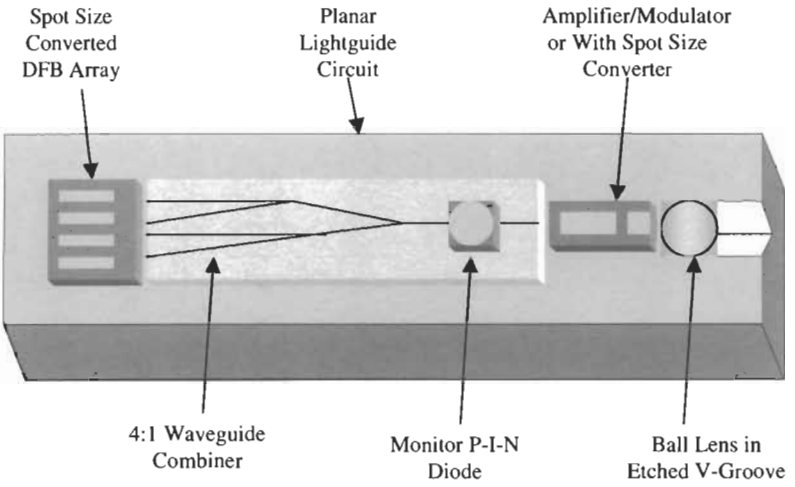


Fig. 35 DFB Laser Array. InP DFB array produces light at four different wavelengths. Glass on Si waveguides combine the light then couple it into an amplifier/modulator chip that boosts the optical power and encodes data. Wavelength tuning is accomplished through array element selection and temperature changes.

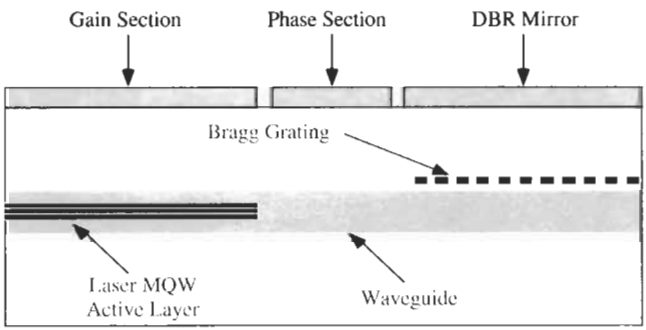


Fig. 36 Cross-section of three-segment distributed Bragg reflector laser. The gain section on the left provides optical gain, while the grating on the right provides a wavelength selectable reflection. The central section is adjusted to assure that the roundtrip phase condition is met.

for a grating with a period Λ is given by Eq. 22 where l is the order of the grating and $\bar{n}$ is the effective index of the waveguide mode.

$$2\beta_B = l\frac{2\pi}{\Lambda} \quad \text{or} \quad \lambda_B = \frac{2\Lambda\bar{n}}{l} \tag{22}$$

Injecting carriers into the waveguide lowers its index of refraction, which reduces the effective index of the mode. This causes the Bragg peak to tune to shorter wavelengths. In this type of tuning, the maximum change in the laser

wavelength is directly proportional to the maximum change in the effective index the so-called $\Delta n/n$ limit. The coupling coefficient and the mirror reflectivity for a grating of length L_g can be calculated from coupled mode theory using Eqs. 23 and 24 respectively, where U represents the optical mode profile and $\Delta\varepsilon$ is the periodic index perturbation caused by the grating. For Eq. 24 δ is the normalized detuning factor represented by $\beta - \beta_B$.

$$\kappa_{\pm l} = \frac{\kappa_o^2}{2\beta} \frac{\int \Delta\varepsilon_{\pm l}(x,y)|U|^2 dA}{\int |U|^2 dA} \tag{23}$$

$$\tilde{r}_g = -j \frac{\tilde{\kappa}_{-1} \tanh\left(\tilde{\sigma} L_g\right)}{\tilde{\sigma} + j\tilde{\delta} \tanh\left(\tilde{\sigma} L_g\right)}, \quad \tilde{\sigma} = \tilde{\kappa}^2 - \tilde{\delta}^2 \tag{24}$$

DBR lasers use a long grating with a weak coupling coefficient κ. This gives the mirror a high reflectivity and a narrow reflection spectrum. Since the laser operates at the minimum threshold gain it will tend to lase at the wavelength corresponding to the longitudinal cavity mode, which is nearest to the Bragg wavelength. To achieve continuous wavelength coverage, a phase control section is typically added to these lasers. This is an additional passive waveguide section, which can be tuned to align the cavity mode to the Bragg peak. This not only allows continuous tuning of the laser, it also increases the side mode suppression ratio and reduces the threshold current for a given wavelength. Fabrication of a DBR laser requires the integration of an active gain section, comprised of either quantum wells or bulk material having a bandgap wavelength around 1550 nm, with passive waveguide regions for the mirror and phase sections. The waveguide for the passive sections must have a bandgap energy, which is greater than the wavelength of the laser to minimize loss in these sections. However to optimize tuning response a low bandgap material is desirable. The best tradeoff between index tuning and increased loss with injected current is found with a waveguide having a bandgap wavelength around 1400 nm. DBR lasers are still not widely available commercially due mainly to the increased fabrication complexity associated with the formation of active and passive sections within a single laser cavity. This is a difficulty that affects most tunable lasers. The best DBR lasers that have been demonstrated to date have tuning ranges of 17 nm with output powers of greater than 100 mW (Delorme et al., 1997). They have also been integrated with on-chip semiconductor optical amplifiers and electroabsorption modulators to make complete tunable transmitter photonic integrated circuits (Fig. 7) (Ketelsen et al., 2000). The simple structure and monotonic tuning characteristic of these devices makes them attractive for telecom applications. However their relative tuning range $\Delta\lambda/\lambda$ is equal to the relative effective index in the waveguide $\Delta n/n$ which limits them to sub-20 nm tuning ranges. To achieve even greater wavelength coverage several new types of devices have been developed which can overcome this limit.

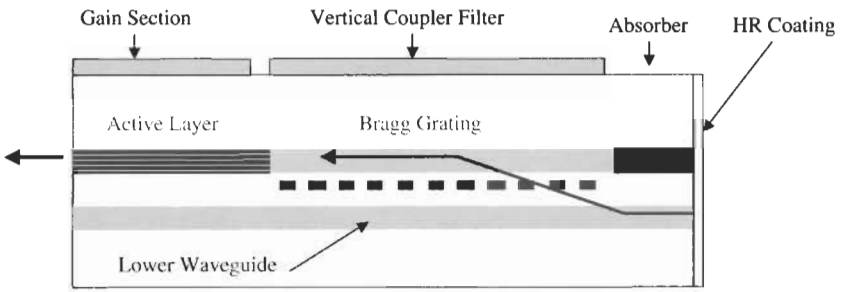


Fig. 37 Grating assisted codirectional coupler laser.

8.1.3 GACC Laser

The grating-assisted codirectional coupler GACC laser is fundamentally different from the DBR laser in design and structure but very similar in operation. The tuning element in the GACC is a pair of vertically stacked waveguides each made of a different material composition and thickness (Fig. 37). Because of the different propagation constants for the two waveguides, light will not couple evanescently between them. However by introducing a periodic index perturbation in the form of a grating, we can obtain efficient codirectional coupling between the two waveguides (Alferness et al., 1992). In this case, the period of the grating is given by Eq. 25, which is identical to Eq. 22, except that the coupling is now between the two waveguide modes and not between the forward and backward traveling modes of the same waveguide.

$$|\beta_1 - \beta_2| = l\frac{2\pi}{\Lambda} \quad \text{so} \quad \lambda_B = l\frac{\Lambda}{\Delta\bar{n}} \tag{25}$$

This subtle change leads to a large tuning range enhancement, since the wavelength shift is now given by Eq. 26, where γ is a bandwidth narrowing factor resulting from the difference in the dispersion of the two waveguides.

$$\frac{\delta\lambda}{\lambda} = \gamma\frac{\delta(\Delta\bar{n})}{\Delta\bar{n}} \quad \text{where} \quad \gamma = \left[1 - \Lambda\frac{\partial(\Delta\bar{n})}{\partial\lambda}\bigg|_{\lambda=\lambda_B}\right]^{-1} \tag{26}$$

The nominal tuning enhancement or F factor for this design versus a standard DBR is inversely proportional to the effective index difference between the two guides:

$$F = \gamma\frac{\beta}{\Delta\beta} \tag{27}$$

The smaller the index difference the greater the tuning enhancement factor and the wider the tuning range. Unfortunately, the bandwidth of the filter degrades as the tuning enhancement increases. This can be seen from Eq. 28, which gives

the full width at half maximum for the wavelength selective interwaveguide coupling.

$$\frac{\Delta\lambda}{\lambda} = 0.8\frac{\Lambda}{L}\gamma \tag{28}$$

Using this approach, Kim et al. (1994) have demonstrated a tuning range of more than 74 nm using a tuning enhancement factor of ~12. However, due to the large passband associated with the GACC filter element, this device had poor side mode suppression, between 25 and 34 dB depending on the wavelength, and was susceptible to mode-hopping. To improve the side mode suppression ratio and the stability of the device, an extra tuning element in the form of a sampled grating mirror can be added to the device to form a grating-assisted codirectional coupler with sampled reflector or GCSR laser.

8.1.4 GCSR

This device as the name indicates uses a grating assisted codirectional coupler within the laser as a frequency discriminator like the GACC, however it also includes a sampled grating DBR mirror (Fig. 38). The device is fabricated with two vertically stacked waveguides with different modal propagation constants. A long-period grating is fabricated on top of the upper waveguide, which provides coupling between these two guides for wavelengths at the Bragg frequency given by Eq. 25. Injecting carriers into one of the guides changes the index difference and tunes the coupling wavelength of the device. Because the Bragg wavelength is now dependent on the difference in index between the two guides, a much greater tuning range can be achieved with a given amount of index shift. This addition of the sampled grating mirror on the back of the laser compensates for the poor mode selection capability of the coupler. The sampled grating mirror consists of a grating that has been periodically sampled or blanked out to produce a set of equally spaced short grating bursts. The effect of this sampling is to produce a periodic set of reflection peaks that are equally

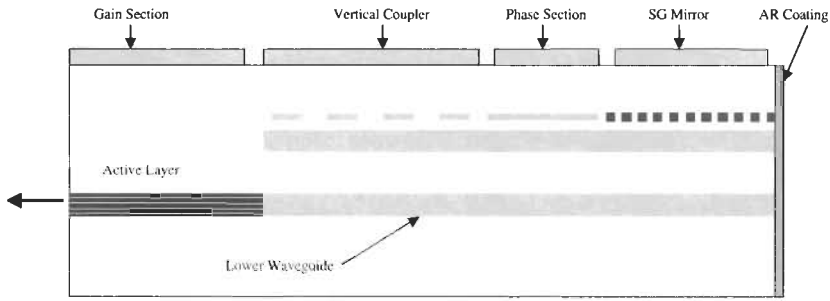


Fig. 38 Grating assisted codirectional coupler with sampled reflector laser.

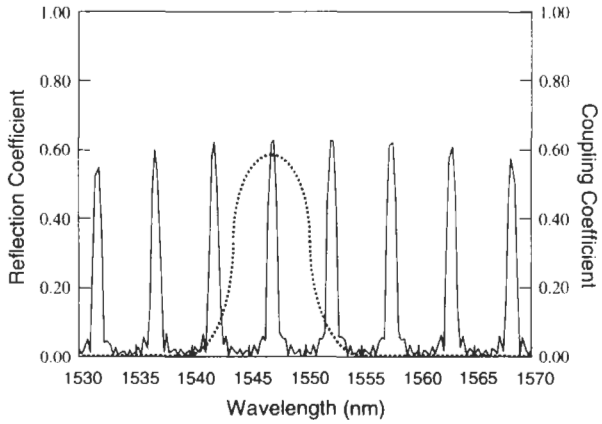


Fig. 39 Overlap of the GACC coupler bandwidth with the SG mirror spectrum.

spaced in frequency. The spacing of these peaks is inversely proportional to the sampling interval according to the relationship given in Eq. 29,

$$\Delta\lambda = \lambda^2/(2n_g\Lambda_s) \tag{29}$$

The reflection peak spacing is chosen so that only a single peak falls within the passband for the grating-assisted codirectional coupler. The reflection peaks for the sampled grating are very narrow so that typically only a single longitudinal mode falls within their reflection band (Fig. 39). This greatly improves the side mode suppression ratio for the device. The laser is tuned by aligning the GACC filter with a chosen peak of the SG-DBR mirror. The sampled grating mirror peaks can be tuned by current injection in the same manner as the mirror in a DBR laser. By successively tuning the GACC filter to different sampled grating peaks it is possible to cover a wide wavelength range with good SMSR. These lasers typically also employ a phase section, which facilitates alignment of the cavity mode with the sampled grating reflection peaks. Lasers of this type have been demonstrated with more than 40 nm tuning ranges (Oberg et al., 1995; Rigole et al., 1996). They are now available commercially with integrated wavelength control circuitry and are starting to be used in systems trials.

8.1.5 SG-DBR

The sampled grating DBR is similar in structure to a DBR laser except that it employs a Vernier tuning mechanism to increase its tuning range. The SG-DBR uses a pair of grating mirrors at either end of the cavity (Fig. 40). The gratings in the mirrors are periodically sampled or blanked out, which results in a sequence of equally spaced short grating bursts. These bursts give the

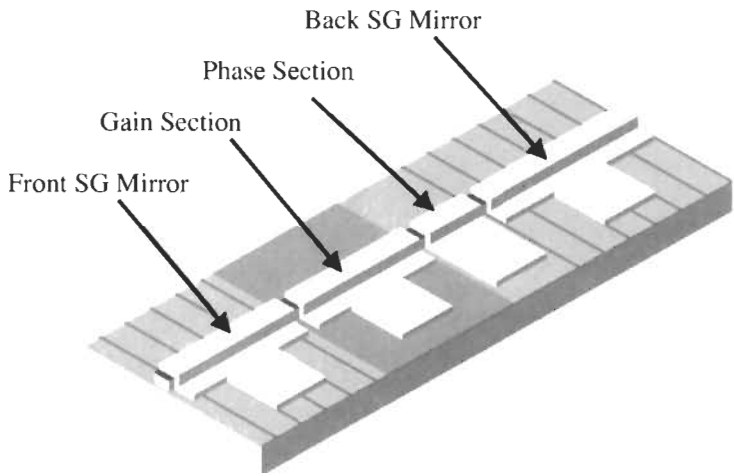


Fig. 40 Sampled grating distributed Bragg reflector.

mirrors a comblike reflection spectrum with multiple equally spaced peaks whose spacing is controlled by the sampling period. Different sampling periods are used in the front and back mirror to give them slightly different peak spacing. This allows only a single pair of peaks to be aligned concurrently, forcing the laser to operate at only this wavelength.

Just as in the DBR laser the mirrors can be tuned by carrier injection. This causes all of the peaks to shift together. Tuning the front and back mirrors in tandem enables the laser wavelength to be continuously tuned over a narrow range. By differentially tuning the mirrors a new set of peaks can be brought into alignment, allowing the laser wavelength to be tuned over a new range. Taking advantage of this Vernier tuning mechanism allows the laser to be tuned over a range of $M \cdot \Delta n/n$ where M is the number of peaks used. This gives it a much wider tuning range than is possible with the simple DBR laser. The theoretical limit on the number of peaks that can be used and thus the tuning range is determined by the repeat mode spacing. The maximum number of peaks is given by Eq. 30, where $\Delta\lambda_P$ is the peak spacing and $\delta\lambda_S$ is the difference in peak spacing between the front and back mirrors.

$$M = \frac{\Delta\lambda_P}{\delta\lambda_S} \tag{30}$$

The practical limit for $\delta\lambda_S$ is determined by the minimum side mode suppression ratio and depends on a number of mirror parameters. For a typical design, a minimum value of 0.5 nm is required for a side mode suppression ratio greater than 35 dB.

The mirror design is a crucial part of the SG-DBR laser. There is a wide variety of possible designs for any given tuning range. However the most efficient

design can be derived with a few simple rules. The best performance for a given tuning range will be achieved by using a mirror structure which yields the lowest total tuning current, since the loss increases with increased carrier injection into the tuning sections. The key design parameters for the mirrors are the grating burst length d, the sampling interval Λ_S and the number of sampling periods M. The reflection spectrum of a SG-DBR is the convolution of the reflection of a continuous grating with a periodic sampling function. Every element of the mirror spectrum has a corresponding element in physical space, which follows a reciprocal relationship. The real space sampling of the grating leads to multiple reflection peaks with a spacing that is inversely proportional to the sampling interval. The width of the reflection envelope for the mirror peaks is likewise inversely proportional to the burst length. As the burst length is increased, the reflectivity of the central peak increases; however the rolloff in the strength of the side peaks also increases.

The maximum tuning range for a given mirror design, with reflection peak spacing $\Delta\lambda_F$ and $\Delta\lambda_B$ for the front and back mirrors respectively, is limited by the repeat mode spacing (Jayaraman et al., 1993):

$$\lambda_{RMS} = \frac{\Delta\lambda_F \Delta\lambda_B}{\delta\lambda_S} \quad \text{where} \quad \delta\lambda_S = \Delta\lambda_F - \Delta\lambda_B \tag{31}$$

Typically for these lasers the peak spacing is on the order of 5 to 7 nm and the tuning enhancement factor is in the range between 10 and 15. This gives them a tuning range capability of between 50 and 100 nm. Devices with quasi-continuous tuning ranges of up to 74 nm have been demonstrated by Mason et al. (2000), and several groups have shown greater than 40 nm tuning ranges (Delorme et al., 1998; Ougier et al., 1996). The SG-DBR is the most prolific of the widely tunable lasers for a number of reasons. It has robust fabrication technology and, since it does not require a cleaved mirror, it can be monolithically integrated with external components like on-chip semiconductor optical amplifiers or electroabsorption modulators.

8.1.6 SSG-DBR

The superstructure grating DBR is a modification of the sampled grating DBR design. The two are functionally identical except for the design of the mirrors. In the SSG-DBR the gratings in the mirrors are periodically chirped instead of sampled, Fig. 41. As in the sampled grating DBR, the period of the chirp determines the peak spacing. The advantage of this approach is that the grating occupies the entire length of the mirror so that a much higher reflectivity can be achieved with a lower κ in the grating. The other main advantage is that the reflectivity of the individual peaks can be tailored such that all of the reflection peaks have the same magnitude and no spurious peaks are produced outside the tuning range. There are other novel designs for producing so-called super period DBR lasers, which involve using binary superimposed gratings

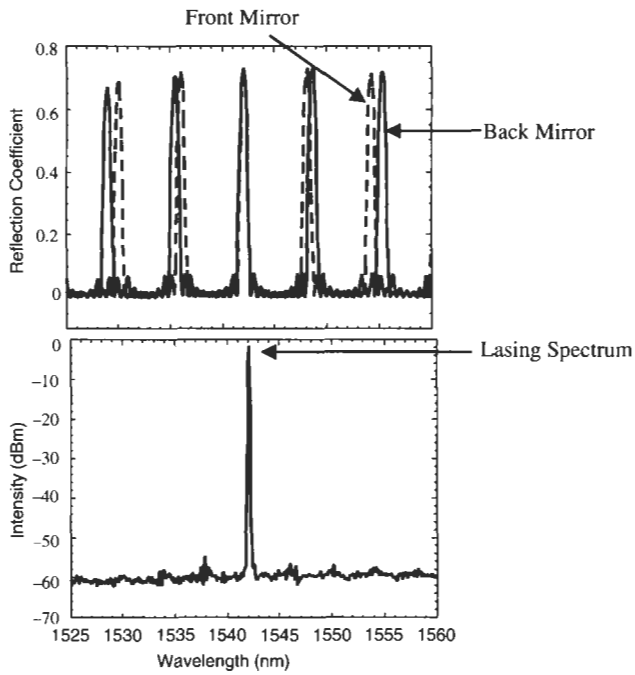


Fig. 41 SG-DBR tuning characteristics.

to form the mirror structure. In all of these it is evident that improved mirror characteristics can be obtained using an arbitrary grating shape function. However, it is not clear whether or not the performance improvements merit the increased complexity and cost required in the fabrication. Superstructure grating DBR lasers are also available commercially with tuning ranges of up to 35 nm and output powers of a few mW, although much wider tuning ranges have been demonstrated in research devices.

8.1.7 MFL

The multifrequency laser is more of a wavelength-selectable laser than a tunable laser, but it offers many of the same capabilities as a tunable laser. This device is based on an arrayed waveguide grating router, which acts as a dispersive element in much the same manner as a grating. Active sections in the input arms are pumped, and the light is coupled through the router to a common output. Only light at a wavelength that falls within the passband of the router for a given arm is coupled back to the input guide. This acts as a narrowband spectral filter for light in the cavity, resulting in a minimum loss at the peak of the router passband. Selecting the wavelength of the MFL is accomplished by pumping the input waveguide that corresponds to a given frequency band and then temperature tuning the device to adjust the wavelength within that range

(Doerr et al., 1996). The latest generation of MFL type lasers has demonstrated more than 40 wavelength channels, which are individually selected by using different combinations of gain sections located at either end of a specially designed arrayed waveguide grating cavity (Doerr et al., 1999). Due to the long length of this device, resulting in closely spaced cavity modes, side mode suppression is a challenge and direct modulation is limited to relatively low frequency (<1 GHz).

8.2 INDEX TUNING IN SEMICONDUCTOR WAVEGUIDES

There are two classes of effects that can be used for index tuning in semiconductors, field effects and carrier effects. The field effects include the linear electrooptic or Pockels effect and the quadratic electrooptic or Kerr effect. Both of these mechanisms produce a change in effective index with an applied field. This makes them more suitable for use in reverse-biased devices that require small index shifts. The field effects are not limited by carrier lifetime or carrier transport effects and thus can be used in applications requiring high speed but relatively low index shifts such as Mach-Zehnder modulators. Carrier injection changes the absorption characteristics of the material, which changes the index of refraction. There are four different components to carrier effect tuning. These include the contributions from free electrons, free holes, bandgap shrinkage and band filling. The amount of index tuning in a forward-biased semiconductor waveguide can be calculated by first determining the absorption change due to carrier injection, and then using the Kramers–Kronig relationship to calculate the associated index change. The free electron and free hole contributions can be calculated independently and then simply added in but the band gap shrinkage and band filling effects are interdependent and must be considered together.

8.2.1 Field Effect Mechanisms

The Pockels effect is anisotropic and gives a positive index shift for TE polarized light propagating along the [011] direction and a negative index shift for light propagating along the [011]. The TM polarization is unaffected by the Pockels effect. The index shift is proportional to the intensity of the electric field and is given by

$$\Delta n_p = \tfrac{1}{2} n^3 r_{41} E \tag{32}$$

where E is the applied field, n is the index of refraction, and r_{41} is the Pockels coefficient for InGaAsP/InP materials. The best measured value for this coefficient is 1.4×10^{-10} cm/V. For forward biased tuning the change in the electric field E is very low, and the index shift due to the Pockels effect can be neglected. The Kerr effect is independent of the polarization of the light and produces an index shift that is proportional to the square of the electric field:

$$\Delta n_K = \tfrac{1}{2} n^3 R_{Kerr} E^2 \tag{33}$$

The Kerr coefficient is dependent on the difference between the bandgap energy of the material and the photon energy of the guided light ΔE. It has been found to obey the empirical relation in

$$R_{Kerr} = 1.5 \times 10^{-15} \exp(-8.85\Delta E)\,\mathrm{cm}^2/\mathrm{V} \tag{34}$$

8.2.2 Carrier-induced Index Tuning

The carrier-induced change in the index of refraction for a material can be calculated from the change in absorption using Eq. 35, derived from the Kramers–Kronig relation where P denotes the principal value of the integral.

$$\Delta n_r = \frac{\hbar c}{\pi} P \int_0^\infty \frac{\Delta\alpha(E')}{E'^2 - E^2} dE' \tag{35}$$

The contributions from the free electron and free hole plasma effect can be calculated separately and then added to the index shift due to band filling and bandgap shrinkage. Throughout this section n_r will be used for the index of refraction, and N will be used for the electron density.

8.2.3 Free-Carrier Absorption

Free carriers can absorb photons and move to a higher energy state within a band. This intraband free-carrier absorption, also known as the plasma effect, can be modeled as being directly proportional to the concentration of electrons and holes and to the square of the wavelength. The corresponding change in refractive index is given by Eq. 36. This equation uses the fact that the concentrations of the heavy and light holes are proportional to their effective masses raised to the 3/2 power.

$$\Delta n_r = -\left(\frac{e^2\lambda^2}{8\pi^2 c^2 \varepsilon_o n_r}\right)\left[\frac{N}{m_e} + P\left(\frac{m_{hh}^{1/2} + m_{lh}^{1/2}}{m_{hh}^{3/2} + m_{lh}^{3/2}}\right)\right] \tag{36}$$

This formula gives the index of refraction change for the free holes and free electrons for all semiconductor materials lattice matched to InP. In general, the free carrier effects are small compared to the band filling and bandgap shrinkage effects; however they can make a significant contribution for optical wavelengths which are far from the band edge of the material.

8.2.4 Band Filling and Bandgap Shrinkage

The index shift due to the free electrons and free holes can be considered independently and simply added to the total shift. However, the bandgap shrinkage and band filling effects combine when carriers are injected into the semiconductor. The bands are filled as the bandgap shrinks so these two

effects must be treated together. For a direct bandgap semiconductor with parabolic bands the absorption spectrum (neglecting the Urbach tail) can be approximated by

$$\alpha_o(E) = \frac{C(y)}{E}\sqrt{E - E_g(y)} \quad \text{for} \quad E > E_g \tag{37}$$

The absorption constant C depends on the material and can be estimated for $In_{1-x}Ga_xAs_yP_{1-y}$ materials lattice matched to InP ($x \approx 0.47y$) using

$$C(y) = \left(1.004 - 1.318y + 0.517Y^2\right) \times 10^5 \ [\text{cm}^{-1}\text{eV}^{1/2}] \tag{38}$$

The absorption coefficient for a semiconductor with carrier injection is given by Eq. 39 if we consider the contributions from the light holes and heavy holes separately.

$$\begin{aligned}\alpha(N,P,E) = {} & \frac{C_{hh}}{E}\sqrt{E - E_g - \Delta E_g}\,[f_v(E_{vh}) - f_c(E_{ch})] \\ & + \frac{C_{lh}}{E}\sqrt{E - E_g - \Delta E_g}\,[f_v(E_{vl}) - f_c(E_{cl})]\end{aligned} \tag{39}$$

The ΔE_g term has been included to account for the bandgap shrinkage effect, which can be modeled using

$$\Delta E_g(N) = -\frac{0.13}{\varepsilon_s}\left(\frac{N}{N_{cr}} - 1\right)^{1/3} \tag{40}$$

where ε_s is the relative static dielectric constant of the material (in units of ε_o, the vacuum permittivity), and N_{cr} is a critical carrier density below which there is no shrinkage its value is given by Eq. 41 where m_c^* is the conduction band effective mass (in units of m_o).

$$N_{cr} = 1.6 \times 10^{30}\left(\frac{m_c^*}{1.4\varepsilon_s}\right)^3 \ [\text{m}^{-3}] \tag{41}$$

The heavy hole and light hole absorption constants can be determined from the experimental absorption constant $C(y)$ and the density of states effective masses.

$$\begin{aligned} C_{hh} &= \left(\frac{\mu_h^{3/2}}{\mu_h^{3/2} + \mu_l^{3/2}}\right)C(y) \qquad \mu_h = \left(\frac{1}{m_c} + \frac{1}{m_{hh}}\right)^{-1} \\ C_{lh} &= \left(\frac{\mu_l^{3/2}}{\mu_h^{3/2} + \mu_l^{3/2}}\right)C(y) \qquad \mu_l = \left(\frac{1}{m_c} + \frac{1}{m_{lh}}\right)^{-1} \end{aligned} \tag{42}$$

The probabilities f_c and f_v are given by the Fermi-Dirac distribution functions, where F_c and F_v are the quasi-Fermi levels for the conduction band and valence band respectively.

$$f_c(E) = \left(1 + \exp\left[\frac{E - F_c}{k_B T}\right]\right)^{-1}$$
$$f_v(E) = \left(1 + \exp\left[\frac{E - F_v}{k_B T}\right]\right)^{-1} \tag{43}$$

The Fermi levels can be obtained from the carrier densities by solving Eq. 44 where E_c and E_v are the potential energies of the conduction and valence band edges respectively.

$$N = \frac{1}{2\pi^2}\left(\frac{2m_c}{\hbar^2}\right)^{3/2} \int_{F_c}^{\infty} \frac{\sqrt{E - E_c}}{1 + e^{(E-F_c)/k_B T}} dE$$
$$P = \frac{1}{2\pi^2}\left(m_{lh}^{3/2} + m_{hh}^{3/2}\right)\left(\frac{2}{\hbar^2}\right)^{3/2} \int_{F_v}^{\infty} \frac{\sqrt{E - E_v}}{1 + e^{(E-F_v)/k_B T}} dE \tag{44}$$

The four transition energies used in Eq. 39 can be determined by momentum conservation and are given in Eq. 45. The conduction band and valence band energies are given relative to the respective band edges.

$$E_{ch} = (E - E_g - \Delta E_g)\left(\frac{m_{hh}}{m_{hh} + m_c}\right) \qquad E_{cl} = (E - E_g - \Delta E_g)\left(\frac{m_{lh}}{m_{lh} + m_c}\right)$$
$$E_{vh} = (E - E_g - \Delta E_g)\left(\frac{m_c}{m_{hh} + m_c}\right) \qquad E_{vl} = (E - E_g - \Delta E_g)\left(\frac{m_c}{m_{lh} + m_c}\right) \tag{45}$$

An excellent global approximation to Eq. 44 was developed by Nilsson which can be used to quickly calculate the Fermi level from the carrier density (Eq. 46). It is complicated but has less than 0.06% error over the entire range of carrier densities.

$$\left(\frac{F_c - E_c}{k_B T}\right) = \frac{\ln(N/N_c)}{1 - (N/N_c)} + \frac{f_h}{1 + (0.24 + 1.08 f_h)^{-2}}$$
$$\text{where} \quad N_c = 2\left(\frac{m^* k_B T}{2\pi\hbar^2}\right)^{3/2} \quad \text{and} \quad f_h = \left(\frac{3\sqrt{\pi}}{4}\left(\frac{N}{N_c}\right)\right)^{2/3} \tag{46}$$

Using the set of relations defined in this section, the change in the absorption spectrum can be calculated for a given change in the carrier density. Then by inserting Eq. 39 into the Kramers–Kronig relation (Eq. 35), and integrating numerically we can determine the corresponding refractive index change.

8.2.5 Thermal Compensation

The carrier-induced index change in a semiconductor waveguide can be partially compensated by the temperature change that occurs due to device self-heating. The temperature dependence of the index of refraction can be modeled for InGaAsP-based materials using Eq. 47 where the high-frequency dielectric constant is given by Eq. 48, as a function of the y parameter and its variation with temperature is given by Eq. 49 (Sundaresan and Henning, 1991).

$$\Delta n_r = \left(\frac{1}{\varepsilon_o} \frac{\partial \varepsilon_\infty}{\partial T} (T - 300) \right) \tag{47}$$

$$\varepsilon_\infty = (9.55 + 2.2y)\varepsilon_o \tag{48}$$

$$\frac{\partial \varepsilon_\infty}{\partial T} = 5.16 \times 10^{-4} \varepsilon_o \quad [\mathrm{K}^{-1}] \tag{49}$$

Estimation of the temperature change in the device as a function of current density can be done by first calculating the power dissipation using Eq. 50, where V_D is the ideal diode voltage and V_S is a current-independent series voltage. The optical output power from the tuning sections is neglected. This result is then multiplied by the thermal impedance, which can be modeled using the simple analytic approximation in Eq. 50 for a waveguide of width w and length l having a total substrate thickness of h (Coldren and Corzine, 1996). For InP the thermal conductivity ξ is about 0.6 W/(cm · K). The series resistance for a typical buried heterostructure device with a width of 2 μm, a length of 350 μm, and a substrate thickness of 100 μm is around 5 Ω.

$$P_D = I^2 R_S + I V_D + I V_S \tag{50}$$

$$Z_T \approx \frac{\ln(4h/w)}{\pi \xi l} \tag{51}$$

8.2.6 Bandgap Dependence of Index Tuning

Using the analysis developed in the previous section it is possible to calculate the tuning efficiency for materials of different bandgap. For an optical wavelength of 1.56 μm, as would be expected, there is a considerable increase in the index tuning for materials that have a smaller separation between the bandgap energy and the photon energy. It is therefore desirable to use a fairly low bandgap material for the waveguide to enhance the tuning efficiency. However it cannot be too close to the photon energy, because there is a significant increase in the waveguide loss as well. The best ratio of index tuning to absorption increase for the 1.55 μm wavelength range is achieved by using a material with a 1.42 μm bandgap. For optimum device design it is preferable to use the highest bandgap waveguide, which can provide sufficient index shift to meet the requirements, since this will minimize the internal waveguide losses.

8.2.7 Recombination

The carrier density in the tuning waveguide is a function of the current density and the recombination rate $R(N)$ that is given by Eq. 52. This equation includes the contributions from nonradiative recombination AN, bimolecular recombination BN^2 and Auger recombination CN^3.

$$R(N) = AN + BN^2 + CN^3 \tag{52}$$

Current estimates for the coefficients in InGaAsP materials lattice matched to InP are given below.

$$A \approx 1 \times 10^8\,[\mathrm{s}^{-1}] \qquad B \approx 2 \times 10^{-10}\,[\mathrm{cm}^3\mathrm{s}^{-1}]$$
$$C \approx 2 \times 10^{-29}\,[\mathrm{cm}^6\mathrm{s}^{-1}] \quad \text{at} \quad 1.3\,\mu\mathrm{m}$$
$$C \approx 5 \times 10^{-29}\,[\mathrm{cm}^6\mathrm{s}^{-1}] \quad \text{at} \quad 1.4\,\mu\mathrm{m}$$
$$C \approx 7.5 \times 10^{-29}\,[\mathrm{cm}^6\mathrm{s}^{-1}] \quad \text{at} \quad 1.55\,\mu\mathrm{m}$$

8.3 TUNABLE LASER SUMMARY

The tremendous progress made in tunable laser development, both in research and commercialization, has finally made widely tunable lasers a practical reality. Devices with tuning ranges in excess of 100 nm are now possible, and the dream of systems that route data in the wavelength domain now seems to be in reach. Currently tunable lasers are just beginning to be deployed in long-haul DWDM systems. This early deployment is being led by DBR lasers capable of covering ten to twenty channels; these will soon be joined by more widely tunable devices capable of covering 80 to 100 channels. Right now, tunable lasers represent a small portion of the market for long-haul transmitters; however in five years it is likely that they will be the standard source, replacing DFB lasers, which currently dominate the market. Early applications for tunable lasers are primarily for inventory reduction in large DWDM systems. However, many new companies are aggressively targeting the wavelength routing and bandwidth provisioning system applications, which tunable lasers enable. With so much progress having been made both in the design and fabrication technologies for tunable lasers, it is difficult to see where the new innovations will come from. Current devices are capable of tuning over hundreds of channels, have output powers as high as 40 mW, and can be fabricated with integrated optical amplifiers and modulators. The next great challenge for tunable lasers and the area from which the really important innovations will come from will be in fast wavelength switching for optical packet routing. Most of the work in this area will be in the control and stabilization electronics and systems for

fast wavelength switching. If these challenges are met, we will see an even more rapid advance in the capabilities of fiber networking systems enabled by advanced tunable lasers.

9.0 Summary

Telecom lasers have progressed along three axes: lower cost, higher modulation rates, and increased functionality. Lower cost has driven technology toward wide temperature range operation and increased manufacturing yields. As a result, lasers that were considered beyond the state of the art a few years ago are commodity items today. Further progress along this axis is likely to require improved understanding of laser thermal properties, more exotic material systems, and advances in manufacturing yields. Increased modulation rate brings many challenges. In general, device structures that work well at 2.5 Gb/s can be extended to 10 Gb/s but not to 40 Gb/s. In response to this, many R&D institutions are now developing low-capacitance device architectures. Pushing even these designs beyond 40 Gb/s is likely to require traveling wave configurations or on-chip optical multiplexing. In either solution, device complexity, size, and level of monolithic integration will increase. The axis of increased functionality, however, is perhaps the most interesting. Along with higher functionality comes increased monolithic integration, need for closed loop electronic control, and architectural uncertainty. An excellent example of these issues is the development of tunable lasers. The commercially available EA-DBR described in Section 8 monolithically integrates 5 separate components, a 2.5-fold increase in the current manufacturing state of the art. Controlling this complex device requires 3 electronic feedback loops, a threefold increase in control technology. While monolithic integration and closed-loop control distinguish tunable laser technology from components with lower functionality, the issues of architectural uncertainty is the most profound difference. As component functionality increases, so does the number of ways to achieve this functionality. This is clearly evident in tunable lasers. Although research in this field is well over a decade old and the global sprint toward a commercially viable device has been under way for 5 years, there is still no consensus on the preferred basic device architecture. This is in stark contrast to historical telecom laser evolution, in which near-global agreement was reached as a device transitioned from research to development, and development focused primarily on implementation details and fine levels of performance improvement. This level of architectural uncertainty is likely to increase as optical components assume more functionality, as in the complex devices designed to regenerate signal solely through optical processes. While this uncertainty results in lack of R&D clarity, it also serves to make optical sources a dynamic field in which thorough understanding and control of device physics and fabrication is essential.

Acknowledgments

The authors would like to thank D. P. Wilt, H. Blauvelt, M. S. Hybertsen, V. Kozlov, R. L. Hartman, and W. A. Asous for helpful contributions to this chapter.

References

Ackerman, D. A., Zhang, L., Morton, P. A., Shtengel, G. E., and Sheridan-Eng, J. A. (1996). "Real-world distributed feedback lasers: Measured, modeled and applied to CATV," presented at *Physics and simulation of optoelectronic devices IV*, San Jose, CA.

Agrawal, G. P. (1997). *Fiber-Optic Communication Systems*, John Wiley and Sons.

Agrawal, G. P. and Dutta, N. K. (1993). *Long-Wavelength Semiconductor Lasers*, Van Nostrand Reinhold, New York.

Alferness, R. C., Koren, U., Buhl, L. L., Miller, B. I., Young, M. G., Koch, T. L., Raybon, G., and Burrus, C. A. (1992). "Broadly tunable IlnGaAsP/InP laser based on a vertical coupler filter with 57.1 nm tuning range," *Appl. Phys. Lett.*, Vol. 60, No. 26, 1992.

Amann, M.-C. and Buus, J. (1998). *Tunable Laser Diodes*, Artech House, Boston.

Aoki, M., Suzuki, M., Sano, H., Kawano, T., Ido, T., Taniwatari, T., Uomi, K., and Takai, A. (1993). *IEEE J. Quantum Electron.*, Vol. 29, No. 6, pp. 2086–2096.

Asada, S., Sugou, S., Kasahara, K., and Kumashiro, S. (1989). "Analysis of leakage current in buried heterostructure lasers with semiinsulating blocking layers," *IEEE J. Quantum Electron.*, Vol. 25, No. 6, pp. 1362–1368.

Atlas, D. A., Pidgeon, R., and Little, F. (1995). *IEEE J. Lightwave Tech.*, Vol. 13, No. 5, pp. 933–946.

Becker, P. C., Olsson, N. A., and Simpson, J. R. (1999). *Erbium-doped fiber amplifiers: fundamentals and technology*, Academic Press.

Belenky, G. L., Kazarinov, R. F., Lopata, J., Luryi, S., Tanbun-Ek, T., and Garbinski, P. A. (1995). "Direct measurement of the carrier leakage out of the active region in InGaAsP/InP laser hetrostructures," *IEEE Trans. Electron. Devices*, Vol. 42, No. 2, pp. 215–218.

Bennett, B. R., Soref, R. A., and Delalamo, J. A. (1990). "Carrier-Induced Change in Refractive Index of InP, GaAs, and InGaAsp." *IEEE J. Quantum Electron.*, 26(1), pp. 113–122.

Bhat, R., et al. (1994). "High performance 1.3 mm AlGaInAs/InP strained quantum well lasers grown by organometallic chemical vapor deposition," *J. Crystal Growth*, Vol. 145, pp. 858–865.

Bigan, E., Ougazzaden, A., Huet, F., Carre, M., Carenco, A., and Mircea, A. (1991). *Electron. Lett.*, Vol. 27, No. 18, pp. 1607–1608.

Bissessur, H. (1992). "Effects of hole burning, carrier induced losses and the carrier dependent differential gain on the static characteristics of DFB lasers," *IEEE J. Lightwave Tech.*, Vol. 10, No. 11, pp. 1617–1630.

Blauvelt, H. A. Private communication.

Blauvelt, H. A., Chen, P. C., Bar-Chaim, N., Jiang, P., Lu, C. J., and Ury, I. (1992). "High-power 1310 nm DFB lasers for AM video transmission," *IEEE J. Lightwave Tech.*, Vol. 10, No. 11, pp. 1766–1772.

Blauvelt, H. A., Kwong, N. S., Chen, P. C., and Ury, I. (1993). "Optimum range for DFB laser chirp for fiber-optic AM video transmission," *IEEE J. Lightwave Tech.*, Vol. 11, No. 1, pp. 55–59.

Buus, J. (1985). "Mode selectivity in DFB lasers with cleaved facets," *Electron. Lett.*, Vol. 21, pp. 179–180.

Chen, Y. J. (1998). "Compromise between the performance of AM and QAM signals in hybrid multichannel AM/QAM," Proceedings, *International Conference on Communication Technology*, Beijing, China.

Chinone, N., Saito, K., Shige, N., and Aiki, K. (1997). "High-efficient (GaAl)Al buried-heterostructure lasers for analog communications," Proceedings, *Optical Communication Conference*, Eindhoven, Netherlands.

Chraplyvy, A. R. (1990). "Limitations on Lightwave Communications Imposed by Optical-Fiber Nonlinearities," *J. Lightwave Tech.*, Vol. 8, pp. 1548–1557.

Chraplyvy, A. R. and Tkach, R. W. (1998). "Terabit/Second Transmission Experiments," *J. Quantum Electron.*, Vol. 34, pp. 2103–2108.

Chiddix, J. A., Laor, H., Pangrac, D. M., Williamson, L. D., and Wolfe, R. W. (1990). *IEEE J. Sel. Areas Comm.*, Vol. 8, No. 7, pp. 1229–1239.

Chiddix, J. A., Vaughan, J. Z., and Wolfe, R. W. (1993). *IEEE J. Lightwave Tech.*, Vol. 11, No. 1, pp. 154–166.

Chiddix, J. A., Wolf, H., Boyer, R., Darcie, T. E., Vecchi, M., and Warner, T. (2000). *Info*, Vol. 2, No. 2, pp. 123–129.

Coldren, L. A. and Corzine, S. W. (1995). *Diode lasers and photonic integrated circuits*, J. Wiley and Sons.

Corvini, P. J. and Koch, T. L. (1987). *J. Lightwave Tech.*, 5, p. 1591.

Darcie, T. E. (1990). "Subcarrier multiplexing for lightwave networks and video distribution systems," *IEEE J. Sel. Areas Comm.*, Vol. 8, No. 7, pp. 1240–1248.

Darcie, T. E. (1991). "Subcarrier Multiplexing for Multiple-Access Lightwave Networks," *J. Lightwave Tech.*, Vol. LT5, pp. 1103–1110.

Darcie, T. E. and Bodeep, G. E. (1990). *IEEE Trans. Microwave Theory and Tech.*, Vol. 38, No. 5, pp. 524–533.

Darcie, T. E., Tucker, R. S., and Sullivan, G. J. (1985). "Intermodulation and harmonic distortion in InGaAsP lasers," *Electron. Lett.*, Vol. 21, No. 16, pp. 665–666.

Delorme, A., Alibert, G., Boulet, P., Grosmaire, S., Slempkes, S., and Ougazzaden, A. (1997). "High reliability of high-power and widely tunable 1.55-μm distributed Bragg reflector lasers for WDM applications," *IEEE J. Selected Topics in Quantum Electronics*, Vol. 3, No. 2, pp. 607–614.

Delorme, F., et al. (1998). "Sampled-grating DBR lasers with 101 wavelengths over 44 nm and optimised power variation for WDM applications," *Electron. Lett.*, Vol. 34, No. 3, pp. 279–281.

Devaux, F., Bigan, E., Rose, B., McKee, M., Huet, F., and Carre, M. (1991). *Electron. Lett.*, Vol. 27, No. 21, pp. 1926–1927.

Devaux, F., Sorel, Y., and Kerdiles, J. F. (1993). *J. Lightwave Tech.*, Vol. 11, No. 12, pp. 1937–1940.

Doerr, C. R., et al. (1996). "Chirped waveguide grating router multifrequency laser with absolute wavelength control," *IEEE Photon. Tech. Lett.*, Vol. 8, No. 12, pp. 1606–1608.

Doerr, C. R., Joyner, C. H., and Stulz, L. W. (1999). "40-Wavelength Rapidly Digitally Tunable Laser," *IEEE Photon. Tech. Lett.*, Vol. 11, No. 11.

Flynn, E. J. and Ackerman, D. A. (1997). "Electron leakage and the excess voltage at p-p heterojunctions in InGaAsP/InP lasers," IEDM Technical Digest, *International Electron. Devices 1997*, Washington, DC.

Frigo, N. J., et al. (1994). "A wavelength-division multiplexed passive optical network with cost-shared components," *IEEE Photon. Tech. Lett.*, Vol. 6, No. 11, pp. 1365–1367.

Fuse, M., Kudo, Y., Maeda, K., Ohya, J., Ishino, M., and Fujito, K. (1996). "Development of 128 optical distribution systems of 150 chs AM/QAM hybrid signals," *Electronics and Communications in Japan, Part 1 (Communications)*, Vol. 79, No. 11, pp. 65–77.

Futakuchi, N., Taguchi, T., and Nakano, Y. (1998). "Reduction of second- and third-order harmonic distortion by nonlinear absorption in gain-coupled DFB lasers," presented at *IEEE 16th International Semiconductor Laser Conference*, Conference Digest, Nara, Japan.

Ghafouri-Sharaz, H. (1996). *Fundamentals of Laser Diode Amplifier*, J. Wiley and Sons.

Girardin, F., Duan, G., Gallion, P., Talneau, A., and Ougazzaden, A. (1995). "Experimental investigation of the relative importance of carrier heating and spectral-hole-burning on nonlinear gain in bulk and strained 1.55 μm lasers," *Appl. Phys. Lett.*, Vol. 67, No. 6, pp. 771–773.

Grubb, S. G., DiGiovanni, D. J., Simpson, J. R., Cheung, W. Y., Sanders, S., Welch, D. F., and Rockney, B. (1996). "Ultrahigh power diode-pumped 1.5-μm fiber amplifiers," In: (eds.) San Jose, pp. 30–31.

Guthrie, J., Tan, G. L., Ohkubo, M., Fukushima, T., Ikegami, Y., Ijichi, T., Irikawa, M., Mand, R. S., and Xu, J. M. (1994). "Beam instability in 980 nm high power lasers," *IEEE Photon. Tech. Lett.*, Vol. 6, No. 12, pp. 1409–1411.

Haisch, H., Cebulla, U., Zielinski, E., Bouayad-Amine, J., Klenk, M., Laube, G., Mayer, H.-P., Weinmann, R., and Speier, P. (1993). *IEEE J. Quantum Electron.*, Vol. 29, No. 6, pp. 1782–1791.

Helms, J. (1991). "Intermodulation and harmonic distortions of laser diodes with optical feedback," *IEEE J. Lightwave Tech.*, Vol. 9, No. 11, pp. 1567–1575.

Henry, C. H. (1985). "Performance of distributed feedback lasers designed to favor the energy gap mode," *IEEE J. Quantum Electron.*, Vol. 21, No. 12, pp. 1913–1918.

Hirota, O. and Suematsu, Y. (1979). *IEEE J. Quantum Electron.*, Vol. QE-15, No. 3, pp. 142–149.

Hong, J., Cyr, M., Kim, H., Jatar, S., Rogers, C., Goodchild, D., and Clements, S. (1999). "Cascaded Strongly Gain-Coupled (SGC) DFB Lasers with

15-nm Continuous-Wavelength Tuning," *IEEE Photon. Tech. Lett.*, Vol. 11, No. 10, pp. 1214–1216.

Ishii, H., et al. (1996). Quasicontinuous Wavelength Tuning in Super-Structure-Grating (SSG) DBR Lasers," *IEEE J. Quantum Electron.*, Vol. 32, No. 3, pp. 433–441.

Ito, K., Fujita, S., and Miyake, Y. (1979). "Optical fiber transmission of ITV video signal by analog baseband modulation of laser diodes," Proceedings, *Optical Communication Conference*, Eindhoven, Netherlands.

Jayaraman, V., Chuang, Z. M., and Coldren, L. A. (1993). "Theory, Design, and Performance of Extended Tuning Range Semiconductor Lasers With Sampled Gratings," *IEEE J. Quantum Electron.*, Vol. 29, No. 6, pp. 1824–1834.

Joyce, W. B. and Dixon, R. W. (1975). "Thermal resistance of heterostructure lasers," *J. Appl. Phys.*

Kaminov, I. P. and Koch, T. L., eds. (1997). *Optical Fiber Telecommunications*, Academic Press, San Diego, Chapters 2, 9, 10, 13–15.

Kawamura, Y., Wakita, K., Yoshikuni, Y., Itaya, Y., and Asahi, H. (1987). *IEEE J. Quantum Electron.*, Vol. QE-23, No. 6, pp. 915–918.

Ketelsen, L. J. P., et al. (2000). "Multiwavelength DFB Laser Array with Integrated Spot Size Converters," *IEEE J. Quantum Electron.*," Vol. 36, No. 6, pp. 641–648.

Ketelsen, L. J. P., Hoshino, I., and Ackerman, D. A. (1991). "The role of axially nonuniform carrier density in altering the TE-TE gain margin in InGaAsP-InP DFB lasers," *IEEE J. Quantum Electron.*, Vol. 27, pp. 957–964.

Ketelsen, L. J. P., Johnson, J. E., Ackerman, D. A., Zhang, L., Hybertsen, M. S., Glogovsky, K. G., Focht, M. W., Reynolds, C. L., Lentz, C. W., Koch, T. L., Kamath, K. K., Asous, W. A., Ebert, C. W., Park, M., and Hartman, R. L. (2000). "2.5 Gb/s transmission over 680 km using a fully stabilized 20 channel DBR laser with monolithically integrated semiconductor optical amplifier, photodetector, and electroabsorption modulator." In: (eds.) Baltimore, pp. paper PD14.

Ketelsen, L. J. P., Johnson, J. E., Ackerman, D. A., Zhang, L., Kamath, K. K., Hybertsen, M. S., Glogovsky, K. G., Focht, M. W., Asous, W. A., Reynolds, C. L., Ebert, C. W., Park, M., Lentz, C. W., Hartman, R. L., and Koch, T. L. (2000). "2.5 Gb/s transmission over 680 km using a fully stabilized 20 channel DBR laser with monolithically integrated semiconductor optical amplifier, photodetector and electroabsorption modulator," *Optical Fiber Communication Conference*, 2000.

Kim, I., Alferness, R. C., Koren, U., Buhl, L. L., Miller, B., Young, M. G., Chien, M. D., Koch, T. L., Presby, H. M., Raybon, G., and Burrus, C. A. (1994). "Broadly tunable vertical-coupler filtered tensile-strained InGaAs/InGaAsP multiple quantum well laser," *Appl. Phys. Lett.*, Vol. 64, No. 21.

Kishino, K., Aoki, S., and Suematsu, Y. (1982). "Wavelength variation of 1.6 mm wavelength buried heterostructure GaInAsP/InP lasers due to direct modulation," *IEEE J. Quantum Electron.*, Vol. QE18, pp. 343–351.

Kito, M., Sato, H., Otsuka, N., Takenaka, N., Ishino, M., and Matsui, Y. (1994). "The effect of dynamic spatial hole burning on the second and third order intermodulation distortion in DFB lasers," Proceedings, *14th IEEE International Semiconductor Laser Conference*, Maui.

Koyama, F. and Iga, K. (1988). *J. Lightwave Tech.*, Vol. 6, No. 1, pp. 87–93.

Kuindersma, P. I., Dijksterhuis, W., Willekes, A. H. E., and Ambrosius, H. P. M. (1985). "Leakage currents in 1.3 μm buried heterojunction laser diodes," presented at Gallium Arsenide and Related Compounds 1984, *Eleventh International Symposium*, Biarritz, France.

Kuindersma, P. I., Valster, A., and Baks, W. (1985). "1.3 μm buried heterostructure laser diodes under high electrical stress: leakage currents and aging behavior," *IEEE J. Quantum Electron.*, Vol. QE-21, No. 6, pp. 726–736.

Kuo, C. Y. (1992). "Fundamental second-order nonlinear distortions in analog AM CATV transport systems based on single frequency semiconductor lasers," *IEEE J. Lightwave Tech.*, Vol. 10, No. 2, pp. 235–243.

Kuo, C. Y., Piehler, D., Gall, C., Kleefeld, J., Nilsson, A., and Middleton, L. (1996). "High-performance optically amplified 1550-nm lightwave AM-VSB CATV transport system," *Optical Fiber Communication Conference*, Vol. 2, 1996 Technical Digest Series, San Jose.

Lau, K. Y. and Yariv, A. (1984). "Intermodulation distortion in a directly modulated semiconductor laser," *Appl. Phys. Lett.*, Vol. 45, No. 10, pp. 1034–1036.

Lee, S., Kim, K., Oh, K. R., Kim, J. S., Oh, D. K., Kim, H. M., and Pyun, K. E. (1998). "Comparison of Fe- and p/n/p-doped InP as a current blocking layer for monolithically integrated InGaAsP/InP laser diodes with passive waveguides," *Jpn. J. Appl. Phys.*, Part 2, Vol. 37, No. 2B, pp. L221–224.

Li, G. P., Makino, T., Moore, R., Puetz, N., Leong, K. W., and Lu, H. (1993). "Partly Gain-Coupled 1.55 μm Strained-Layer Multiquantum-Well DFB Lasers," *IEEE J. Quantum Electron.*, Vol. 29, No. 6, pp. 1736–1742.

Linke, R. A. (1985). *IEEE J. Quantum Electron.*, QE21, pp. 593–597.

Majewski, M. L. (1990). "Nonlinear effects associated with the large-signal direct modulation of semiconductor laser diodes," Conference Digest, LEOS Summer Topical on Broadband Analog Optoelectronics – Devices and Systems, Monterey, CA.

Malik, R. J., ed. (1989). *III-V semiconductor materials and devices*, Elsevier Science Pub., New York.

Mason, B., Barton, J., Fish, G. A., Coldren, L. A., and Denbaars, S. P. (2000). "Design of sampled grating DBR lasers with integrated semiconductor optical amplifiers," *IEEE Photon. Tech. Lett.*, Vol. 12, No. 7.

Mason, B., Fish, G. A., DenBaars, S. P., and Coldren, L. A. (1999). "Widely tunable sampled grating DBR laser with integrated electroabsorption modulator," *IEEE Photon. Tech. Lett.*, Vol. 11, No. 6.

Michaelis, T. D. (1979). *IEEE Trans. Cable TV*, Vol. CATV-4, No. 1, pp. 30–42.

Mickelson, A. R., Basavanhally, N. R., and Lee, Y. C., eds. (1997). *Optoelectronic Packaging*, New York: J. Wiley & Sons.

Miller, D. A. B., Chemia, D. S., Damen, T. C., Gossard, A. C., Wiegmann, W., Wood, T. H., and Burrus, C. A. (1984). *Phys. Rev. Lett.*, Vol. 53, No. 22, pp. 2173–2176.

Miller, D. A. B., Chemia, D. S., Damen, T. C., Gossard, A. C., Wiegmann, W., Wood, T. H., and Burrus, C. A. (1985). *Phys. Rev. B.*, Vol. 32, No. 2, pp. 1043–1060.

Moerman, I., Daele, P. V., and Demeester, P. M. (1997). "A review on fabrication technologies for the monolithic integration of tapers with III-V semiconductor devices," *IEEE J. Sel. Topics Quantum Electron.*, Vol. 3, pp. 1308–1320.

Morthier, G. (1994). "Design and optimization of strained-layer-multiquantum-well lasers for high-speed analog communications," *IEEE J. Quantum Electron.*, Vol. 30, No. 7, pp. 1520–1528.

Morthier, G. and Baets, R. (1996). "Analytical formulation of the spatial hole burning induced distortion in DFB lasers and some implications for CATV laser design," Conference Digest, *15th IEEE International Semiconductor Laser Conference*, Haifa, Israel.

Morthier, G., David, K., Vankwikelberge, P., and Baets, R. (1990). "A new DFB-laser diode with reduced spatial hole burning," *IEEE Photon. Tech. Lett.*, Vol. 2, No. 6, pp. 388–390.

Morthier, G. and Vankwikelberg, P. (1997). *Handbook of Distributed Feedback Laser Diodes*, Artec House.

Morton, P., et al. (1992). "25 GHz Bandwidth 1.55 μm GaInAsP p-doped strained mutiquantum-well laser," *Electron. Lett.*, Vol. 28, No. 23.

Morton, P. A., Ormondroyd, R. F., Bowers, J. E., and Demokan, M. S. (1989). "Large-signal harmonic and intermodulation distortions in wide-bandwidth GaInAsP semiconductor lasers," *IEEE J. Quantum Electron.*, Vol. 25, No. 6, pp. 1559–1567.

Morton, P. A., Shtengel, G. E., Tzeng, L. D., Yadvish, R. D., Tanbun-Ek, T., and Logan, R. A. (1997) "38.5 km Error Free Transmission at 10 Gb/s in Standard Fiber Using a Low Chirp Spectrally Filtered, Directly Modulated, 1.55 μm DFB Laser," *Electron. Lett.*, Vol. 33, No. 4, pp. 310–311.

Nakamura, M., Suzuki, N., Uematsu, Y., Ozeki, T., and Takahashi, S. (1984). "Modal noise elimination in multimode fiber analog transmission using DFB lasers," *Conference on Optical Fiber Communication*, Digest of Technical Papers, New Orleans.

Nash, F. (1993). *Estimating Device Reliability: Assessment of Credibility*, Kluwer.

Oberg, M., et al. (1995). "Complete Single Mode Wavelength Coverage Over 40 Nm With a Super Structure Grating DBR Laser," *J. Lightwave Technol.*, Vol. 13, No. 9, pp. 1892–1898.

O'Brien, S., Shealy, J. R., and Wicks, G. W. (1991). "Monolithic integration of an (Al)GaAs laser and an intracavity electroabsorption modulator using selective partial interdiffusion," *Appl. Phys. Lett.*, Vol. 58, pp. 1363–1365.

Ogawa, K. (1982). "Analysis of mode partition noise in laser transmission systems," *IEEE J. Quantum Electron.*, Vol. QE18, pp. 849–855.

Oshiba, S., et al. (1993). *Electron. Lett.*, Vol. 29, p. 1528.

Ougier, C., et al. (1996). "Sampled-grating DBR lasers with 80 addressable wavelengths over 33 nm for 2.5 Gbit/s WDM applications," *Electron. Lett.*, Vol. 32, No. 17, pp. 1592–1593.

Panish, M. B. and Temkin, H. (1993). *Gas source molecular beam epitaxy*, Springer-Verlag, Berlin.

Parker, E. H. C., ed. (1985). *The technology and physics of molecular beam epitaxy*, Plenum Press, New York.

Peale, D. R., Ackerman, D. A., Brener, I., and Hobson, W. S. (1999). "0.84 terahertz self modulation in a dual transverse mode 980 nm laser," Technical Digest, *Conference on Lasers and Electro-Optics*, Washington, DC.

Phillips, M. R. (1997). "Fiber effects in amplified 1550-nm CATV lightwave systems," *LEOS '97, 10th Annual Meeting, IEEE Lasers and Electro-Optics Society*, San Francisco.

Phillips, M. R. and Darcie, T. E. (1991). "Numerical simulation of clipping-induced distortion in analog lightwave systems," *IEEE Photon. Tech. Lett.*, Vol. 3, No. 12, pp. 1153–1155.

Phillips, M. R., Darcie, T. E., and Flynn E. J. (1992). "Experimental measure of dynamic spatial hole-burning in DFB lasers," *IEEE Photon. Tech. Lett.*, Vol. 4, No. 11, pp. 1201–1203.

Phillips, M. R. and Sweeney, K. L. (1997). "Distortion by stimulated Brillouin scattering effect in analog video lightwave systems," Technical Digest, Postconference Edition, *Conference on Optical Fiber Communications*, Dallas.

Plumb, R. G., Armistead, C. J., Collar, A. J., Butler, B. R., and Kirkby, P. A. (1986). "The effect of facets on the stability of DFB lasers," *Proceedings of the SPIE – The International Society for Optical Engineering*, Vol. 587, No. 21–26.

Po, H., Cao, J. D., Laliberte, B. M., Minns, R. A., Robinson, R. F., Rockney, B. H., Tricca, R. R., and Zhang, Y. H. (1993). "High power neodymium-doped single transverse mode fiber laser," *Electron. Lett.*, Vol. 29, pp. 1500–1501.

Pollock, C. R. (1995). *Fundamentals of Optoelectronics*, R. D. Irwin.

Rabinovich, W. S. and Feldman, B. J. (1989). "Spatial hole burning effects in distributed feedback lasers," *IEEE J. Quantum Electron.*, Vol. 25, No. 1, pp. 23–30.

Raddatz, L. (1999). Private communication.

Ramachandran, M. and Jasti, C. S. (1997). "Dispersion compensation techniques for CATV lightwave transmission systems," Technical Digest, Postconference Edition, *Conference on Optical Fiber Communications*.

Rigole, P. J., et al. (1996). "Quasi-continuous tuning range from 1560 to 1520 nm in a GCSR laser, with high power and low tuning currents," *Electron. Lett.*, Vol. 32, No. 25, pp. 2352–2354.

Sahara, R., Morito, K., and Soda, H. (1994). *Electron. Lett.*, Vol. 30, No. 9, pp. 698–699.

Saleh, A. A. M. (1989). "Fundamental limit on number of channels in subcarrier-multiplexed lightwave CATV system," *Electron. Lett.*, Vol. 25, No. 12, pp. 776–777.

Saleh, B. E. A. and Teich, M. C. (1991). *Fundamentals of Photonics*, J. Wiley and Sons, p. 296.

Sato, K. and Asatani, K. (1981). *Trans. Inst. Electron. and Comm. Eng. Jpn., Sect. E*, Vol. E64, No. 10, pp. 646–652.

Schatz, R. (1995). "Dynamics of spatial hole burning effects in DFB lasers," *IEEE J. Quantum Electron.*, Vol. 31, No. 11, pp. 1981–1993.

Schemmann, M. F. C., v. d. Poel, C. J., v. Bakel, B. A. H., Ambrosius, H. P. M., Valster, A., v. d. Heijkant, J. A. M., and Acket, G. A. (1995). "Kink power in weakly index guided semiconductor lasers," *Appl. Phys. Lett.*, Vol. 66, No. 8, pp. 920–922.

Shtengel, G. E., Belenki, G. L., Hybertsen, M. S., and Ackerman, D. A. (1998). "Advances in measurements of physical parameters of semiconductor lasers," *International Journal of High Speed Electronics and Systems*, Vol. 9, No. 4, pp. 901–940.

Shtengel, G., Kazrinov, R., and Eng, L. (1998) International Laser Conference.

Shuke, K., Yoshida, T., Kanazawa, T., and Shibutani, M. (1995). "Development of 1 GHz almighty fiberoptic transmitter using 550 MHz LD for CATV system," *21st European Conference on Optical Communication*, ECOC '95, Brussels.

Soda, N., Furutsu, M., Sato, K., Okazaki, N., Yamazaki, S., Nishimoto, H., and Ishikawa, H. (1990). "High-power and high-speed semi-insulating BH structure monolithic electroabsorption modulator/DFB laser light source," *Electron. Lett.*, Vol. 26, pp. 9–10.

Soda, H., Kotaki, Y., Sudo, H., Ishikawa, H., Yamakoshi, S., and Imai, H. (1987). "Stability in single longitudinal mode operation in GaInAsP/InP phase-adjusted DFB lasers," *IEEE J. Quantum Electron.*, Vol. 23, pp. 804–814.

Straus, J. and Szentesi, O. I. (1975). "Linearized transmitters for optical communications," *IEEE International Symposium on Circuits and Systems Proceedings*, Phoenix.

Stringfellow, G. B. (1989). *Organometallic vapor-phase epitaxy: theory and practice*, Academic Press.

Sundaresan, H. and Henning, I. D. (1991). "Measurements On Competition Between Carrier-Induced and Thermally-Induced Wavelength Tuning in Ridge Waveguide DBR Lasers," *Electron. Lett.*, Vol. 27, No. 21, pp. 1897–1899.

Susaki, W. (1991). "Recent progress in superlinear InGaAsP laser diodes," Technical Digest Series, *Conference on Optical Fiber Communications*, San Diego, CA.

Suzuki, M., Noda, Y., Tanaka, H., Akiba, S., Kushiro, Y., and Isshiki, H. (1987). "Monolithic integration of InGaAsP/InP distributed feedback laser and electroabsorption modulator by vapor phase epitaxy," *J. Lightwave Tech.*, Vol. LT5, pp. 1277–1285.

Suzuki, M., Tanaka, H., Utaka, K., Edagawa, N., and Matusuchima, Y. (1992). *Electron. Lett.*, Vol. 28, No. 11, pp. 1007–1008.

Sze, S. M. (1983). *VLSI technology*, McGraw-Hill.

Takemoto, A., Ohkura, Y., Kimura, T., Yoshida, N., Kakimoto, S., and Susaki, W. (1989). "1.3 μm distributed feedback laser diode with grating accurately controlled by new fabrication technique," *Electron. Lett.*, Vol. 25, pp. 220–221.

Takemoto, A., Watamabe, H., Nakajima, Y., Sakakibara, Y., Kakimoto, S., Yamashita, J., Hatta, T., and Miyake, Y. (1990). "Distributed feedback laser diode and module for CATV systems," *IEEE J. Sel. Areas Commun.*, Vol. 8, No. 7, pp. 1359–1364.

Tamanaka, Y., et al. (1993). "Theoretical Study on Enhanced Differential Gain and Extremely Reduced Linewidth Enhancement Factor in Quantum-Well Lasers," *IEEE JQE*, Vol. 29, No. 6.

Tanbun-Ek, T., Chen, Y. K., Grenko, J. A., Byrne, E. K., Johnson, J. E., Logan, R. A., Tate, A., Sergent, A. M., Wecht, K. W., Sciortino, P. F., and Chu, S. N. G. (1994). "Integrated DFB-DBR laser modulator by selective area metalorganic vapor phase epitaxy growth technique." *J. Crystal Growth*, Vol. 145, pp. 902–906.

Tatsuno, K., Kato, T., Hirataka, T., Yoshida, K., Ishii, T., Kono, T., Miura, T., Sudo, R., and Shimaoka, M. (1997). "High performance and low-cost plastic optical modules for the access network applications," In: (eds.) Dallas, pp. 111–112.

Thomas, G. A., Shraiman, B. I., Glodis, P. F., and Stephen, M. (2000). "Toward the Clarity Limit in Optical Fiber," *Nature*, Vol. 404, pp. 262–264.

Thrush, E. J., Glew, R. W., Greene, P. D., Gibbion, M. A., Armistead, C. J., Smith, A. D., Briggs, A. T. R., Scarrott, K., Czajkowski, I. K., Jones, C. J., and Patel, B. L. (1994). "The Growth of 1550 nm Integrated Laser/Modulator Structures by MOCVD," *Sixth International Conf. of Indium Phosphide and Related Materials*, ME2, p. 72.

Tsang, W. T., ed. (1985). *Lightwave communications technology*, Academic Press.

Vinzio, R., Toffano, Z., and Destrez, A. (1998). "Analytical formulation of distortion and chirp in CATV DFB lasers including spatial hole burning," *IEEE J. Quantum Electron.*, Vol. 34, No. 2, pp. 311–317.

Wakita, K. (1998). *Semiconductor Optical Modulators*, Kluwer Academic Publishers.

Watanabe, H., Aoyagi, T., Takemoto, A., Takiguchi, T., and Omura, E. (1995). "Low distortion laser diodes for optical analog transmission application," Conference Proceedings, *IEEE Lasers and Electro-Optics Society*, 8th Annual Meeting.

Watanabe, H., Takiguchi, T., and Omura, E. (1995). "DFB laser diodes for analog transmission," *Rev. Laser Eng.*, Vol. 23, No. 7, pp. 522–530.

Weber, J. P. (1998). "Optimization of the carrier-induced effective index change in InGaAsP waveguides-application to tunable Bragg filters," *IEEE J. Quantum Electron.*, Vol. 30, No. 8, pp. 1801–1816.

Whiteaway, J. E. A., Thompson, G. H. B., Collar, A. J., and Armistead, C. J. (1989). "The design and assessment of $\lambda/4$ phase-shifted DFRB laser structures," *IEEE J. Quantum Electron.*, Vol. 24, No. 8, pp. 1261–1279.

Willardson, R. K. and Beer, A., eds. (1990). *Indium phosphide: crystal growth and characterization*, Academic Press.

Wilmsen, E. W., Temkin, H., and Coldren, L. A. (1999). *Vertical-cavity surface-emitting lasers*, Cambridge University Press.

Wilt, D. P. (1998). "Reliability of 1.55 μm DFB Lasers for use in Dense WDM Systems," *Proceedings of Indium Phosphide and Related Materials (IPRM)*, p. 825.

Williams, P. J., Robbins, D. J., Robson, F. O., and Whitbread, N. D. (2000). "High Power and Wide Quasi-Continuous Tuning, Surface Ridge SG-DBR Lasers," *Proceedings of 26th European Conference on Optical Communcation*, Ref No. 6.3.3.

Wilson, G. (1999). "Predistortion techniques for linearization of external modulators," Conference Digest, *1999 LEOS Summer Topical Meetings*, San Diego, CA.

Wood, T. H. (1988). *J. Lightwave Tech.*, Vol. 6, No. 6, pp. 743–757.

Wood, T. H., Pastalan, J. Z., Burrus, C. A., Johnson, B. C., Miller, B. I., deMiguel, J. L., Koren, U., and Young, M. G. (1990). *Appl. Phys. Lett.*, Vol. 57, No. 10, pp. 1081–1083.

Wu, S., Yariv, A., Blauvelt, H. A., and Kwong, N. (1991). "Theoretical and experimental investigation of conversion of phase noise to intensity noise by Rayleigh scattering in optical fibers," *Appl. Phys. Lett.*, Vol. 59, No. 10, pp. 1156–1158.

Yamada, H., Okuda, T., Huang, T., Muroya, Y.. and Torikai, T. (1997). "Partially corrugated waveguide laser diodes for use in CATV networks," *Proceedings of SPIE – International Society of Optical Engineers*, San Jose, CA.

Yamanaka, T., Yoshikuni, Y., Yokoyama, K., Lui, W., and Seki, S. (1993). "Theoretical Study on Enhanced Differential Gain and Extremely Reduced Linewidth Enhancement Factor in Quantum-Well Lasers," *IEEE JQE*. Vol. 29, No. 6.

Yariv "Optical Electronics"

Yariv, A., Blauvelt, H. A., Huff, D., and Zarem, H. (1997). "An experimental and theoretical study of the suppression of interferometric noise and distortion in AM optical links by phase dither," *IEEE J. Lightwave Tech.*, Vol. 15, No. 3, pp. 437–443.

Yasaka, H., Yoshikuni,Y., and Watanabe, M. (1991). "Measurement of gain saturation coefficient of a DFB laser for lasing mode control by orthogonal polarization light," *IEEE J. Quantum Electron.*, Vol. 27, No. 10, pp. 2248–2255.

Yonetani, H., Kamite, K., and Ushijima, I. (1993)."DFB laser diode and module for analog application," *Fujitsu Scientific and Tech. J.*, Vol. 29, No. 4, pp. 377–389.

Young, M. G., Koren, U., Miller, B. I., Chien, M., Koch, T. L., Tennant, D. M., Feder, K.., Dreyer, K., and Raybon, G. (1995). "Six wavelength laser array with integrated amplifier and modulator," *Electron. Lett.*, Vol. 31. pp. 1835–1836.

Zhang, L. and Ackerman, D. A. (1995). "Second- and third-order harmonic distortion in DFB lasers," *IEEE J. Quantum Electron.*, Vol. 31, No. 11, pp. 1974–1980.

Zory. P. S. (1993). *Quantum Well Lasers*, Academic Press, Chapter 3.

Chapter 13 | VCSEL for Metro Communications

Connie J. Chang-Hasnain

Department of Electrical Engineering and Computer Science, University of California, Berkeley, California and Bandwidth 9, Inc., Fremont, California

Vertical cavity surface emitting lasers (VCSELs) emitting at 850 nm wavelength regime are now key optical sources in optical communications. Presently, their main commercial applications are in local area networks (LAN) and storage area networks (SAN) using multi-mode optical fibers and ribbon fibers. There are several books published with details on the fundamental design and characteristics of VCSELs [1–4]. In this chapter, I will focus specifically on recent advances of long-wavelength VCSELs, with emission wavelength range from 1.3 to 1.6 micron for single-mode fiber telecommunication applications. Key device parameters that impact transmission performance will also be discussed.

1. Metro Applications

The telecommunications infrastructure and market have grown explosively over the last decade, fueled by the burgeoning demand for communications and information. A series of optical inventions enabled much of this unprecedented expansion, including the erbium-doped fiber amplifier (EDFA), high-power pump lasers, high-speed distributed-feedback (DFB) laser diode transmitters, and dispersion-shifted fiber. These technologies were primarily developed and deployed in the backbone of the telecommunication infrastructure to provide long-haul transmission of data and voice.

While many innovations and developments have been focused on the deployment of long-haul optical systems, there also exists a growing demand for dense WDM (DWDM) systems optimized for metro environments. Traffic patterns in metropolitan areas can be distinguished from long-haul data flow patterns largely by the uneven flow of data traffic, because metro traffic is more data-centric. In addition, metro systems need to accommodate multiple protocols, multiple data rates, and different quality of service (QoS) requirements. It is essential that metro systems are designed for scalability with minimum startup needs and expenses. Interoperability of equipment from different vendors will be an important consideration. Metro systems need to be far more cost-effective, and more compact and consume less operational power. Finally,

OPTICAL FIBER TELECOMMUNICATIONS, VOLUME IVA

ISBN: 0-12-395172-0

reconfigurable systems are highly desirable to facilitate service provisioning and bandwidth-on-demand services. These differences require different hardware solutions. In short, metro DWDM systems need to be smart, flexible, and cost-effective.

1.1 TRANSMITTER PERFORMANCE REQUIREMENTS

There are no established performance requirements for metro WDM transmitters at the present. Depending on the network architecture and system designs, the required performance can vary a great deal. The distance requirement also varies, from a few kilometers in a campuswide network to 120 km or even longer, approaching wide area network distances. Though there are systems built based on long-haul systems and thus using long-haul requirements on metro components, new systems are being developed based on innovations specific to metro applications to reduce costs. New single-mode fibers are being deployed with negative dispersion to compensate for chirp-induced propagation power penalty for diode lasers under direct modulation [5–6]. Compact and cost-effective single-channel and multichannel optical amplifiers are being developed for metro applications. Both of these developments will impact the transmitter performance requirements, more specifically on power and chirp.

One major cause for the high costs in long-haul components is the extensive reliability qualification process, known as Telcordia standards. Since metro traffic is data-centric and LAN-like, we expect different standards to emerge to reduce the extensive testing required on optical components and bring the metro system costs closer to those of LAN systems.

To limit downtime resulting from system failures, carriers use spare components to establish infrastructure redundancy. In the case of laser sources, this requires a redundant line card with fixed-wavelength transmitters for each working wavelength, which represents a large capital investment. Additional costs associated with this redundancy include space and power costs. This expenditure can be greatly reduced by using broadly tunable sources, so that a single spare line card can be substituted for a wide range of wavelengths. Further cost reductions are possible if system reconfigurations are implemented remotely and quickly.

Another cost issue is that many customers employ systems operating with slightly different wavelengths. This limits the flexibility and optimization of network designs. The availability of tunable wavelength sources would eliminate such mismatch issues.

1.2 TUNABLE LASERS FOR RECONFIGURABLE NETWORK

Metro DWDM systems based on fixed wavelength add/drop nodes work well when a limited number of wavelengths and system reconfigurations are required. However, as traffic and customer requirements change, costly and time-consuming system reconfigurations are needed for such systems.

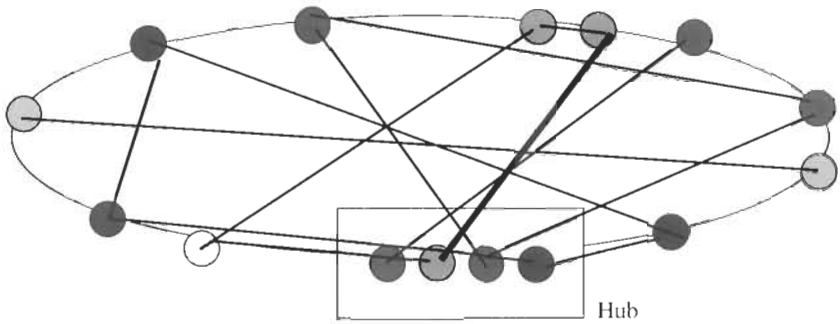


Fig. 1 A physically connected fiber ring can be made into a fully connected mesh network based on wavelengths.

Wavelength-tunable lasers enable simple network reconfiguration, potentially providing wavelengths on demand and dynamic provisioning of data flow. In essence, a physically connected fiber ring can be made into a fully connected mesh network based on wavelengths. Figure 1 shows a schematic for such a design.

Reconfigurability presents one major advantage—greatly reduced time and labor to remove or add subscribers to the network. Using tunable sources and filters, changes in configuration can be accomplished remotely in less than one millisecond. The revenue loss due to network outage caused by operator errors can be largely eliminated. In addition, new services to increase revenue are also likely to be enabled.

2. Introduction

2.1 ADVANTAGES OF VCSEL

A typical VCSEL (Fig. 2) consists of two oppositely-doped distributed Bragg reflectors (DBR) with a cavity layer between. In the center of the cavity layer resides an active region, consisting of multiple quantum wells. Current is injected into the active region via a current guiding structure either provided by an oxide aperture or proton-implanted surroundings [7–13]. As the entire cavity can be grown with one-step epitaxy, these lasers can be manufactured and test on a wafer scale. This presents a significant manufacturing advantage.

The VCSEL cavity is very short, 100–1000 times shorter than that of a typical edge-emitting laser. There is typically only one Fabry–Perot (FP) wavelength within the gain spectrum; hence the FP wavelength (and not the gain peak) determines the lasing wavelength. The optical thickness variation of

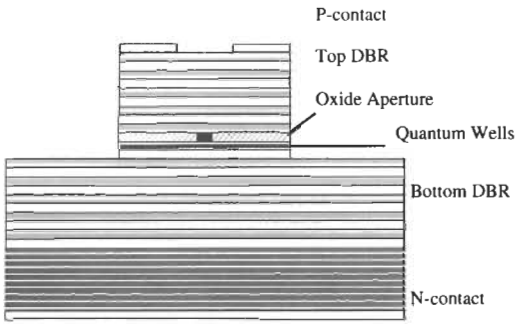


Fig. 2 Schematic of a typical oxide-confined VCSEL, consisting of two oppositely-doped distributed Bragg reflectors (DBR) with a cavity layer in between. In the center of the cavity layer resides an active region, consisting of multiple quantum wells. Current is injected into the active region via a current guiding structure either provided by an oxide aperture. As the entire cavity can be grown with one-step epitaxy, these lasers can be manufactured and tested on a wafer scale.

the layers in a VCSEL changes the lasing wavelength. The position of the layers with thickness variation relative to the center of the cavity is crucial for the resulting wavelength variation; the closer they are to the cavity center, the larger the wavelength change. This property lends to simple designs of wavelength-tunable VCSEL and multiple wavelength VCSEL arrays [14–33].

2.2 CHALLENGES IN LONG-WAVELENGTH VCSEL DESIGN

Progress on long-wavelength VCSELs has been hindered by poor optical and thermal properties of conventional InP-based III-V materials, i.e., InGa(Al)As and InGaAsP quaternary compound lattice matched to an InP substrate. Comparing to the GaAs-based materials, the InP-based materials have lower optical gain, higher Auger recombination, smaller difference in refractive indices, higher free carrier absorption, and much lower thermal conductivities. Table 1 lists the key parameters of these compounds as well as the new ones discussed in sections 3 and 4. These facts translate into major challenges in searching for a promising gain material and DBR designs. In addition, a suitable device structure with a strong current confinement is required. The oxide-confined structure suitable for GaAs-based VCSELs, which led to greatly increased output power and efficiency, is readily extended for many of the InP-based VCSEL designs.

Significant research efforts were carried out for about a decade. There were two major designs using InGa(Al)As strained quantum wells (QW) on an InP substrate. The variation was on the material choice for the mirrors: dielectric or wafer-fused GaAs/AlGaAs DBRs. Using wafer fusion, 1.3 and 1.55 micron VCSELs with steady improvements have been achieved over the past few years.

Table 1 **Key Material Parameters of III-V Compounds**

Wavelength (microns)	*Material*	n_H	*Thermal Conductivity (W/cm K)*	*Material*	n_L	*Thermal Conductivity (W/cm K)*	Δn *(%)*
1.3	GaAs	3.41	0.440	AlAs	2.92	0.900	15.5
	InGaAsP	3.37	0.043	InP	3.21	0.680	4.9
	AlGaInAs	3.35	0.042	$Al_{0.48}In_{0.52}As$	3.24	0.045	3.3
	AlGaAsSb	3.57	0.050	$AlAs_{0.56}Sb_{0.44}$	3.15	0.057	12.5
	AlGaPSb	3.57	0.038	$AlP_{0.4}Sb_{0.6}$	3.10	0.039	14.1
	$In_{0.2}Ga_{0.8}As$	3.47	0.060	$Al_{0.2}In_{0.8}As$	3.04	0.070	13.2
1.5	GaAs	3.37	0.440	AlAs	2.89	0.900	15.3
	InGaAsP	3.45	0.045	InP	3.17	0.680	8.5
	AlGaInAs	3.47	0.045	$Al_{0.48}In_{0.52}As$	3.21	0.045	7.8
	AlGaAsSb	3.60	0.062	$AlAs_{0.56}Sb_{0.44}$	3.10	0.057	14.9
	AlGaPSb	3.55	0.046	$AlP_{0.4}Sb_{0.6}$	3.05	0.039	15.2

Wavelength (microns)	*Material*	*QW Thickness (nm)*	*Strain (%)*	*Jtr*	*dg/dN*	*Auger Coeff B*
0.85	GaAs	8.00	0	110	1300	
0.98	$In_{0.2}Ga_{0.8}As$	8.00	2	50	1200	
1.55	$In_{0.53}Ga_{0.47}As$	7.00	0	81	583	
1.55	$In_{0.53}Ga_{0.47}As$	3.00	1			

We shall not include these developments in this chapter because the technique and theory have been well summarized previously [34].

Recently, breakthrough results were achieved with some very different new approaches [35–46], as well as a design using dielectric mirror [47–48]. The new approaches center on a main theme: the VCSEL structure is grown by a single step of epitaxy. Hence, these devices have promising prospects for higher manufacturability.

The new approaches can be grouped into two major directions: new active materials or new DBRs. The former approach is typically GaAs-based (i.e., on a GaAs substrate) and heavily leveraged on the mature GaAs/AlGaAs DBR and thermal AlO_x technologies. The new active material includes InGaAs quantum dots (QD) and GaInNAs and GaAsSb quantum wells. By and large, this group has focused on extending the active material to longer wavelengths. Currently, 1.3 micron wavelength has been achieved and efforts on 1.55 micron are still at a very early stage. The second approach focuses on engineering of new DBRs. The material system is primarily InP-based, leveraging on heavily documented understandings of InGa(Al)As quantum wells (QW) in the 1.55 micron wavelength range [49]. The DBRs include InGaAsSb

and metamorphic GaAs/AlGaAs. In the following section, I will summarize some representative designs and results.

Key features such as single epitaxy and top-emission appear to be the necessary conditions for the success of a particular long-wavelength VCSEL design. One-step epitaxy greatly increases device uniformity and reduces device or wafer handling and thus testing time. Similarly, top emission (as opposed to bottom emission) enables wafer-scale testing before devices are packaged. It reduces wafer handling, improves coupling efficiency into a single-mode fiber, and eliminates the potential of reliability concern of soldering metal diffusion into the top DBR. Further, to date, only top-emitting VCSEL geometry has proven reliability. Top-emitting structure also facilitates an easy integration with a microelectromechanical (MEM) structure to result in a wavelength-tunable VCSEL [46].

3. 1.3 Micron VCSEL

As mentioned in 2.1, the choice of DBR material is critical for the optimization of VCSEL performance. Using a GaAs substrate, the typical GaAs/AlGaAs DBRs used for commercial 850 nm VCSELs can be applied for 1.3–1.6 micron emission. Another DBR combination that can provide a very high reflectivity is GaAs/(Ga)AlO_x. The lateral current confinement can be provided using an AlO_x aperture or proton implant, both well established for 850 nm VCSELs. Hence, the DBR design and VCSEL processing are all proven. The most challenging task is to extend the wavelength of a new active material. Most of the current results are presently centered on 1.3 micron emission.

3.1 GaInNAs ACTIVE REGION

$Ga_{1-x}In_xN_yAs_{1-y}$ is a compound semiconductor that can be grown to lattice match to a GaAs substrate by adjusting the compositions of N and In. The direct bandgap decreases with increasing N as well as In content. Figure 3 shows theoretical calculation of GaInNAs compounds' bandgap energy versus lattice constant, together with those of other well characterized III-V compounds. For example, typical 1.3 micron emission can be obtained with 1.5–2% N and 35–38% of In. As it is very challenging to incorporate a higher nitrogen content, due to the miscibility gap, it has been difficult to obtain longer wavelength material with high photoluminescence efficiency. Initial results appeared to indicate that 1.2 microns might be the longest wavelength for a good-performance VCSEL using both MOCVD and MBE [51–53].

Recently, almost simultaneously, three groups reported GaInNAs top-emitting VCSEL at 1.3 micron grown by MBE [35–37]. Figure 4 shows one published VCSEL. Lateral contacts were used in this structure for electrical injection. The current is confined into a small aperture using AlO_x aperture.

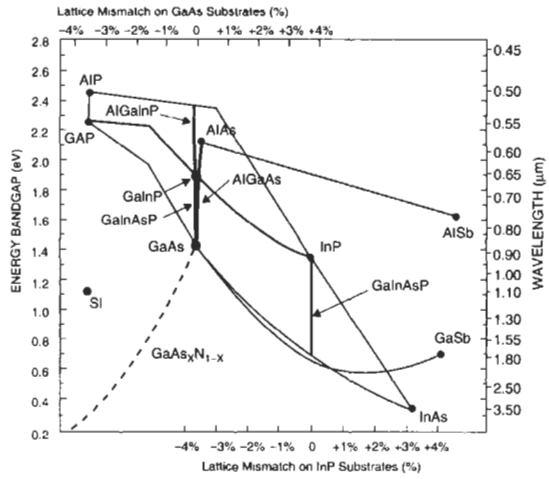


Fig. 3 Theoretical calculation of GaInNAs and GaAsSb bandgap energy versus lattice constant, together with other well characterized III-V compounds [3].

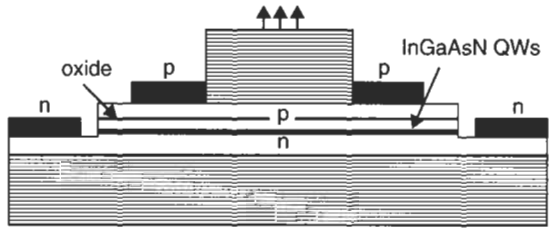


Fig. 4 Schematic of a 1.3 micron VCSEL with GaInNAs active region [36]. Lateral contacts were used in this structure for electrical injection. The current is confined into a small aperture using AlO_x aperture. The DBRs consist of *xx* and *yy* pairs of undoped GaAs/AlAs layers for top and bottom, respectively. The reported output power was ~0.8 mW.

The DBRs consists of undoped GaAs/AlAs layers. The reported single-mode output power was ~0.7 mW under room temperature CW operation. High-speed digital modulation at 10 Gbps were also reported [35,37].

One advantage this material can potentially provide is a very high characteristic temperature T_o. As the laser heatsink temperature is increased, the threshold current can be expressed as exponential term on the heatsink temperature as in Eq. 1.

$$I_{th}(T + \Delta T) = I_{th}(T)e^{(\Delta T/T_o)} \tag{1}$$

A higher T_o indicates a lower sensitivity to temperature. As the GaInNAs QW is expected to have a very high conduction band offset with GaAs barrier, T_o

is expected to be very high. This is indeed observed for an edge-emitting laser at 1.3 micron [49].

Extending wavelength still longer, a group from University of Wurzburg demonstrated edge-emitting lasers emitting at 1.55 micron. Though the results are still preliminary, it is expected that further development in this material will bring interesting future prospects.

3.2 InGaAs QUANTUM DOTS ACTIVE REGION

Quantum confinement has long been known as an efficient method to improve the performance of optoelectronic devices [54–55]. The suggestion of increased gain and differential gain due to the reduced dimensionality in density of states were key motivating factors for many research efforts. However, the compelling reason for introducing quantum well lasers and strained quantum well lasers to the market place was their facilitation to engineer laser wavelength. Similar motivation is now observed for QD lasers.

As demonstrated in InGaAs strained quantum well lasers, with the increase of In, the bandgap of the material moves towards longer wavelength, and the critical thickness of the material that can be grown on a GaAs substrate reduces. Interestingly, the longest wavelength using this approach to obtain a good performance VCSEL is also approximately 1.2 micron. Increasing the In content further, three-dimensional growth was observed, and islands of high In content material were formed among GaAs material. A detailed review of this field can be found in [39]. Very recently, 1.3 micron VCSELs operated at room passage continuous wave was reported [40].

Figure 5 shows a typical VCSEL structure using InGaAs QD as active material. GaAs/AlO_x was used as the DBRs. Lateral contacts and AlO_x aperture were used to provide current injection and confinement. Output power at the level of 0.8 mW under CW room temperature operation was demonstrated with a threshold current and voltage of 1 mA and 3 V. Rapid developments are also expected in this area.

3.3 GaAsSb ACTIVE REGION

Strained GaAsSb QWs have been considered as an alternative active region for 1.3 micron VCSEL grown on a GaAs substrate [57]. Due to the large lattice mismatch, only a very limited number of QWs can be used. In a recent report, a VCSEL emitting at 1.23 microns was reported to operate under room temperature and CW operation using two $GaAs_{0.665}Sb_{0.335}$ QWs as the active region. Typical GaAs/AlGaAs DBRs were used with AlO_x as current confinement aperture. Very low threshold of 0.7 mA was achieved, although the output power is relatively lower at 0.1 mW.

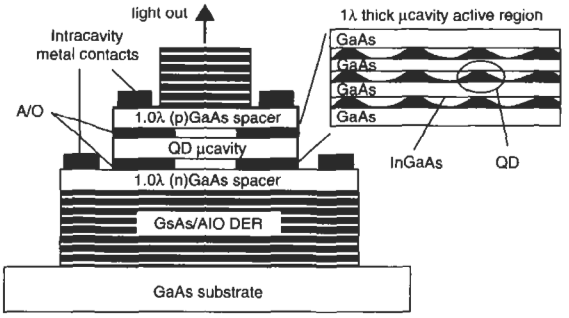


Fig. 5 Schematic of VCSEL structure using InGaAs QD as active material [40]. GaAs/AlO_x was used as the DBRs. Lateral contacts and AlO_x aperture were used to provide current injection and confinement.

4. 1.55 Micron Wavelength Emission

Long-wavelength VCSEL on an InP substrate using conventional InGaAs or InGaAlAs strained QWs as active region possesses one key advantage that the gain region is well understood and the reliability concern is minimum. However, the major obstacles include the small thermal conductivity and index difference in quaternary materials and the strong temperature dependence of optical gain. Current and optical confinements are also a major issue to be resolved to reduce the excessive heating generated at the active junction. Much of the engineering for 1.55 micron VCSEL have focused on these issues. Here we discuss three very different approaches.

4.1 DIELECTRIC MIRROR

Though the dielectric mirror was one of the oldest approaches for making VCSELs, remarkable results were recently published by the group led by M.-C. Amann et al., shown in Fig. 6 [47–48]. In this design, the bottom and top DBRs are InGa(Al)As/InAlAs and dielectric/Au, respectively. Strained InGa(Al)As QWs were grown on top of the bottom n-doped DBR, all lattice matched to an InP substrate.

There are several unique new additions in this design. First, on top of the active region, an n^+-p^+-p tunnel junction was used to provide current injection. A buried heterostructure was regrown to the VCSEL mesa to provide a lateral current confinement. The use of buried tunnel junction (BTJ) provided an efficient current injection mechanism and resulted in a very low threshold voltage and resistance. Secondly, a very small number of pairs of dielectric mirror was used, typically 1.5–2.5 pairs. The dielectric mirror is mounted directly on a Au heatsink, and the resulting net reflectivity is approximately 99.5–99.8%. The few dielectric pairs used here enabled efficient heat removal, which made

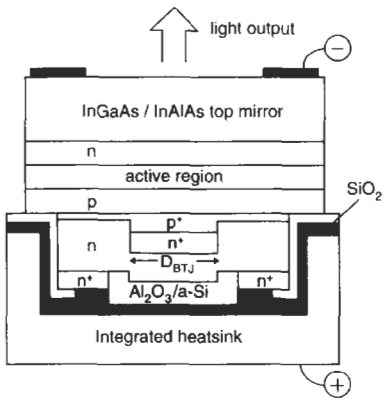


Fig. 6 Schematic of bottom-emitting 1.5 micron VCSEL structure using InGa(Al)As/InAlAs and dielectric/Au as bottom and top DBRs, respectively [47]. Strained InGa(Al)As QWs were grown on top of the bottom n-doped DBR, all lattice matched to an InP substrate.

a strong impact on the laser power and temperature performance. Finally, the substrate is removed to reduce the optical loss and the laser emission is taken from the substrate side.

Bottom-emitting VCSELs with emission wavelength from 1.45–1.85 micron were achieved with this structure. The 1.55 micron wavelength VCSEL with a 5 micron aperture emits a single transverse mode. The highest single-mode power reported is ~1 mW at 20°C under CW operation. A larger 17 micron aperture VCSEL emits above 2 mW under the same condition. Maximum lasing temperatures around 90°C was also obtained.

4.2 AlGaAsSb DBR

As listed in Table 1, AlGaAsSb appears to be a promising DBR material lattice matched to InP. The large bandgap energy difference of AlAsSb and GaAsSb gives rise to a large refractive index difference. For a DBR designed for 1.55 microns, the index difference is approximately 0.5 or 15% between AlGaAsSb (at 1.4 micron bandgap) and AlAsSb. This is nearly the same as the difference between AlAs and GaAs, and much larger comparing to InGaAs/InAlAs at 7.8% and InP/InGaAsP at 8.5%. Also shown in Table 1 are the thermal conductivities. Similar to all quaternary materials, the thermal conductivities are one order of magnitude worse comparing to GaAs and AlAs.

Using AlGaAsSb/AlAsSb as DBRs, a bottom-emitting 1.55 micron VCSEL with one single MBE growth was achieved [41–42]. The device structure is shown in Fig. 7. The active region consists of InGaAsAs strained QWs. The top and bottom DBRs consists of 32.5 and 21.5 pairs of AlGaAsSb/AlAsSb layers. Since the thermal conductivities for the DBRs are very low, this design

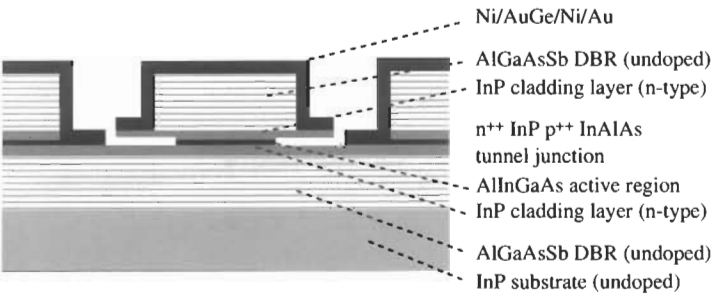


Fig. 7 Schematic of bottom-emitting 1.5 micron VCSEL structure with AlGaAsSb/AlAsSb DBRs lattice matched to an InP substrate [42].

utilized a number of schemes to reduce the heat generated at the active region. First, a tunnel junction was used to reduce the overall p-doping densities, which in turn reduce free carrier absorption. Secondly, intracavity contacts were made for both p- and n-side to further reduce doping-related optical absorption. A wet-etched undercut airgap was created surrounding the active region to provide lateral current and optical confinements.

CW operation at room temperature was reported for these devices. A single-mode VCSEL with 0.9 mW at 25°C was reported. This device operates up to 88°C.

4.3 METAMORPHIC DBR

As mentioned earlier, GaAs/AlGaAs is excellent material combination for DBR mirror because of the large refractive index difference and high thermal conductivities. However, the use of AlGaAs DBRs with InP-based active region by wafer fusion raised concerns on device reliability. This is because in the wafer fusion design, the active region is centered by two wafers fused lattice-mismatched DBRs and current injects through both fusion junctions. A new design using metamorphic DBR [43–46], however, can alleviate such concerns. In the metamorphic design, the active region is grown on top of a lattice-matched DBR and is never removed from it. A metamorphic DBR is deposited some distance away from the active region, serving a function similar to that of a conductive dielectric mirror.

Figure 8 shows the device schematic for a VCSEL consisting of metamorphic DBR [44]. The heterostructure consists of a lattice matched InGaAlAs bottom n-DBR, strained InGaAs QWs, a small number of pairs of p-InGaAlAs DBR, and finally lattice-relaxed, metamorphic GaAs/AlGaAs p-DBR. The use of metamorphic material relaxes the constraints imposed by lattice matching and allows high reflectivity and direct current injection to be achieved simultaneously. The entire structure is grown in a single epitaxy run.

The processing follows that of a conventional 850 nm top-emitting VCSEL on wafer-scale. Selective oxidation is used to provide electrical and optical

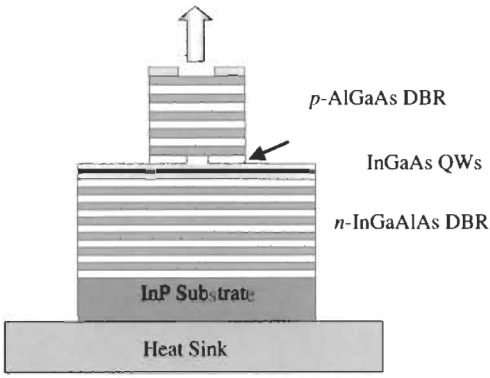


Fig. 8 Schematic for a VCSEL consisting of metamorphic DBR[44]. The heterostructure consists of a lattice-matched InGaAlAs bottom n-DBR, strained InGaAs QWs, a small number of pairs of p-InGaAlAs DBR, and finally lattice-relaxed, metamorphic GaAs/AlGaAs p-DBR. The use of metamorphic material relaxes the constraints imposed by lattice matching and allows high reflectivity and direct current injection to be achieved simultaneously. The entire structure is grown in a single epitaxy run.

confinement. This is presently the only 1.55 micron VCSEL structure that emits from the top of the substrate.

Top-emitting VCSEL with emission wavelength from 1.53–1.62 μm was reported. The devices were grown by a single epitaxy step. Output power up to 0.9 mW in single mode was obtained at 25°C under CW operation. The devices operate up to 75°C. Better than 45 dB side mode suppression ratio was obtained.

The device structure is very similar to that of an 850 nm VCSEL. High-speed data transmission was readily achieved at OC-48 (2.488 Gbps). Transmission performance of a 1590 nm VCSEL at 2.5 Gbps over 150 km of single-mode fiber (Corning LEAF™ fiber) was reported [45].

5. Wavelength Tunable 1.55 Micron VCSEL

As discussed earlier, wavelength-tunable long-wavelength VCSELs that are electrically pumped and directly modulated are highly desirable for WDM applications. Using the metamorphic DBR structure described in 4.3, tunable VCSELs with continuous tuning have been demonstrated in a 1530–1620 nm wavelength regime [46]. The VCSELs exhibit a continuous, repeatable, and hysteresis-free wavelength-tuning characteristic. Further, the VCSELs are directly modulated at 2.5 Gbps and have >45 dB side mode suppression ratio (SMSR) over the entire tuning range. Accurate wavelength locking is achieved in <175 μs by a simple universal locker.

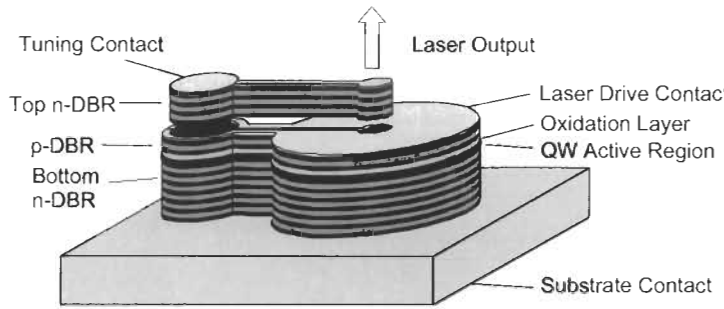


Fig. 9 Schematic of a tunable cantilever VCSEL (c-VCSEL). The device consists of a bottom n-DBR, a cavity layer with an active region, and a top mirror. The top mirror, in turn, consists of three parts (starting from the substrate side): a p-DBR, an airgap, and a top n-DBR, which is freely suspended above the laser cavity and supported via a cantilever structure. Laser drive current is injected through the middle contact via the p-DBR. An oxide aperture is formed on an AlAs layer in the p-DBR section above the cavity layer to provide efficient current guiding and optical index guiding. A top tuning contact is fabricated on the top n-DBR [14].

5.1 TUNABLE CANTILEVER-VCSEL DESIGN

Figure 9 shows a top-emitting VCSEL with an integrated cantilever-supported movable DBR, referred to as cantilever-VCSEL or c-VCSEL. The device consists of a bottom n-DBR, a cavity layer with an active region, and a top mirror. The top mirror, in turn, consists of three parts (starting from the substrate side): a p-DBR, an airgap, and a top n-DBR, which is freely suspended above the laser cavity and supported by the cantilever structure. The heterostructure is similar to that described in 4.3. It was grown in one single step, resulting in a highly accurate wavelength tuning range and predictable tuning characteristics. The active region consists of lattice-matched strained InGaAs quantum wells to ensure high device reliability.

Laser drive current is injected through the middle contact via the p-DBR. An oxide aperture is formed on an Al-containing layer in the p-DBR section above the cavity layer to provide simultaneous current and optical confinements. A tuning contact is fabricated on the top n-DBR. The processing steps include a cantilever formation and release step.

5.2 CONTINUOUS TUNING

Wavelength tuning is accomplished by applying a voltage between the top n-DBR and p-DBR, across the airgap. A reverse bias voltage is used to provide the electrostatic force, which attracts the cantilever downward to the substrate and thus tuning the laser wavelength toward a shorter wavelength. As the movement is elastic, there is no hysteresis in the wavelength-tuning curve. The cantilever returns to its original position once the voltage is removed.

Figure 10 shows a typical c-VCSEL wavelength vs. tuning voltage in the L-band. The continuous, repeatable, and hysteresis-free tuning curve enables the transmission system to lock onto a channel well ahead of its data transmission, known as dark tuning. Dark tuning is important for reconfigurable metro networks when activation and redirection of broadband optical signal must be accomplished without interference to other operating channels.

Figure 11 shows the emission spectra of a c-VCSEL, tuned across 20 channels of C-band at 100 GHz ITU spacing. Single mode with greater than 45 dB side mode suppression ratio (SMSR) was obtained throughout tuning. The

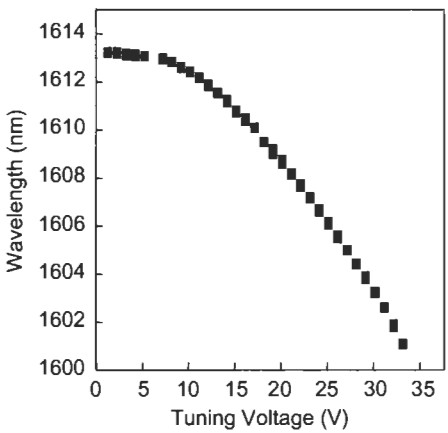


Fig. 10 Emission wavelength as a function of tuning voltage for an L-band tunable c-VCSEL [46]. The curve is continuous, repeatable, and hysteresis-free. This makes it possible to design a WDM system that is upgradable with interleaved channels. Also it will be relatively easier to design a locker capable to support dark tuning.

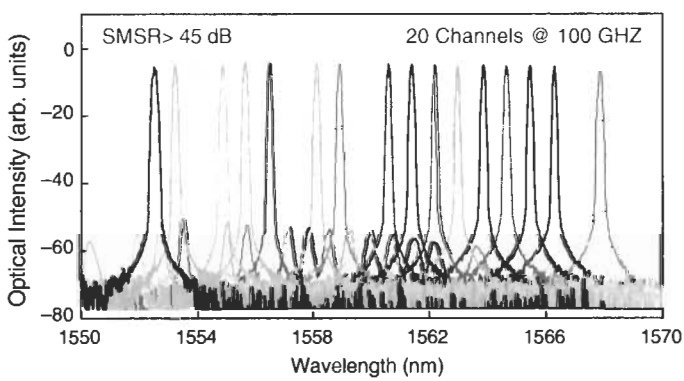


Fig. 11 The emission spectra of a c-VCSEL, tuned across 20 channels in C-band at 100 GHz ITU spacing. Single mode with greater than 45 dB side mode suppression ratio (SMSR) was obtained throughout tuning [46].

tuning range shown here is 16 nm. A larger range over the entire C- or L-band can be achieved with variations of design.

The c-VCSEL is an electrically pumped VCSEL suitable for high-speed direct modulation. Figure 12 shows a typical oscilloscope trace of an eye-diagram for a c-VCSEL under direct modulation at 2.5 Gbps (OC-48). A SONET mask is also shown to illustrate that there are no bit errors. An open and error-free eye diagram is obtained over the full tuning range.

Single-mode fiber transmission performance at 2.5 Gbps (OC-48) of a directly modulated c-VCSEL is exhibited in Fig. 13 [56]. Bit-error-rate (BER)

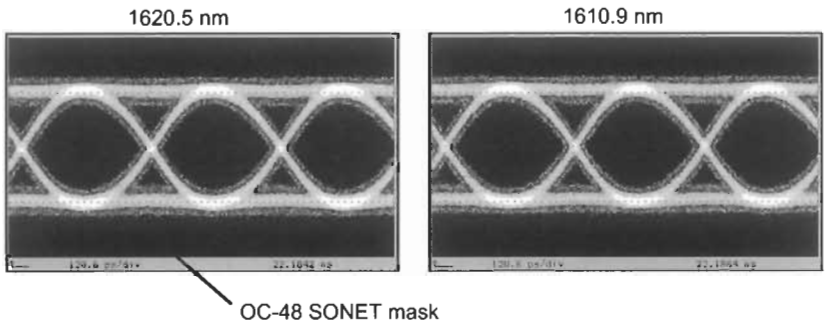


Fig. 12 Typical oscilloscope trace of an eye-diagram for a c-VCSEL under direct modulation at 2.5 Gbps (OC-48). A SONET mask is also shown to illustrate that there are no bit errors. An open and error-free eye diagram is obtained over the full tuning range [46].

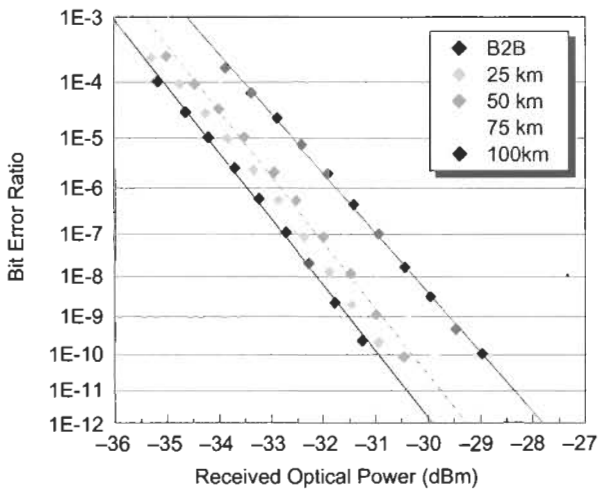


Fig. 13 Typical bit-error-rate (BER) as a function of minimum receiver power for a c-VCSEL under direct modulation at 2.5 Gbps (OC-48) transmitted through various lengths of standard single-mode fiber. No BER floor was observed for any fiber lengths.

data as a function of minimum receiver power for various lengths of standard single-mode fiber (Corning SMF-28^{TM}) is shown. Without the insertion of fiber, the minimum receiver power at 10^{-10} BER is −31.5 dBm, determined by the APD used. As the propagation distance increases, the minimum required receiver power at 10^{-10} BER increases, indicating a power penalty attributed to laser chirp and fiber dispersion (17 ps/nm/km at 1550 nm). However, no BER floor (or saturation) was observed for all fiber lengths tested. Power penalties of 1.4 and 2.3 dB were obtained for transmission through 75 and 100 km of standard single-mode fiber, respectively, with greater than 8.2 dB extinction ratio. We interpolate a power penalty less than 2 dB for 85 km of single-mode fiber, fulfilling one important SONET long-reach requirement.

5.3 TUNING SPEED AND WAVELENGTH LOCKING

The tuning speed is determined by the cantilever dimensions. Since the cantilever structure weighs very little, tuning speed is very fast compared to other MEMS structures and is measured to be 1–10 μs [17].

The continuous wavelength-tuning characteristic enables a simple and cost-effective design of a universal wavelength locker that does not require individual adjustments or calibration for each laser. It also enables rapid locking, resulting in a fast usable tuning speed.

A typical c-VCSEL with a closed-loop feedback locking can be tuned over multiple channels on the 100 GHz ITU grid and locked to within 2.5 GHz accuracy in 175 μs (see Fig. 14). In this case, the lock consists of a reference wavelength filter for identification of absolute wavelength and a comb filter for ITU-grid locking. No special electronics is required.

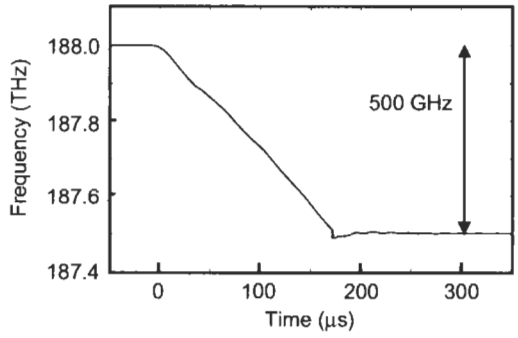


Fig. 14 Optical frequency as a function of time as a c-VCSEL is tuned over five 100 GHz ITU-grid channels with a closed-loop feedback wavelength locker. An accuracy of within 2.5 GHz is achieved in 175 μs. Transient ringing was eliminated using a simple filtering circuitry [46].

A continuous tuning curve further implies that a tunable transmitter that is set at a given ITU-grid spacing can be easily programmed to lock onto a denser spacing without concerns of mode hopping or significant changes in hardware. This enables a WDM system built with such transmitters to upgrade gracefully and cost-effectively in both channel counts and wavelength plans.

5.4 OTHER TUNABLE DIODE LASERS

Using optical pumping (by a high-power 980 nm laser), a VCSEL with a different MEMS structure based on dielectric DBRs was reported in the 1.55 μm wavelength regime [31–32]. In this design, both top and bottom DBRs are dielectric mirrors and the substrate is etched with a via hole to accommodate the deposition of the bottom DBR. A wide tuning range of 43 nm, centered at 1550 nm, was achieved with a pumping threshold of 9 mW and an efficiency of 0.1 mW/mW. Since the pumping source is another laser external to the tunable VCSEL, this approach is not monolithic and is likely to be more expensive to manufacture. However, a higher output power (6 mW at the peak wavelength) was reported resulting from less heat generated at the active region.

There are many rapid developments in the area of widely tuned multisection DBR lasers [58–61]. A multisection DBR laser typically requires three or more electrodes to achieve a wide tuning range of approximately 60 nm with full coverage of wavelengths in the range. The tuning characteristics are discontinuous if only one tuning electrode is used. Hence, exact knowledge of the wavelengths at which the discrete steps occur is critical for precise wavelength control.

6. Transmission Characteristics

Currently, long-wavelength VCSELs are designed primarily for digital modulation applications with nonreturn-to-zero (NRZ) format at 2.5 Gbps and 10 Gbps, for protocols such as synchronous optical network (SONET) for MAN and wide area networks (WAN) and 10G Ethernet for LAN. There are also applications for high-frequency broadband analog modulation applications for cable TV, radar distribution, and antenna remoting. Though the performance requirements are very different for the two modulation schemes, there are overlaps in the fundamental properties required on VCSELs. As much of the long-wavelength VCSEL development is recent and some by start-up companies which do not wish to publish at present, we use published data of 850 nm VCSELs to discuss the fundamental properties, while pointing out where the similarities and differences may be. In this section, we intend to discuss relevant system impairments specific for VCSELs in these applications.

We do not intend to cover detailed analyses nor various modulation formats and system architectures. More can be found in references 62–65.

6.1 SMALL-SIGNAL MODULATION

Small-signal modulation (SSM) response is an important performance indicator for both analog and digital applications. Although large-signal modulation is used in applications, small-signal modulation provides a clean experimental verification for the simple first-order (linear) analysis of laser design. Hence, the primary use of SSM is for guiding laser design and fabrication.

A family of frequency response curves is typically obtained for the laser under various DC biases (CW output powers). In summary, the frequency response can be described by [65]

$$|H(f)|^2 = \frac{1}{1+\left(\frac{f}{f_{RC}}\right)^2} \cdot \frac{1}{\left[1-\left(\frac{f}{f_{RC}}\right)^2\right]^2 + \left(\frac{\gamma}{2\pi f_r}\right)^2 \left(\frac{f}{f_{RC}}\right)^2}$$

$$f_{RC} = \frac{1}{2\pi RC} \tag{2}$$

$$f_r = \frac{1}{2\pi}\left[\eta_i \frac{\Gamma v_g}{qV}\frac{dg}{dN}(I - I_{th})\right]^{1/2}$$

where the f_{RC} and f_r describe the laser parasitic and intrinsic frequency response, respectively. f_{RC} relates to the carrier delivery to the active region, whereas f_r provides the information of electron–photon interaction in the laser cavity (as deciphered by the rate equations). Here, R and C are series resistance and capacitance, respectively; $\eta_i, \Gamma, v_g, q, V, dg/dN, I, I_{th}$ are internal efficiency, optical confinement factor, group velocity, unit charge, active volume, differential gain, DC bias current, and threshold current, respectively.

Based on characterization of f_{RC}, one may improve the laser current guiding structure, reduce the contact pad area, vary the doping densities in the p- and n-doped area surrounding the quantum wells, reduce DBR resistance, and improve on laser packaging.

The laser intrinsic response f_r is determined by a second-order equation with a resonance frequency referred as relaxation oscillation f_r. At f_r, the modulation response shows a peak due to the resonance of electron and photon exchange. At frequencies above f_r, the optical output cannot respond to the modulation and delays at 40 dB/decade. The 3 dB frequency f_{3dB} is often used as a comparison indicator. Providing the damping and parasitics are not too large, f_{3dB} is typically 1.5 times f_r.

With increasing photon density (or DC bias I), f_r increases and the peak response decreases. Using the family curves at various DC biases, important parameters of the active region such as differential gain dg/dN and nonlinear gain can be extracted. In general, high-frequency lasers require a gain medium with a high differential gain and low nonlinear gain coefficient and a laser cavity with a short photon lifetime. The former frequently leads to the use of strained and/or multiple quantum wells and, the latter, a higher output coupling in the output mirror.

There are many publications reporting high frequency SSM of 850 and 980 nm VCSELs. As shown in Table 1, strained quantum wells have a higher differential gain and thus 980 nm VCSELs have typically higher 3 dB responses than 850 nm VCSELs. With careful optimization of various parameters, a 3 dB frequency as high as 21 GHz was reported [68] (Fig. 15). There have not been many publications on frequency response of long wavelength VCSELs. Figure 16 shows a recent result published on 1.3 micron GaInNAs VCSEL [38]. A 3 dB bandwidth of 7 GHz was achieved. However, with careful optimizations of all design parameters, 3 dB response in the high teens of GHz can be expected.

The 3 dB SSM frequency does not convert into large-signal modulation frequency with a simple conversion factor. Depending on the actual time trace of digital signals used to modulate a laser, detailed computer simulations can be done to predict the correlation between the small-signal and large-signal responses. However, the NRZ bit rate is estimated to be 1.5 times 3 dB frequency, largely based on experimental findings.

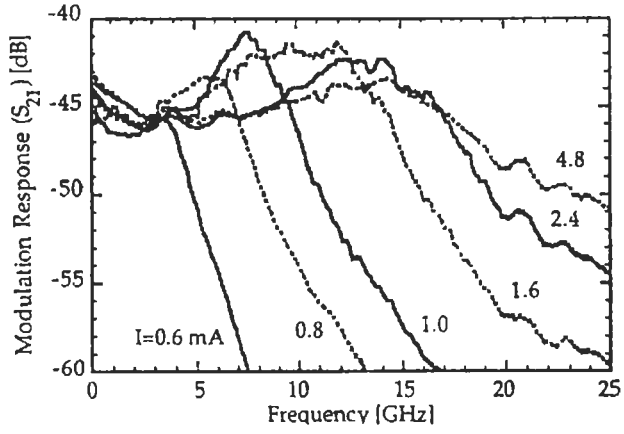


Fig. 15 Small-signal modulation frequency response of a 980 nm oxide-implant VCSEL with optimized parameters. A 3 dB frequency as high as 21 GHz was reported [66].

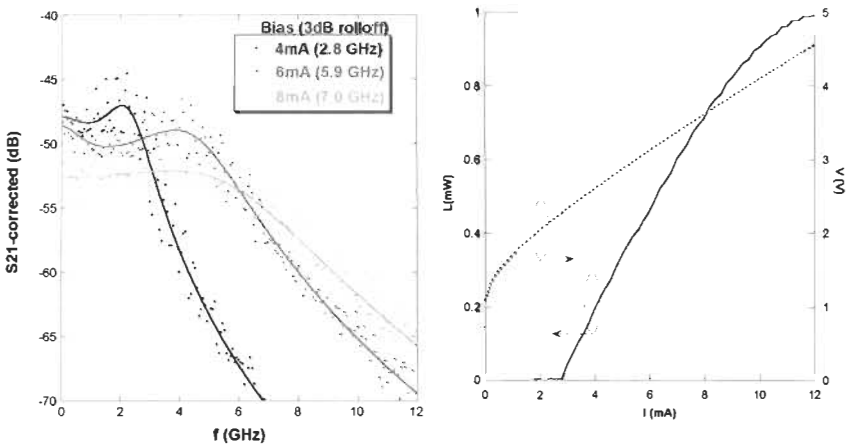


Fig. 16 Small-signal modulation frequency response of a 1.3 micron oxide GaInNAs VCSEL with lateral contacts. A 3 dB frequency of 7 GHz bandwidth was reported [37].

6.2 BIT-ERROR-RATE (BER)

Bit-error-rate (BER) is defined as the probability of incorrect identification of a bit by the decision circuit of the receiver, i.e., the sum of the probabilities of a bit 1 being read as 0 and vice versa. BER is the ultimate test for a digital transmission link. Depending on the specific system designs, a typical BER of 10^{-12} to 10^{-15} is required. A transmission link is typically characterized by BER as a function of the minimum average power required at the receiver (P_{rec}), compared to the case where the laser is tested without the link. If for a given BER, more power is required at the receiver, due to the insertion of components and fibers, the additional power is called power penalty. For a SONET system, the power penalty allowed is 1–2 dB for different link lengths (Table 2).

In addition to power penalty, a BER floor may occur due to various sources of impairment. A BER floor refers to a saturation of BER regardless of increased power P_{rec}, which could be detrimental for system performance. Hence, a thorough BER characterization is very important. A power penalty may be caused by fiber dispersion, laser frequency chirp (in presence of fiber dispersion), jitter, extinction ratio, laser noise, or other optical elements in a given link. For a VCSEL in single-mode fiber metro applications, the first two factors are most significant.

Assuming noise distribution is Gaussian, BER can be expressed as the complementary function of the optical signal-to-noise ratio (Q), as in Eq. 3. Note, the commonly referred signal-to-noise ratio (SNR) is defined as electrical power. The SNR is proportional to the square of the optical signal-to-noise

Table 2 **SONET Specifications per GR-253-CORE Issue 3, September 2000 for Short, Intermediate, Long, and Very Long Reach Transmission Links**

	Link Class[1–4]	*Short (1300 nm)*	*Intermediate*	*Long*	*Very Long*[5]
Transmitter	Loss budget (dB)	7	12	28	33
	Distance (km)	12	42	85	120
	RMS spectral width (nm)	4	N/A	N/A	not specified
	20 dB spectral width (nm)	N/A	1	<1	not specified
	SMSR (dB)	N/A	30	30	not specified
	Max transmitted power (dBm)	−3	0	3	not specified
	Min transmitted power (dBm)	−10	−5	−2	not specified
	Extinction ratio (dB)	8.2	8.2	8.2	10
Optical path	Chromatic dispersion max (ps/nm)	12	not specified	1600	2400
	Attenuation (max) (dB)	7	12	24	33
Receiver	Max Rx power (dBm)	−3	0	−9	−9
	Min Rx power (dBm)	−18	−18	−28	−25
	Power penalty max (dB)	1	1	2	2

[1] All measurements at BER of 10-10, SONET traffic.
[2] Compliance with SONET mask test is also required.
[3] Short reach is only defined in SONET for 1300 nm at OC-48.
[4] SONET spec is intended for single-channel transmission.
[5] VLR only is specified for OC192 transmission, externally modulated.

ratio (OSNR). Equation 5 is a good approximation for a convenient calculation of BER, given Q. For a BER $= 10^{-9}$, $Q = 6$ and BER $= 10^{-12}$, $Q = 7$.

$$BER = \frac{1}{2} erfc\left(\frac{Q}{\sqrt{2}}\right)$$
$$Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0} \tag{3}$$

$$SNR = \frac{P_1 - P_0}{\sigma_1^2 + \sigma_0^2} \approx 2Q^2 \tag{4}$$

$$BER \cong \frac{e^{(-Q^2/2)}}{Q\sqrt{2\pi}} \tag{5}$$

For a given Q, the minimum required power P_{rec} at the receiver can be expressed as

$$P_{rec} = \frac{Q}{R_{\text{det}}}\left(q\frac{B}{2}F_A Q + \frac{\sigma_T}{M}\right) \tag{6}$$

where $R_{\text{det}}, B, F_A, \sigma_T, M$ are detector responsivity, bit rate, amplifier noise figure (APD), thermal noise, and gain (APD). For a p-I-n detector where $M = 1$ and thermal noise dominates, P_{rec} linearly depends on Q. Typically, in a BER vs. P_{rec} plot, Log(Q) is plotted against log (P_{rec}), with the y axis labeled as BER.

The power penalties due to chirp and intensity noise are discussed in the next section. The power penalty δ_{ex} due to extinction ratio (r_{ex}) is expressed by Eq. 7.

$$\begin{aligned} r_{ex} &= P_0/P_1 \\ Q &= \left(\frac{1-r_{ex}}{1+r_{ex}}\right)\frac{2R_{\text{det}}P_{rec}}{\sigma_1+\sigma_0} \\ \delta_{ex} &= 10\log\left[\frac{P_{rec}(r_{ex})}{P_{rec}(r_{ex}=0)}\right] = 10\log\left[\frac{1+r_{ex}}{1-r_{ex}}\right] \end{aligned} \tag{7}$$

where P_0, P_1, and P_{rec} are optical power for bit 0, bit 1, and the average power, respectively.

Note the extinction ratio is different from SNR or Q. Depending on the relaxation frequency of the laser, it is sometimes necessary to bias the laser at high enough powers to increase its speed. Clearly, if the optical power for 0's is substantial (for reduction of rise and fall times), the extinction ratio is reduced. An extinction ratio of 6.5 and 8.2 dB corresponds to a power penalty of 2 and 1.3 dB, respectively.

A large timing jitter can cause power penalty or BER floor. Jitters can be caused by laser biasing at a level that results in a laser being turned off for 0 bit streams. Mode partition noise can also lead to a large timing jitter. (This is when the laser supports more than one mode and the link contains a mode selective element.) Mode partition noise does not cause any concern for a single mode VCSEL with a side mode suppression ratio (SMSR) >25 dB.

6.3 RELATIVE INTENSITY NOISE (RIN)

The output of a diode laser exhibits fluctuations in intensity, phase, and frequency, even when biased by a fixed current source with negligible current fluctuations. The two major noise mechanisms are spontaneous emission and shot noise. Noise in semiconductor lasers is dominated by spontaneous emission. The intensity fluctuations limit SNR, whereas phase fluctuations lead to a broadened linewidth for a laser operated under CW operation. Both affect the performance of a lightwave system.

Relative intensity noise (RIN) of a laser is defined as the spectra density of the time-averaged power fluctuation normalized by the CW average power, with unit of dB/Hz:

$$RIN = \frac{S_P(\omega)}{P^2}$$
$$S_P(\omega) = \int_{-\infty}^{\infty} \langle \delta P(t+\tau)\delta P(t)\rangle e^{-i\omega\tau} d\tau = \langle |\delta P(\omega)|^2 \rangle \tag{8}$$

Note the definition used here is based on electrical power detected by a photodetector, which is proportional to the square of the incident optical power. Hence, in Eq. 8, RIN is expressed as the square of the optical power. The total noise can be calculated by integrating RIN over the bandwidth of interest (typically DC to the carrier frequency or the bit rate). If the spectrum is fairly flat, the relative noise is simply the spectral density times the bandwidth.

RIN is an important characteristic to examine because it provides a simple measurement for a laser's intrinsic speed. The relaxation oscillation frequency f_r can be measured under CW operation, free of possible parasitic effects. RIN is dominated by the P^{-3} term at low power or low frequency, and by the P^{-1} term at higher power or frequency above f_r. Hence, the higher the average power and the higher the relaxation frequency, the lower the relative intensity noise. Figure 17 shows typical RIN spectra for a single-mode 850 nm VCSEL. The state-of-the-art VCSEL has RIN of approximately −140 to −145 dB/Hz

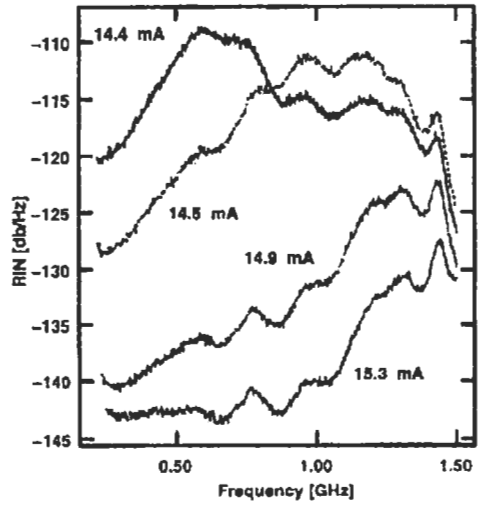


Fig. 17 RIN spectra for a single-mode 850 nm VCSEL. The state-of-the-art VCSEL has RIN of approximately −140 to −145 dB/Hz [67].

[70]. Though this number is higher than the best reported for DFB lasers (−150 to −160 dB/Hz), it is more than sufficient for most applications.

The RIN requirement for a digital system is very lenient. For a given bit error rate, the power penalty due to the laser's relative intensity noise, δ_{ex}, can be expressed by

$$r_{RIN}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_P(\omega)\, d\omega$$
$$\delta_{ex} = 10\log\left[\frac{P_{rec}(r_{RIN})}{P_{rec}(r_{RIN}=0)}\right] = 10\log\left(1 + r_{RIN}^2 Q^2/2\right) \tag{9}$$

For a 10 Gbps system with 10^{-12} BER (corresponding to $Q = 7$), RIN values of −117 and −127 dB/Hz correspond to power penalties of 1 and 0.1 dB, respectively. The typical RIN value for an 850 nm VCSEL is thus more than adequate for digital transmission. Although there are limited reports on RIN of long-wavelength VCSELs, particularly the recent structures mentioned in this chapter, we do not anticipate much deviation from that of 850 nm VCSELs. This makes the RIN value of long-wavelength VCSELs well within the acceptable range.

For an analog system with optical power $P(t)$, where $P(t) = P_0 + mP_1 \sin \omega t$, and m is the modulation depth, the signal to noise ratio (SNR) is determined by RIN:

$$SNR = \frac{\langle i_s^2\rangle}{\langle i_n^2\rangle} = \frac{\langle (P_1 \sin \omega t)^2\rangle}{\langle \delta P(t)^2\rangle} = \frac{m^2 P_0^2}{2\langle \delta P(t)^2\rangle} = \frac{m^2 r_{RIN}^2}{2} \tag{10}$$

For an 8-bit system, the SNR needs to be better than 48 dB. Thus, for $m = 1$, total noise r_{RIN}^2 needs to be less than −51 dB. At 1 GHz, the laser RIN spectral density needs to be less than −141 dB/Hz; and for 10 GHz, this number needs to be better than −151 dB/Hz. Thus, for analog transmission, it is essential to develop VCSELs with both high power and relaxation frequency.

Another important figure of merit for an analog system is the spur-free dynamic range (SFDR). SFDR is the input power range over which the output power at the carrier frequency is above the noise floor while the third-order distortion products remain below noise. The lower the RIN, the larger the SFDR. For a 3 dB increase of total noise, e.g., with 3 dB/Hz increase in RIN or double of the bandwidth, the SFDR is reduced by 2 dB. State-of-the-art 850 nm VCSELs were demonstrated to have a SFDR of 100–108 $dBHz^{2/3}$ over 1 km of multimode fiber [68]. This number is at present still 10–15 dB less than state-of-the-art DFB lasers.

6.4 LINEWIDTH

The fundamental linewidth of a laser is governed by the random nature of spontaneous emissions related to the coherent nature of stimulated emissions, as first derived by A. Schawlow and C. Townes, known as the Schawlow-Townes limit ($\Delta\nu_0$). For a semiconductor laser, the linewidth $\Delta\nu$ is broadened resulting from phase fluctuations caused by spontaneous emissions and a strong coupling of real and imaginary parts of the refractive index, expressed in Eq. 11 [62–63].

$$\Delta\nu = \Delta\nu_0 \left(1 + \alpha^2\right) \tag{11}$$

where $\Delta\nu_0$ is the Schawlow-Townes linewidth

$$\Delta\nu_0 = \frac{1}{4\pi N_P} \frac{\Gamma\beta_{sp}\eta_{rad}N}{\tau_e}$$

$N_P, \Gamma, \beta_{sp}, \eta_{rad}, N, \tau_e$ are the number of photons in the cavity, optical confinement factor, spontaneous emission factor, internal radiation efficiency (the radiation recombination portion of all carrier recombination), carrier numbers and carrier lifetime, respectively. Physically, the Schawlow-Townes linewidth has two components. It is proportional to the amount of spontaneous emission (N/τ_e) coupled to the given mode ($\Gamma\beta_{sp}$) and inversely proportional to optical power (photon numbers). The carrier density N is exactly N_{th} if gain is clamped at the threshold (i.e., for a system with very little spontaneous emission). For practical purposes in diode lasers, even though N is not clamped, it is not far off from N_{th}. Hence, $\Delta\nu_0$ is narrower for low-threshold, high-power lasers.

The linewidth enhancement factor α is defined as

$$\alpha = \frac{4\pi}{\lambda} \frac{dn_r/dN}{dg/dN}$$

where dn_r/dN corresponds to the change in the real part of the refractive index as a function of change in carrier numbers, and dg/dN is the differential gain. α is thus strongly wavelength-dependent. For a given material, it is found more effective to increase the differential gain in order to reduce α. Hence, α can be made smaller by using multiple strained quantum wells, and by making the VCSEL emit on the blue side of the gain spectrum.

A narrow laser linewidth is very important for a transmission link where a Mach-Zender type of external modulator is used, as a broad linewidth translates into a low extinction ratio and signal distortion. However, for direct-modulated lasers, a narrow linewidth per se is not required, albeit a small α is very important for lowering the wavelength chirp.

Linewidth of a diode laser can be measured by delayed self-heterodyne, homodyne, or scanning Fabry-Perot interferometry [69]. The linewidth of a single-mode 850 nm VCSEL was reported to be approximately 20 MHz at an output power level of 0.8 mW. Figure 18 illustrates the experimentally

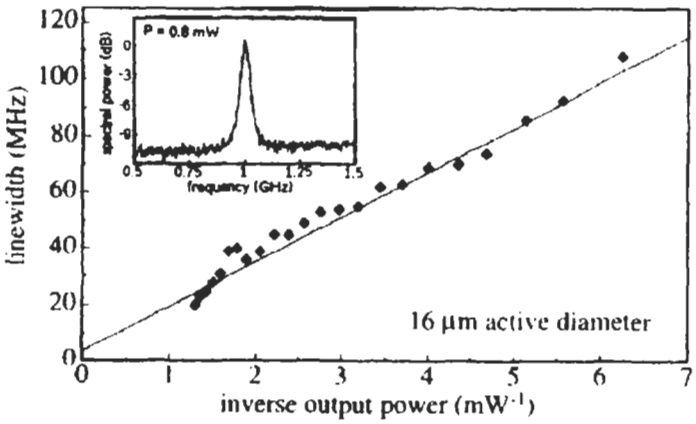


Fig. 18 Plot of measured linewidth versus inverse input power for a 16 micron diameter proton-implanted VCSEL [70]. The inset shows spectral power plotted against frequency.

measured VCSEL linewidth as a function of inverse output power. Linewidth measurements allows simple extraction of the value of α. This is usually very useful for determining the amount of wavelength chirp, whose direct measurement is more complex and expensive (see next section). There has been a limited number of reports on measurements of the linewidth enhancement factor. A value of 3.7 was reported for 850 nm VCSELs [70].

6.5 WAVELENGTH CHIRP

For a digital transmission link, the wavelength chirp (also known as frequency chirp) in a semiconductor laser is one of the most significant sources limiting the transmission distance. Wavelength chirp is defined as the change of the emission wavelength with time of a diode laser driven by a time-varying current source. It arises because the change in carrier density in the laser cavity results in a change of refractive index.

Under digital modulation, two types of chirp have been observed: transient and adiabatic chirp. The former is caused by the fast rise and fall time required for digital modulation and the associated relaxation oscillation. The latter can be caused by nonlinear gain and transient heating, and can have more dependence on signal patterns (e.g., a long strings of 1s vs. mostly 0s with a single 1). For most cases, the transient chirp dominates in magnitude:

$$\Delta\nu_c(t) = -\frac{\alpha}{4\pi}\left(\frac{1}{P(t)}\frac{dP(t)}{dt}\right)$$
$$\Delta\lambda_c(t) = \frac{\lambda^2}{c}\frac{\alpha}{4\pi}\left(\frac{1}{P(t)}\frac{dP(t)}{dt}\right) \tag{12}$$

The amount of chirp, thus, depends on the linewidth enhancement factor, the optical power (photon numbers) and the modulation wave form. The wavelength blueshifts at the beginning of the pulse and redshifts at the tail.

As the different wavelengths generated within a bit period propagate through a single mode fiber with a certain dispersion coefficient (D measured in ps/km/nm), the different wavelengths will propagate at different speeds, causing pulse broadening and distortion, which manifest in power penalty.

In general, the power penalty δ_c depends on the pulse broadening ratio, the ratio of pulse spread due to dispersion and the bit period. This pulse spread is $LD\Delta\lambda_c$, where L is the fiber length, and the bit period is simply the inverse of the bit rate $1/B$. The pulse broadening ratio is thus $BLD\Delta\lambda_c$. The power penalty δ_c also depends on the ratio of the transient chirp duration and the bit period, BT_c. T_c, in turn, is determined by the relaxation oscillation frequency of the laser and its bias condition. Though there have been significant efforts to analyze the power penalty, it is difficult to establish a general model, as there are many interdependent experimental parameters.

Figure 19 shows a calculated power penalty using a rather simplified model and treating BT_c and $BLD\Delta\lambda_c$ as two independent parameters. However, it is very useful as a rough estimate. A 2 dB power penalty is expected for $BT_c = 0.1$ and $BLD\Delta\lambda_c = 0.35$. This corresponds to a 2.5 Gbps modulated laser with a transient chirp of 40 ps and 0.1 nm propagating over 80 km of single-mode fiber with $D = 17$ ps/km/nm (e.g., Corning SMF-28 at 1.55 micron wavelength).

Chirp can be measured by examining the CW spectra or by time-resolved spectra to reveal the transient time constant. A maximum chirp value of 0.1 nm was observed for 1.55 micron VCSELs, similar to or slightly broader than a

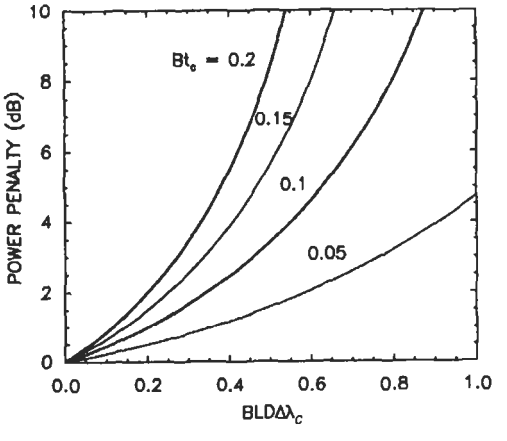


Fig. 19 Chirp-induced power penalty as a function of $BLD\Delta\lambda_c$ for several values of the parameter BT_c. $\Delta\lambda_c$ is the wavelength shift occurring because of wavelength chirp and T_c is the duration of such a wavelength shift [62].

typical DFB laser used in DWDM applications. Further design to reduce the linewidth enhancement factor is the key to lower the chirp-related system impairment.

7. Conclusion

The history of research of long-wavelength VCSELs started when Dr. K. Iga started his work on surface emitting lasers in 1979. However, it was not until recent, breakthroughs were accomplished. The advances in both 1.3 and 1.55 micron VCSELs are rapid and exciting. It is anticipated that the low-cost manufacturing and single-wavelength emission advantages will be a strong force to drive these lasers to the marketplace, particularly for metro area networks and LAN applications.

The monolithic integration of MEMS and VCSEL has successfully combined the best of both devices and led to superior performance in tunable lasers. The tunable cantilever VCSELs are widely tunable, can be directly modulated at high data rates, have a simple monotonic tuning curve for easy wavelength locking, tune reasonably quickly, and emit a reasonable amount of power. They can be batch processed and tested, essential characteristics of mass manufacturability. Tunable lasers will dramatically reduce system and operating costs and increase the degree of connectivity and reconfigurability. Tunable VCSELs are expected to play an important role in these exciting new applications.

References

1. C.J. Chang-Hasnain, Advances of VCSELs, Optical Society of America, Trends in Optics and Photonics Series, 1997.
2. J. Cheng and N.K. Dutta, Vertical Cavity Surface Emitting Lasers: Technology and Applications, Gordon and Breach Science Publishers, 2000.
3. C. Wilmsem, H. Temkin, and Coldren, L.A. Vertical-Cavity Surface-Emitting Lasers, Cambridge University Press, 1999.
4. T.E. Sale, Vertical Cavity Surface Emitting Lasers, Research Studies Press Ltd., 1995.
5. C.-C. Wang, I. Roudas , I. Tomkos, M. Sharma, and R.S. Vodhanel, "Negative dispersion fibers for uncompensated metropolitan networks," Proc. of ECOC, Vol. 1, pp. 97–98, 2000.
6. L. Chrostowski, C.H. Chang, R.J. Stone, and C. Chang-Hasnain, "Study of long-wavelength directly modulated VCSEL transmission using SOA amplifiers," European Conference on Optical Communications, Amsterdam, October 2001.
7. M. Orenstein, A. Von Lehmen, C.J. Chang-Hasnain, N.G. Stoffel, L.T. Florez, J.P. Harbison, J. Wullert, and A. Scherer, "Matrix Addressable Vertical Cavity

Surface Emitting Laser Array," Electron. Lett., Vol. 27, No. 5, pp. 437–438, February 1991.

8. G. Hasnain, K. Tai, L. Yang, Y.H. Wang, R.J. Fischer, J.D. Wynn, B. Weir, N.K. Dutta, and A.Y. Cho, "Performance of gain-guided surface emitting lasers with semiconductor distributed Bragg reflectors," IEEE Journal of Quantum Electronics, Vol. 27, No. 6, pp. 1377–1385, June 1991.
9. K.D. Choquette, K.M. Geib, C.I.H. Ashby, R.D. Twesten, O. Blum, H.Q. Hou, D.M. Follstaedt, B.E. Hammons, D. Mathes, and R. Hull, "Advances in selective wet oxidation of AlGaAs alloys," IEEE Journal of Selected Topics in Quantum Electronics, Vol. 3, pp. 916–926, 1997.
10. D.L. Huffaker, L.A. Graham, H. Deng, and D.G. Deppe, "Sub-40 mu A continuous-wave lasing in an oxidized vertical-cavity surface-emitting laser with dielectric mirrors," IEEE Photonics Technology Letters, Vol. 8, pp. 974–976, 1996.
11. A.E. Bond, P.D. Dapkus, and J.D. O'Brien, "Aperture dependent loss analysis in vertical-cavity surface-emitting lasers," IEEE Photonics Technology Letters, Vol. 11, No. 4, pp. 397–399, April 1999.
12. C. Lei, H. Deng, J.J. Dudley, S.F. Lim, B. Liang, M. Tashima, and R.W. Herrick, "Manufacturing of oxide VCSEL at Hewlett Packard." 1999 Digest of the LEOS Summer Topical Meetings: Nanostructures and Quantum Dots/WDM Components/VCSELs and Microcavities/RF Photonics for CATV and HFC Systems (Cat. No. 99TH8455), San Diego, 26–30 July 1999.
13. Y.A. Wu, G.S. Li, W. Yuen, C.J. Chang-Hasnain, and C. Caneau, "CW Single Mode, Passive Antiguide Vertical Cavity Surface Emitting Laser," IEEE Journal of Selected Topics in Quantum Electronics on Semiconductor Lasers, Vol. 3, No. 2, pp. 429–434, April 1997.
14. C.J. Chang-Hasnain, "Tunable VCSEL," IEEE Journal of Selected Topics in Quantum Electronics, Vol. 6, pp. 978–987, 2000.
15. C.J. Chang-Hasnain, J.P. Harbison, C.E. Zah, M.W. Maeda, L.T. Florez, N.G. Stoffel, and T.P. Lee, "Multiple Wavelength Tunable Surface Emitting Laser Arrays," IEEE J. Quantum Electron, Vol. 27, No. 6, pp. 1368–1376, June 1991.
16. M.S. Wu, E.C. Vail, G.S. Li, W. Yuen, and C.J. Chang-Hasnain, "Tunable micromachined vertical cavity surface emitting laser," Electronics Letters, Vol. 31, pp. 1671–1672, September 14, 1995.
17. E.C. Vail, G.S. Li, W.P. Yuen, and C.J. Chang-Hasnain, "High Performance and Novel Effects of Micromechanical Tunable Vertical Cavity Lasers," IEEE Journal of Selected Topics in Quantum Electronics on Semiconductor Lasers, Vol. 3, No. 2, pp. 691–697, April 1997.
18. M.Y. Li, W. Yuen, G.S. Li, and C.J. Chang-Hasnain, "Top-Emitting Micromechanical VCSEL with a 31.6 nm Tuning Range," IEEE Photonics Technology Letters, Vol. 10, No. 1, pp. 18–20, January 1998.
19. W. Yuen, G.S. Li, and C.J. Chang-Hasnain, "Multiple-wavelength vertical-cavity surface-emitting laser arrays," IEEE J. Select. Topics Quantum Electron., Vol. 3, pp. 422–428, 1997.

20. H. Saito, I. Ogura, and Y. Sugimoto, "Uniform CW operation of multiple-wavelength vertical-cavity surface-emitting lasers fabricated by mask molecular beam epitaxy," IEEE Photon. Technol. Lett., Vol. 8, No. 9, pp. 1118–1120, 1996.
21. F. Koyama, T. Mukaihara, Y. Hayashi, N. Ohnoki, N. Hatori, and K. Iga, "Two-dimensional multiwavelength surface emitting laser arrays fabricated by nonplanar MOCVD," Electron. Lett., Vol. 30, pp. 1947–1948, 1994.
22. G.G. Ortiz, S.Q. Luong, S.Z. Sun, J. Cheng, H.Q. Hou, G.A. Vawter, and B.E. Hammons, "Monolithic, multiple-wavelength vertical-cavity surface-emitting laser arrays by surface-controlled MOCVD growth rate enhancement and reduction," IEEE Photon. Technol. Lett., Vol. 9, pp. 1069–1071, 1997.
23. Y. Zhou, S. Luong, C.P. Hains, and J. Cheng, "Oxide-confined monolithic, multiple-wavelength vertical-cavity surface-emitting laser arrays with a 40-nm wavelength span," IEEE Photon. Technol. Lett., Vol. 10, pp. 1527–1529, 1998.
24. T. Wipiejewski, M.G. Peters, and L.A. Coldren, "Vertical cavity surface emitting laser diodes with post-growth wavelength adjustment," IEEE Photon. Technol. Lett., Vol. 7, pp. 727–729, 1995.
25. C.J. Chang-Hasnain, J.P. Harbison, C.E. Zah, L.T. Florez, and N.C. Andreadakis, "Continuous wavelength tuning of two-electrode vertical cavity surface emitting lasers," Electron. Lett., Vol. 27, No. 11, pp. 1002–1003, 1991.
26. L. Fan, M.C. Wu, H.C. Lee, and P. Grodzinski, "10.1 nm range continuous wavelength-tunable vertical-cavity surface-emitting lasers," Electron. Lett., Vol. 30, No. 17, pp. 1409–1410, 1994.
27. N. Yokouchi, T. Miyamoto, T. Uchida, Y. Inaba, F. Koyama, and K. Iga, "4 Angstrom continuous tuning of GaInAsP/InP vertical-cavity surface-emitting laser using an external cavity mirror," IEEE Photonics Technol. Lett., Vol. 4, No. 7, pp. 701–703, 1992.
28. C. Gmachi, A. Kock, M. Rosenberger, E. Gormick, M. Micovic, and J.F. Walker, "Frequency tuning of a double-heterojunction AlGaAs/GaAs vertical-cavity surface-emitting laser by a serial integrated in-cavity modulator diode," Appl. Phys. Lett., Vol. 62, pp. 219–221, 1993.
29. M.C. Larson, A.R. Massengale, and J.S. Harris, "Continuously tunable micromachined vertical cavity surface emitting laser with 18 nm wavelength range," Electronics Letters, Vol. 32, pp. 330–332, February 15, 1996.
30. F. Sugihwo, M.C. Larson, and J.S. Harris, Jr., "Micromachined widely tunable vertical cavity laser diodes," IEEE Journal of Microelectromechanical systems, Vol. 7, No. 1, March 1998.
31. D. Vakhshoori, P. Tayebati, Chih-Cheng Lu, M. Azimi, P. Wang, Jiang-Huai Zhou, and E. Canoglu, "2mW CW singlemode operation of a tunable 1550 nm vertical cavity surface emitting laser with 50 nm tuning range," Electron. Letters, Vol. 35, pp. 900–901, 1999.
32. D. Vakhshoori, Jiang-Huai Zhou, M. Jiang, M. Azimi, K. McCallion, C.C. Liu, K.J. Knopp, J. Cai, P.D. Wang, P. Tayebati, H. Zhu, and P. Chen, "C-band tunable 6mW vertical cavity surface emitting laser," PD-13, Conference on Optical Fiber Communications, Baltimore, MD, March 2000.

33. J.A. Lott, M.J. Noble, E.M. Ochoa, L.A. Starman, and W.D. Cowan, "Tunable red vertical cavity surface emitting lasers using flexible miro-electro-mechanical top mirrors," IEEE International Conference on Optical MEMS, Kauai, Hawaii, August 2000.
34. D.I. Babic, J. Piprek, and J.E. Bowers, "Long-wavelength vertical-cavity lasers", Chapter 8, ref. [3].
35. G. Steinle, F. Mederer, M. Kicherer, R. Michalzik, G. Kristen, A.Y. Egorov, H. Riechert, H.D. Wolf, and K.J. Ebeling, "Data transmission up to 10 Gbit/s with 1.3 μm wavelength InGaAsN VCSELs," Electron. Lett., Vol. 37, No. 10, pp. 632–634, May 2001.
36. M. Fischer, M. Reinhardt, and A. Forchel, "GaInAsN/GaAs laser diodes at 1.52 μm", Electron. Letters, Vol. 36, 1208, 2000; and M. Reinhardt, M. Fischer, M. Kamp, J. Hofmann, and A. Forchel, "1.3-μm GaInNAs-AlGaAs distributed feedback lasers," IEEE Photonics Technol. Lett., Vol. 12, pp. 239–241, 2000.
37. R.L. Naone, A.W. Jackson, S.A. Feld, D. Galt, C. Embry, K.J. Malone, and J.J. Hindi, "Performance of Monolithic Long-Wave VCSELs as High Speed Sources for Optical Communications," National Fiber Optic Engineers Conference (NFOEC), Baltimore, MD, July 2001.
38. T. Kageyama, T. Miyamoto, S. Makino, Y. Ikenaga, N. Nishiyama, A. Matsutani, F. Koyama, and K. Iga, "Room temperature continuous wave operation of GaInNAs/GaAs VCSELs grown by chemical beam epitaxy with output power exceeding 1 mW," Electron. Lett., Vol. 37, No. 4, pp. 225–227, February 2001.
39. N. Nikolai et al., "Quantum-dot heterostructure lasers," IEEE Journal of Selected Topics in Quantum Electronics, Vol. 6, 439, 2000.
40. Ioffe, QD EL, J.A. Lott, N.N. Ledentsov, V.M. Ustinov, N.A. Maleev, A.E. Zhukov, M.V. Maximov, B.V. Volovik, and D. Bimberg, "Vertical cavity surface emitting lasers with InAs-InGaAs quantum dot active regions on GaAs substrates emitting at 1.3 μm," Conference Digest of International Semiconductor Laser Conference, Monterey, CA, pp. 13–14, September 2000.
41. Coldren, S. Nakagawa, E.M. Hall, G. Almuneau, J.K. Kim, and L.A. Coldren, "Room-temperature CW operation of lattice-matched long-wavelength VCSELs," Electron. Lett., Vol. 36, No. 17, pp. 1465–1467, 2000.
42. S. Nakagawa, et al., "88°C, CW operation of apertured, intracavity contacted, 1.5 mm VCSEL," Appl. Phys. Lett., Vol. 78, March 2001.
43. J. Boucart, C. Stark, F. Gaborit, A. Plais, N. Bouche, E. Derouin, L. Goldstein, C. Fortin, D. Carpentier, P. Salte, F. Brillouet, and J. Jacquet, IEEE Photonics Technology Lett., Vol. 11, p. 629, 1999.
44. W. Yuen, G.S. Li, R.F. Nabiev, J. Boucart, P. Kner, R.J. Stone, D. Zhang, M. Beaudoin, T. Zheng, C. He, M. Jensen, D.P. Worland, and C.J. Chang-Hasnain, "High-Performance 1.6 μm single-epitaxy top-emitting VCSEL," Electron. Lett., Vol. 36, No. 13, pp. 1121–1123, June 22, 2000.
45. R. Stone, R.F. Nabiev, J. Boucart, W. Yuen, P. Kner, G.S. Li, R. Carico, L. Scheffel, M. Jansem, D.P. Worland, and C.J. Chang-Hasnain, "50 km error-free

10 Gb/s WDM transmission using directly modulated long-wavelength VCSELS," Electron. Lett., Vol. 36, pp. 1793–1794, 2000.

46. G.S. Li, R.F. Nabiev, W. Yuen, M. Jansen, D. Davis, and C.J.Chang-Hasnain, "Electrically-Pumped Directly-Modulated Tunable VCSEL For Metro DWDM Applications," European Conference on Optical Communications, Amsterdam, October, 2001.
47. M. Ortsiefer, R. Shau, M. Zigldrum, G. Bohm, F. Kohler, and M.-C. Amann. "90°C Continuous-Wave Operation of 1.83 μm Vertical-Cavity Surface-Emitting Lasers," *IEEE Photonics Technol. Lett.*, Vol. 12, No. 16, p. 1435, November, 2000.
48. M.-C. Amann (invited), "Long-wavelength VCSELs," 13th International Conference on InP and Related Materials (IPRM '01), Nara, Japan, May 14–18, 2001.
49. M. Fischer, D. Gollub, M. Reinhardt, A. Forchel (invited), "GaInAsN based lasers for the 1.3 and 1.5 μm wavelength range," 13th International Conference on InP and Related Materials (IPRM '01), Nara, Japan, May 14–18, 2001.
50. The necessary conditions may not be sufficient, of course, pending on market conditions, timing and applications.
51. M. Kondow, K. Uomi, A. Niwa, T. Kitatani, S. Watahiki, and Y. Yazawa "GaInNAs: Novel material for long-wavelength-range laser diodes with excellent high-temperature performance," Jpn. J. Appl. Phys., Vol. 35, pp. 1273–1275, 1996.
52. M.C. Larson, C.W. Coldren, S.G. Spruytte, H.E. Petersen, and J.S. Harris, "Low Threshold Current Continuous-Wave GaInNAs/GaAs VCSELs," Conference Digest of International Semiconductor Laser Conference, Monterey, CA, pp. 9–10, September 2000.
53. A.J. Fischer, J.F. Klem, K.D. Choquette, O. Blum, A.A. Allerman, I.J. Fritz, S.R. Kurtz, W.G. Breiland, R. Sieg, and K.M. Geib, "Continuous Wave Operation of 1.3 μm Vertical Cavity InGaAsN Quantum Well Lasers," Conference Digest of International Semiconductor Laser Conference, Monterey, CA, pp. 7–8, September 2000.
54. R. Dingle and C.H. Henry, "Quantum effects in heterostructure lasers," U.S. Patent 3982207, September 21, 1976.
55. Y. Arakawa and H. Sakaki, "Multidimensional quantum well lasers and temperature dependence of its threshold current," Appl. Phys. Lett., Vol. 40, pp. 939–941, 1982.
56. G.S. Li, R.F. Nabiev, W. Yuen, M. Jansen, D. Davis, C.J. Chang-Hasnain, "Electrically-pumped directly-modulated tunable VCSEL for metro DWDM applications," ECOC, Amsterdam, September, 2001.
57. M. Yamada, T. Anan, K. Tokutome, A. Kamei, K. Nishi, and S. Sugou, "Low-threshold operation of 1.3-μm GaAsSb quantum-well lasers directly grown on GaAs substrates," IEEE Photonics Technol. Lett., Vol. 12, 774, July, 2000.
58. B. Mason, G.A. Fish, V. Kaman, J. Barton, L.A. Coldren, S.P. DenBaars, and J.E. Bowers, "Characteristics of sampled grating DBR lasers with integrated semiconductor optical amplifiers and electroabsorption modulators,"

Conference on Optical Fiber Communications, TuL6, Baltimore, MD, March 2000.

59. H. Ishii, T. Tanobe, F. Kano, Y. Tohmori, Y. Kondo, and Y. Yoshikuni, "Broad range wavelength coverage (62.4 nm) with superstructure grating DBR laser," Electron. Lett., Vol. 32, 454–455, 1996.
60. M. Oberg, S. Nilsson, K. Streubel, J. Wallin, L. Backbom, and T. Klinga, "74 nm wavelength tuning range of an InGaAsP/InP vertical grating assisted codirectional coupler laser with rear sampled grating reflector," IEEE Photon. Tech. Lett., Vol. 5, 735–738, 1993.
61. R.C. Alferness, U. Koren, L.L. Buhl, B.I. Miller, M.G. Young, T.L. Koch, G. Raybon, and C.A. Burrus, "Broadly tunable InGaAsP/InP laser based on a vertical coupler filter with 57-nm tuning range," Appl. Phys. Lett., Vol. 60, pp. 3209–3211, 1992.
62. G.P. Agrawal, Fiber-Optic Communication Systems, 2nd edition, John Wiley, 1997.
63. G.P. Agrawal and N.K. Dutta, Long-wavelength semiconductor lasers Van Nostrand Reinhold, 1986.
64. S.L. Chuang, Physics of Optoelectronic Devices, John Wiley, 1995.
65. L. Chrostowski, C.H. Chang, R.J. Stone, C. Chang-Hasnain, "Parasitics and Design Considerations on Oxide-Implant VCSELs," Conference Digest of International Semiconductor Laser Conference, Monterey, CA, pp. 7–8, September 2000.
66. K.L. Lear, M. Ochiai, V.M. Hietala, H.Q. Hou, B.E. Hammons, J.J. Banas, J.A. Nevers, "High Speed Vertical Cavity Surface Emitting Lasers," Proceedings of the IEEE/LEOS Summer Topical Meetings, Paper WA-1, pp. 53–54, 1997.
67. D.K. Kuchta, J. Gamelin, J.D. Walker, J. Lin, K.Y. Lau, J.S. Smith, "Relative Intensity Noise of Vertical Cavity Surface Emitting Lasers," Appl. Phys. Lett., Vol. 62, No. 11, 1194–1196, 1993.
68. R.J. Stone, J.A. Hudgings, S.F. Lim, C.J. Chang-Hasnain et al., "Independent phase and magnitude control of an optically carried microwave signal with a three-terminal vertical-cavity surface-emitting laser," IEEE Photonics Technol. Letter, Vol. 11, April 1999; and J.A. Hudgings, R.J. Stone, C.H. Chang, S.F. Lim, K.Y. Lau, and C.J. Chang-Hasnain, "Dynamic behavior and applications of a three-contact vertical cavity surface emitting laser," IEEE J. Selected Topics in Quantum Electronics, Vol. 5, pp. 512–519, 1999.
69. D. Kuksenkov, H. Temkin, "Polarization Related Properties of Vertical Cavity Lasers," in ref. [3].
70. W. Schmid, C. Jung, B. Weigl, G. Reiner, R. Michalzik, K.J. Ebeling, "Delayed Self-Heterodyne Linewidth Measurement of VCSELs," *IEEE Photonic. Technol. Lett.*, Vol. 8, pp. 1288–1290, 1996.

Chapter 14 | Semiconductor Optical Amplifiers

Leo H. Spiekman

Genoa Corporation, Eindhoven, The Netherlands

1. Introduction

As early as the 1960s it was recognized that a gallium arsenide laser diode could be used as an amplifier of infrared light if the reflectivity of its facets was reduced by means of antireflection coatings (Crowe 1966). The laser diode that was used at the time needed an injection current of several amperes and was operated at liquid nitrogen temperatures, to produce gain in a bandwidth that would be considered tiny (30 Å) by present-day standards. It was, however, an early sign of the large interest that would ensue in the semiconductor laser amplifier or semiconductor optical amplifier (SOA), as it would come to be known.

Interest in the SOA took off with the introduction of fiberoptic telecommunication systems, two decades later. A cursory tally of scientific publications over the years shows a great deal of interest in SOAs in the scientific literature, interrupted by a brief dip in the beginning of the 1990s. It is not by accident that this decline of interest coincides with the rise of the erbium-doped fiber amplifier, the EDFA. This device was a gift of nature that provided almost perfect gain in the wavelength band around 1550 nm that is of such high interest to optical telecommunications. It proved to be particularly suitable in the simultaneous amplification of multiple optical channels (wavelength division multiplexing or WDM), which had been troublesome with SOAs because of their speedy gain dynamics, which caused crosstalk between WDM channels. The fact that interest in the SOA picked up after the initial decline was exactly because of its fast gain dynamics, as this allowed all-optical data processing to be carried out in various ways.

The purpose of this chapter, after discussing the basic properties of the SOA, is to present the ways in which these signal-processing capabilities can be used in present and future optical networks, as well as to show how the SOA holds its ground as an amplifier in an EDFA-dominated world.

2. Device Principles

The typical SOA chip is structurally very similar to a laser diode chip (see Chaps. 11 and 12). In fact, the first SOAs were just Fabry–Perot laser chips

OPTICAL FIBER TELECOMMUNICATIONS.
VOLUME IVA

ISBN: 0-12-395172-0

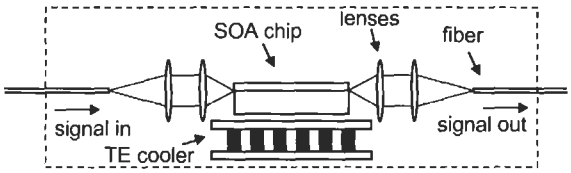


Fig. 1 Typical configuration of a packaged SOA chip. The device has two fiber-chip couplings. Optional items like isolators may be placed in the light path.

with antireflection (AR) coatings applied to their facets to increase the laser threshold current to a level above the operating current of the amplifier. Various advances such as angled stripes and window regions (see below) have improved the suppression of reflections at the facets, but for the most part SOA technology is still semiconductor laser technology.

The common SOA package is very similar to typical laser diode packages. A notable distinction is that two fiber pigtails are found on the SOA—one for the signal entering the amplifier, and another for the amplified signal leaving it (see Fig. 1). All other features usually found in a laser package, such as the thermoelectric cooler (TEC) to keep the chip at room temperature during operation, can also be found in most SOA modules.

2.1. GAIN

The primary property of interest for an amplifier is gain. When semiconductor material with a direct band gap is pumped by injecting a current, a population inversion of electrons and holes can occur in the valence and conduction bands. Upon passing a light signal through the semiconductor, amplification takes place by means of stimulated emission. The part of the electromagnetic spectrum for which emission occurs begins at the band gap energy and goes to shorter wavelengths. The gain medium commonly used for SOAs (as well as telecommunications lasers) is indium gallium arsenide phosphide (InGaAsP). This material has a band gap corresponding to an emission wavelength of 0.9 to 1.65 μm, depending on the composition, which can be smoothly varied from binary InP to ternary InGaAs (Agrawal 1993).

Efficient amplification is obtained by confining the light and the injected carriers to the same thin slab of material (the *active layer*). This is accomplished by enclosing it in two cladding layers of larger band gap and smaller refractive index. The larger band gap forms a barrier that keeps the carriers in the active layer; the smaller refractive index around the active layer forms an optical waveguide (see Fig. 2). The n and p doping of the cladding layers allows for the injection of electrons and holes. In the orthogonal spatial direction, carrier confinement and waveguiding can be achieved by etching a ridge waveguide or by forming a buried waveguide with current blocking regions on both sides. The gain experienced by the optical mode is determined by the material gain

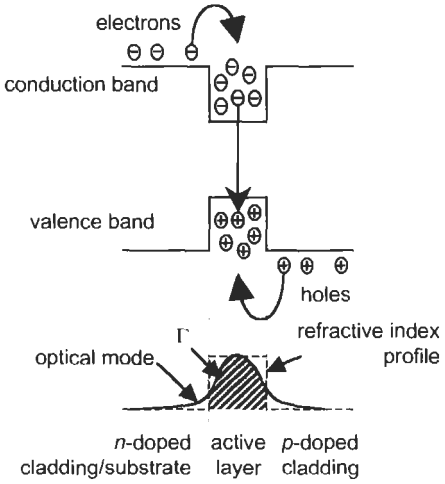


Fig. 2 Schematic of carrier and light confinement in the active layer of a semiconductor optical amplifier. The smaller bandgap confines the electrons and holes, while the larger refractive index introduces optical waveguiding. The fraction of the mode inside the active layer is denoted by the confinement factor Γ.

of the active layer and by the *confinement factor*, denoted by Γ. This is the fraction of the mode inside the active layer (cross-hatched in Fig. 2).

Typically, the fiber-to-fiber gain of a SOA is in the range 15–25 dB, but devices have been made with gains as high as 35 dB (Tiemeijer 1996a). Besides the chip gain, another factor affecting total gain is the fiber-to-chip coupling efficiency. With lensed fibers or lens-fiber assemblies, the coupling loss is reduced to 2–4 dB. The loss associated with butt-coupling the chip directly to a fiber can be lowered by using on-chip spot size converters (Brenner 1993, Kitamura 1999). This method is not yet commonly used in commercial devices.

2.1.1. Gain Ripple

Since the goal is to make an optical amplifier as opposed to a laser, optical feedback that may lead to round-trip resonances must be avoided. The most straightforward method of reducing reflections from the chip facets is the application of AR coatings. Especially for high chip gains, however, an AR coating will not be sufficient to avoid longitudinal resonance that gives rise to *gain ripple* (see Fig. 3). With a chip gain G and a facet reflectivity R, the peak to peak amplitude of the gain ripple will be

$$\text{Ripple} = (1 + GR)^2/(1 - GR)^2$$

(Mukai 1981). Therefore, with a chip gain of 30 dB, a reflectivity smaller than $6 \cdot 10^{-6}$ is needed to achieve a gain ripple less than 0.1 dB.

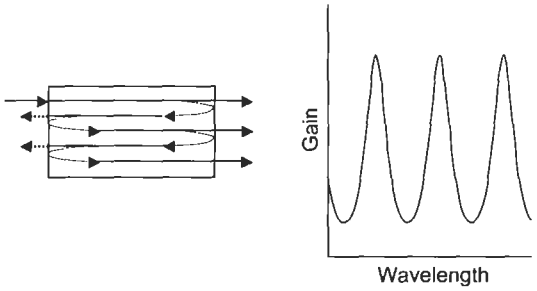


Fig. 3 Reflections at the SOA facets will give rise to round-trip paths which will interfere constructively or destructively depending on the signal wavelength, leading to gain ripple as shown on the right.

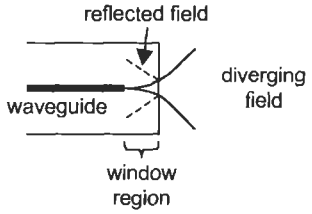


Fig. 4 Top view of an amplifier with window region. The output light diverges before it is reflected back toward the waveguide, thereby reducing the optical power coupled back into it.

A common method to accomplish this is to place the active waveguide stripe on the chip at an angle with respect to the facets. That way, residual reflections returning from the AR-coated facets will be directed away from the waveguide. It has even turned out to be possible to make a SOA without AR coatings this way (Kelly 1996). Obviously, the chip will have to be mounted in its package at an angle in accordance with Snell's law in order to inject and collect the signal light parallel to package coordinates.

A further technique to reduce feedback is the so-called *window region*. This is a section without optical confinement of a few microns length, in which light emanating from the waveguide is allowed to diverge before hitting the facet (see Fig. 4). The reflected light will continue to diverge, and the optical power that is coupled back into the waveguide is minimized. A window region is especially straightforward to fabricate in buried waveguide devices, by simply ending the waveguide stripe a few μm before it reaches the facet.

Combining the three techniques of AR coating, angled stripe, and window region, SOAs can be fabricated with facet feedback close to 10^{-6} (Tiemeijer 1996a).

2.1.2. Polarization Properties

The active layer of a SOA does not have the circular symmetry of the core of an erbium-doped fiber. It customarily has a rectangular shape (typical dimensions: 2 μm wide and a few hundred Ångstroms thick), and in addition is sometimes not made out of bulk InGaAsP but of quantum wells, which adds further anisotropy. This leads to a difference in confinement factor Γ between the two orthogonally polarized modes that is often quite considerable. Without countermeasures, the amplifier would exhibit different amounts of gain for signals in the transverse electric (TE) and transverse magnetic (TM) polarization directions. This polarization dependent gain (PDG) can range from a few to tens of dB, which is clearly not acceptable in any application in which the input signal polarization is not controlled.

Several tactics are possible to remedy this situation. The most obvious one is to restore symmetry between the TE and TM modes by utilizing a square active waveguide. Since a large active waveguide embedded in InP cannot be kept single-mode, the dimensions must be small, e.g., $0.5 \times 0.5\,\mu m$. Although this principle has been demonstrated (Doussiere 1994), a waveguide width this small is hard to control in an industrial process. A second approach is to use the anisotropy of quantum wells to compensate for the difference in confinement factor. It has been shown that quantum well lasers that produce TE polarized light start to emit in the TM polarization when a small amount of tensile strain is introduced in the quantum wells (Thijs 1991). This principle can be used in a SOA by carefully controlling the strain in a stack of tensile strained quantum wells to match the fiber-to-fiber gains of both polarizations (Joma 1993, Ito 1998), or by providing a stack of quantum wells that alternate in tensile and compressive strain, in order to separately control TM and TE gain (Tiemeijer 1993, Newkirk 1993). The latter approach has the advantage of permitting more design freedom, with the associated disadvantage of having more degrees of freedom to control.

It has turned out that the same method also works for bulk active layers. Introducing a small, well controlled amount of tensile strain in bulk InGaAsP allows for compensation of the PDG (Emery 1997). When using this method, the active waveguide must be designed in such a way that the product of layer thickness and required strain does not exceed the critical value for relaxation-free crystal growth. Using any of the methods discussed, the PDG of a SOA can be reduced to a few tenths of a dB.

2.2. OUTPUT POWER

Optical amplifiers are no different from electrical amplifiers in that their output will saturate if too much input power is applied. The power source of an amplifier is finite, and when it is depleted, the gain will drop. This can be seen most easily in a plot of gain versus output power (see Fig. 5). As long as the output power is small, the amplifier is in the small signal regime, and the gain

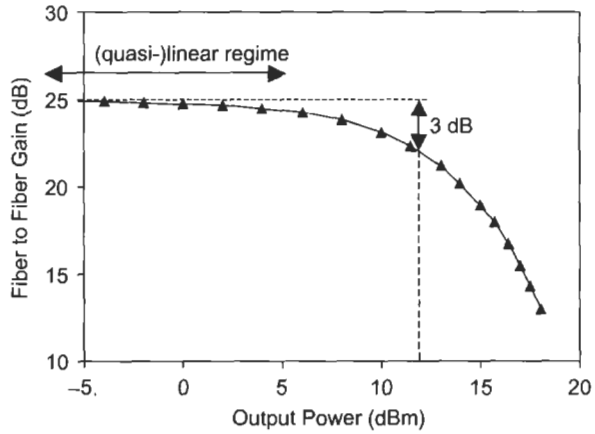


Fig. 5 Gain versus output power curve of a SOA. The gain is quasi-linear in the small signal regime, but saturates at higher power levels. From the gain versus output power graph both the small signal gain and the saturation output power can directly be read.

is constant. With rising input power, the output power will grow to be so large that the small signal gain cannot be sustained. The saturation of an optical amplifier is usually referenced to the output power at which the gain has been compressed by 3 dB, as indicated in the Figure. For a SOA this point is given by

$$P_{3\text{-dB}} = (h\nu A\eta_o \ln 2)/(\tau\Gamma\, dG/dN)$$

which is linear in the active stripe cross-section area A and the output coupling efficiency η_o, and inversely proportional to the carrier lifetime τ, the confinement factor Γ, and the differential gain dG/dN. $h\nu$ is the photon energy (Tiemeijer 1996a). From a gain versus output power measurement, when gain in dB (G[dB]) is plotted against output power in mW (P_{out}), in first-order approximation (that assumes τ and dG/dN to be constants) a linear fit can be made with

$$G[\text{dB}] = G_0[\text{dB}] - (3/P_{3\text{-dB}})P_{out}$$

from which the small signal gain G_0 and the saturation output power $P_{3\text{-dB}}$ are found. The $P_{3\text{-dB}}$ of a SOA is typically between a few and +10 dBm, but values as high as +17 dBm have been demonstrated (Morito 2000).

In dynamic operation, the output power of the amplifier will continually exhibit fluctuations, and since the gain dynamics of the SOA are very fast (typically, $\tau \approx 100$ ps), the instantaneous gain will also swing back and forth in accordance with the saturation curve of the amplifier (Fig. 5). This is the reason a SOA will exhibit crosstalk during multichannel operation, as will be further discussed below.

2.3. *NOISE FIGURE*

In addition to the stimulated recombination of carriers that produces gain, electrons and holes in the active layer are also incessantly recombining spontaneously. The photons emitted in this process are subsequently amplified, giving rise to the *amplified spontaneous emission* (ASE) spectrum of the SOA (see Fig. 6).

In an optically amplified system, the ASE plays the part of the noise that is added to the signal. The noise properties of an amplifier can be quantified by its noise figure (NF), which for optical amplifiers is defined as the deterioration in signal-to-noise ratio (SNR) with a purely shot noise limited input signal. This is not a very practical definition since optical signals are rarely shot noise limited. A more useful definition looks at the signal-spontaneous emission beat noise at the receiver that is caused by the ASE noise of the amplifier, which, in amplified lightwave systems that use a bandpass filter to limit the ASE power reaching the detector, is usually the dominant noise term (Olsson 1989). In this approximation, the noise figure of an optical amplifier is denoted as

$$\mathrm{NF} = 2n_{sp}/\eta_i$$

in which $n_{sp} = N_2/(N_2 - N_1)$ is the inversion parameter of the amplifier (i.e., the degree of population inversion, with N_1 and N_2 the fractional number of carriers in the ground and excited states, respectively), and η_i is the input coupling loss (Walker 1989). The noise figure of an optical amplifier can easily

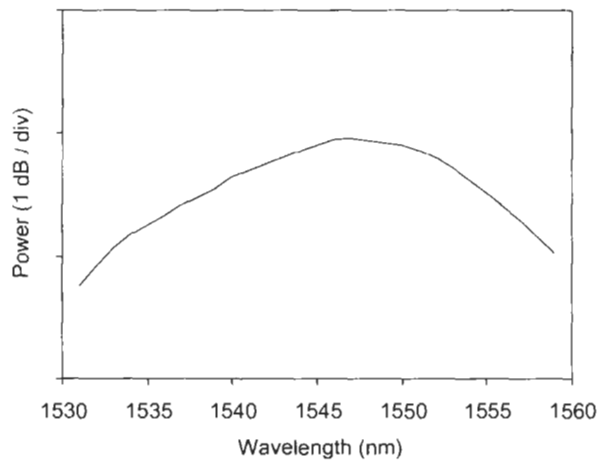


Fig. 6 Amplified spontaneous emission spectrum of a SOA, which typically has a smooth parabola-like shape.

be determined by measuring the gain and the output ASE power and taking

$$\mathrm{NF} = 2\rho_{\mathrm{ASE}\|}/Gh\nu$$

in which $\rho_{\mathrm{ASE}\|}$ is the power density of the part of the ASE that is copolarized with the amplified signal, G is the gain of the amplifier, and $h\nu$ is the photon energy. In a polarization-independent amplifier, $\rho_{\mathrm{ASE}\|}$ is just half the total measured ASE power density. Even though the inversion factor of a SOA can be as high as that of an EDFA, the input (fiber-chip) coupling loss clearly puts the SOA at a disadvantage. Nevertheless, devices with a noise figure as low as 6 dB have been demonstrated (Tiemeijer 1996b).

2.4. GAIN DYNAMICS

2.4.1. Gain Compression and Recovery

In Fig. 5 the reduction of gain was shown that occurs when the SOA is required to generate high output power. This so-called *gain compression* results because carriers (electrons and holes) are burned up faster by stimulated recombination. Under continuous (CW) operation, a new steady state sets in at a reduced degree of population inversion. When the gain is compressed by an intense optical pulse, it recovers with a characteristic timescale called the *carrier lifetime* (the τ previously mentioned), which is typically a few 100 ps (see Fig. 7).

Although gain compression and recovery will cause interchannel crosstalk in a WDM system (Inoue 1989), it can also be put to good use in all-optical

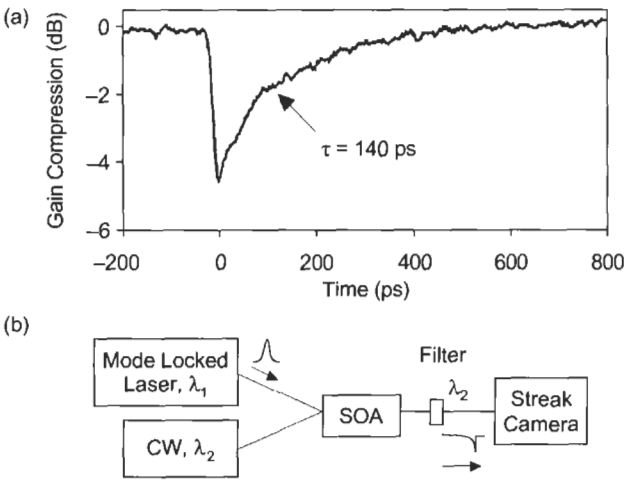


Fig. 7 a) Fast gain compression and slower recovery of a SOA by an intense optical pulse. The gain recovery time is determined by the carrier lifetime. b) Pump-probe setup with which the curve was measured.

signal processing applications such as wavelength conversion. The gain compression due to the intensity of one optical channel can influence the intensity of another channel, which is known as cross-gain modulation (XGM) (Glance 1992). Likewise, it can influence the phase of another channel by means of the associated variations in the carrier density, which lead to variations in refractive index of the active layer: cross-phase modulation (XPM) (DiJaili 1992). Applications that use these mechanisms will be discussed below.

2.4.2. Four-Wave Mixing

A different interaction mechanism occurring in the SOA is four-wave mixing (FWM) (Inoue 1987b). This process occurs when two signals of different wavelength are injected into the SOA. The intensity beating that arises at the difference frequency of the two signals will modulate the carrier distribution in the SOA. At small enough frequency separation, the carrier density will be modulated (interband carrier dynamics). When the detuning between the signals becomes larger, modulation of the intraband carrier distribution will dominate, which is reflected in processes such as spectral hole burning and carrier heating (Wiesenfeld 1996).

The modulation driven in these processes will set up a moving grating in both amplitude and phase in the SOA active stripe, producing both gain and refractive index effects. The input signals will scatter from this grating, producing side bands as shown in Fig. 8, located at higher and lower frequencies according to the difference frequency between the beating input signals. The power in the side bands is usually small relative to the signal power for

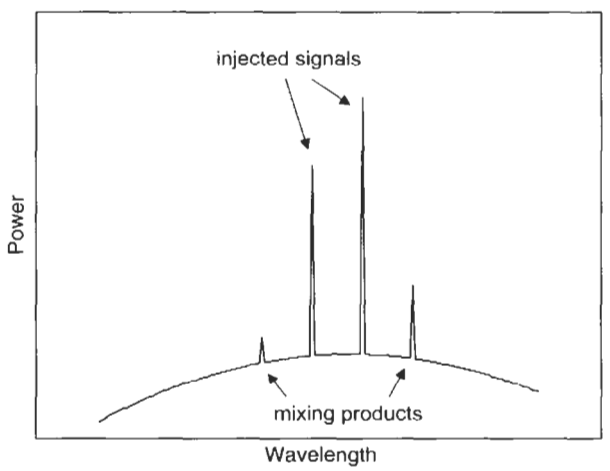


Fig. 8 Four-wave mixing in a semiconductor optical amplifier. Mixing products appear on both sides of two strong injected signals.

detunings larger than a few nanometer, and the side bands ride on a significant background of amplified spontaneous emission noise.

The four-wave mixing process between two input signals requires that they have parallel components. To ensure an efficiency that is independent of the polarization of one of the input signals, as would be needed in nonlinear signal processing applications, therefore requires complicated experimental arrangements. Nonetheless, the response time for four-wave mixing is very fast and both intensity information and phase information are transferred to the mixing products, giving FWM a distinct advantage in some all-optical processing applications.

2.5. *GAIN CLAMPING*

As discussed above, cross-gain modulation will inflict distortion on WDM channels when a SOA is used for multichannel amplification. This could be avoided if a method was available to clamp the carrier concentration in the active layer to a fixed level. That can be accomplished by introducing lasing action into the amplifier (Simon 1994, Bauer 1994), because in a laser the round-trip gain is constant and equal to the round-trip loss by virtue of the lasing condition. A gain-clamped SOA (GC-SOA) can be constructed by appending small distributed Bragg reflector (DBR) mirrors to both ends of the active waveguide (Bachmann 1996), which should produce wavelength-selective feedback at a wavelength just outside the range of interest for amplification (see Fig. 9). That way, the laser light can be separated from the amplified channels using a filter.

A SOA thus gain-clamped exhibits a gain versus output power curve as shown in Fig. 10. With increasing input power, the gain remains constant due to the laser action, and the power in the lasing mode decreases steadily. This continues until the signal has consumed all of the laser power, at which point

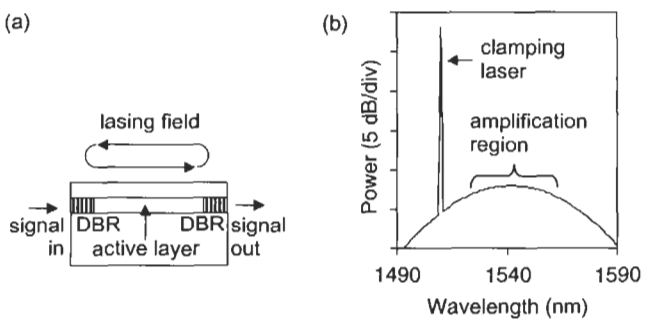


Fig. 9 a) A gain-clamped SOA constructed by introducing DBR mirrors in the waveguide on both sides of the gain stripe. This causes the device to lase at a wavelength chosen to be outside the desired gain region. b) Lasing spectrum of the GC-SOA.

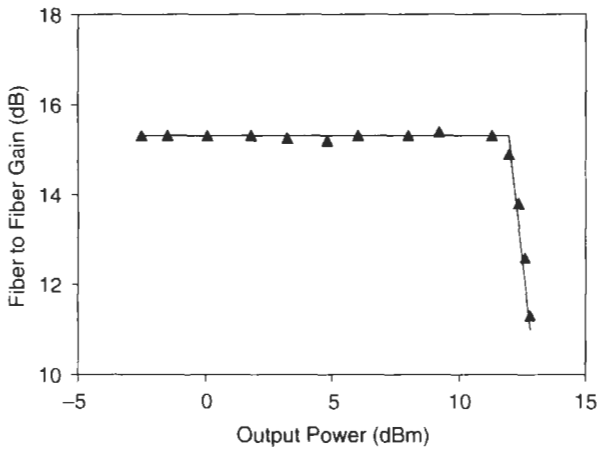


Fig. 10 Gain versus output power of a GC-SOA, showing clamping of the gain to a fixed level up to a certain maximum output power, which represents the condition of the laser going below threshold.

the laser turns off. (This is at $P_{out} = +12$ dBm in Fig. 10.) In a way, the signals being amplified and the laser are communicating reservoirs, keeping the power in the laser cavity constant at all times.

An important drawback of gain clamping a SOA using a laser is the relaxation oscillations that occur under dynamic operation. The laser in the GC-SOA is in fact being modulated at the bitrate of the signals being amplified, and since the typical relaxation oscillation frequency of this type of device is a few to ten GHz, this poses serious problems at the bitrates of interest for optical telecommunications (Pleumeekers 1997, Selbmann 1999).

Recently a device was introduced that aims at solving this problem by using the vertical lasing field of an integrated vertical cavity surface emitting laser (VCSEL) to clamp the gain (Francis 2001). The relaxation oscillation time of the VCSEL is extremely short, allowing operation up to much higher bitrates. Both gain-clamped operation for long (microsecond) timescales with less than 0.05 dB gain variation, and penalty-free WDM operation at 10 Gb/s were demonstrated. An additional advantage of using a vertical laser is that there is no need to filter out the clamping light at the output of the device.

3. Applications

After the preceding descriptions of gain dynamics and the problems associated with gain clamping, it may have become a little clearer why the EDFA was so quick in winning ground on the SOA in the late 1980s. Apart from its somewhat better noise figure and generally higher output power, the EDFA

does not suffer from nonlinearities like the SOA does. Its gain does saturate at higher output powers, but this occurs at a timescale of milliseconds rather than nanoseconds—much slower than the bitrate. A stream of random ones and zeroes will be perceived by an EDFA as continuous light of average power, while a SOA will react to every bit. As said, although this limits the applicability of the SOA as a linear amplifier, it does enable its use as an all-optical processor. Applications in this domain will be dealt with below. First, some applications of the SOA as an amplifier will be looked at in closer detail.

3.1. SINGLE-CHANNEL AMPLIFICATION

3.1.1. Digital Transmission

The limits of single-channel digital transmission using SOAs have been tried in a laboratory experiment (Kuindersma 1996) and a field trial (Reid 1998). In single-channel transmission, no crosstalk with other channels needs to be feared, putting the focus mainly on signal-to-noise ratio degradation along the transmission line. Care must be taken, however, to avoid the occurrence of intersymbol interference (ISI), which arises if the amplifiers are operated in deep saturation and have a gain recovery rate τ comparable to the bitrate (Saleh 1990).

A schematic of the experiment is shown in Fig. 11. The experiment was carried out at a wavelength of 1310 nm, at the dispersion zero of standard fiber. A single channel, modulated with a 10 Gb/s return-to-zero (RZ) data stream, was launched into a transmission line consisting of eleven 38-km spans of fiber, which contained a total of twelve SOAs including the booster amplifier and the receiver preamplifier. The transmission was limited by accumulation

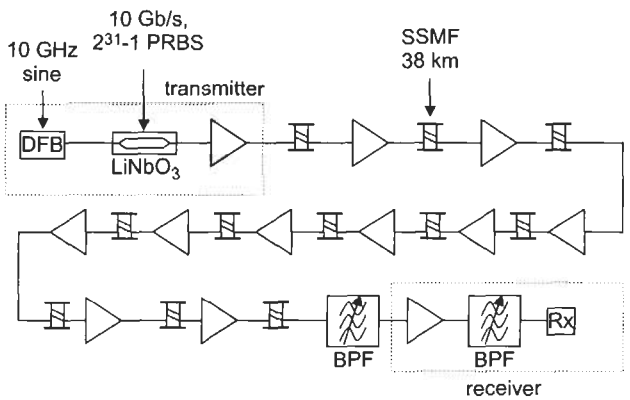


Fig. 11 Single-channel 10 Gb/s RZ transmission at 1310 nm using SOAs. The 11 spans cover a total distance of 420 km. A transmission penalty of 5 dB was observed without a BER floor.

of spontaneous emission in the signal bandwidth, and showed a penalty of 5 dB without an observable BER floor. If it were not for dispersion, the loss budget of over 200 dB associated with the total distance spanned of 420 km would have spanned 1000 km at 1550 nm.

The goal of the field trial was to connect the European capital cities of Madrid, Spain and Lisbon, Portugal, with a fiberoptic link. Although installed fiber plant is much less well behaved than spools of fiber in the lab, and span lengths vary greatly depending on the location of repeater huts, the total distance of 810 km could be spanned using only one regenerator between two all-optical trajectories of 460 and 350 km, respectively (Reid 1998). Video transmission over the link was demonstrated at the European Conference on Optical Communication (ECOC '98) in Madrid and at the World Expo in Lisbon.

3.1.2. Analog Transmission

Analog modulation has far more stringent requirements of linearity than digital transmission. Nonlinearities in an analog transmission line will distort a sine wave, generating unwanted second- and third-order products called composite second-order (CSO) and composite triple-beat (CTB). The variation of gain with output power in a standard SOA would introduce too high a level of distortion for common analog applications; only a gain-clamped SOA offers sufficient linearity (Tiemeijer 1996c).

This has been demonstrated in a cable television transmission experiment using a full 77-channel NTSC load. A gain-clamped SOA module was used to increase the power budget between the transmitter and the receiver (Mutalik 1997) (see Fig. 12). The SOA module delivers CSO/CTB performance comparable to inserting an optical-electrical-optical regenerator in the transmission line. In this application, the relaxation oscillation of the GC-SOA does not pose a problem, because the frequencies used in a cable television system remain well below the point where these relaxation oscillations start to impact performance.

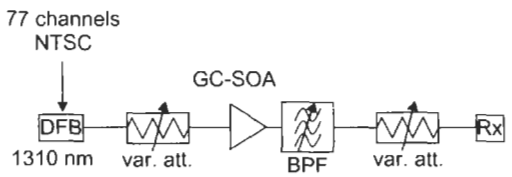


Fig. 12 Experimental setup to evaluate analog performance of GC-SOA using 77 NTSC carriers. The attenuators simulate fiber spans. CSO/CTB distortion values of −55 dBc were measured.

3.2. WDM AMPLIFICATION

Wavelength division multiplexing has long been the exclusive domain of the EDFA. The crosstalk problem of the SOA, caused by its fast gain dynamics, which, by XGM, imprints the data of each channel onto every other channel, seemed insurmountable. Yet, the SOA remains tempting as a cheap and compact amplifier technology. Therefore, several groups have undertaken work to enable operation of SOAs in a WDM environment. Two approaches can be discerned: one, avoiding the crosstalk issue by operating the amplifier at sufficiently low output power that the gain variation due to XGM becomes small, and the other, operating at higher powers and utilizing techniques for suppression of gain variations. Since the noise figure and output power attainable by SOAs still trail the performance of EDFAs, application in long-haul trunk networks is not expected. The results described below indicate, however, that application in modest-distance, moderate-channel-count environments such as access or metro networks may be possible.

3.2.1. Avoiding Interchannel Crosstalk

As was seen in Fig. 5, most of the gain versus output power curve of a SOA is fairly flat. Interchannel crosstalk due to gain variations can be made arbitrarily small by working in a lower output power regime. Therefore, when using SOAs in WDM transmission, there is a tradeoff between number of channels and span length (which determine the required power to be launched in a span of fiber), and the crosstalk distortions that can be tolerated. At the same time, the amplifier gain needs to be carefully matched to the span loss. In an EDFA-based system, the amplifiers would automatically go into an average saturation level in accordance with the span loss, but this method cannot be used in a SOA system, because saturation is exactly what we want to avoid.

With present SOA saturation output powers and noise figures, spans of 80 km present too much loss for all systems except those with the smallest number of channels, but 40-km spans (plus some additional loss budget for dispersion compensation) are feasible. Experiments in several configurations have been undertaken (8×10 Gb/s over 240 km (Spiekman 2000d), 8×20 Gb/s over 160 km (Spiekman 2000a), 8×40 Gb/s over 160 km (Spiekman 2000c), 32×10 Gb/s over 160 km (Spiekman 2000b)), one of which will be discussed here in more detail.

Figure 13 shows the experimental layout of the 8×10 Gb/s experiment. Eight WDM channels spaced at 200 GHz are comodulated and decorrelated, and are launched into a transmission line consisting of six spans of 40 km of standard single mode fiber (SSMF). Each span has an appropriate amount of dispersion-compensating fiber (DCF) to keep the dispersion nominally zero. For each span, a single SOA, packaged in an industrial standard butterfly package, is used to compensate the loss of SSMF and DCF. Biased at 400 mA, the SOAs have noise figures between 8 and 10 dB, saturation output powers of

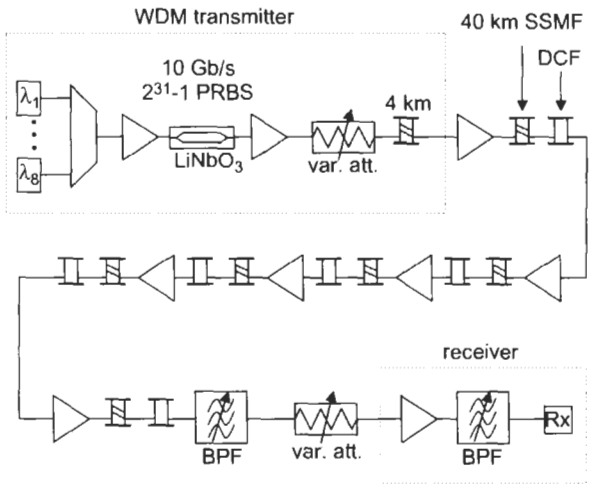


Fig. 13 Transmission of 8 WDM channels modulated at 10 Gb/s across six 40-km spans of standard fiber using a total of 9 SOAs.

+12 dBm, and gains (at the transmission wavelengths) equal to the span loss of about 13 dB. The output power of each SOA is kept about 2 dB below P_{sat}, for a maximum gain compression of 1 dB. The receiver is a SOA-preamplified pin photodiode.

With all eight channels operating simultaneously, bit error rate (BER) curves (see Fig. 14) show a small power penalty relative to baseline of around 1 dB, and go straight down to 10^{-11} without a hint of an error floor. As a function of the power launched into the first fiber span, the BER exhibits a minimum, as shown in Fig. 15. With lower launched power, the signal-to-noise ratio at the receiver declines. With higher launched power, although the received SNR is improved, the SOAs in the transmission line are driven further into saturation, worsening the distortion due to interchannel crosstalk. The optimum operating point of the transmission line represents the best balance between SNR and crosstalk distortion.

3.2.2. Suppressing Interchannel Crosstalk

Several techniques can be used to make a SOA think it is amplifying continuous (CW) light, while in reality it is amplifying data. One is to use polarization modulation instead of intensity modulation. This does have the disadvantage of requiring special transmitters and receivers, and it suffers from the polarization-dependent loss of any component in the transmission line, which converts part of the polarization modulation back into intensity modulation. Operation of the SOA in saturation is possible, however, as has been demonstrated (Yu 2000, Srivastava 2000).

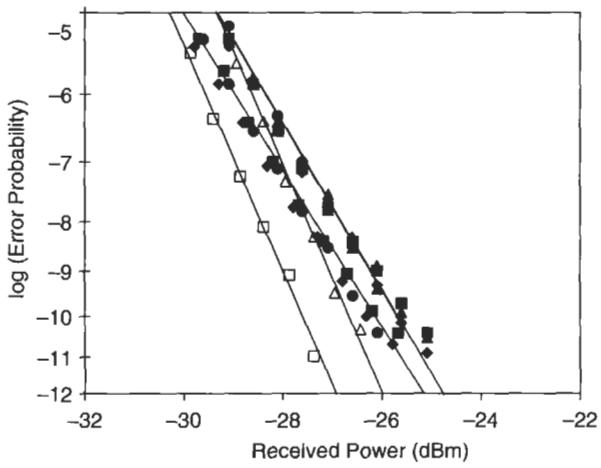


Fig. 14 Bit error rate curves of the eight transmitted WDM channels after 240 km and 9 SOAs. Open symbols are baselines. Transmission penalties are around 1 dB. Lines are fitted to the channel 1 and channel 8 data.

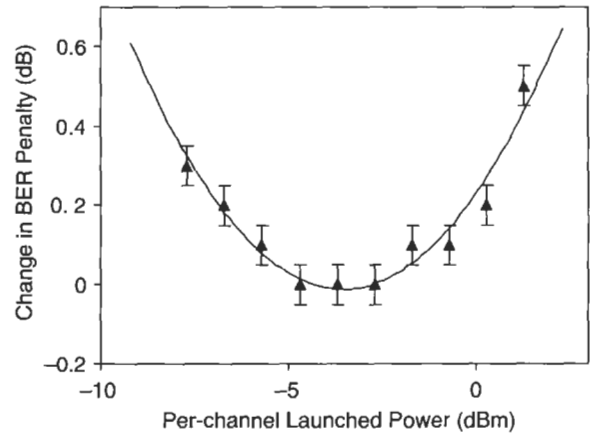


Fig. 15 Sensitivity of transmission performance to launched power. At lower powers, lower OSNR degrades performance. At higher powers nonlinearities in the SOAs dominate.

A similar way is to use two tightly spaced wavelengths for each channel that are modulated 180° out of phase, creating a type of frequency shift keying transmitter (Kim 2000). Thus the optical power in the amplifiers is constant, leading to minimal XGM. After transmission, the data can be demodulated by using a filter. Apart from the obvious disadvantage of this scheme of requiring twice the number of transmitter lasers and a special modulator, the claim of

constant power holds only as far as the two complementary bit streams are not separated by dispersion.

A method that cannot claim as perfect a stabilization of total power but that is more practical is the use of a *reservoir channel.* This is an intense CW optical beam that is colaunched with the data channels and that can be viewed as the addition of a significant number of channels whose data content is always "1." This compresses the statistical distribution of instantaneous output powers in such a way that the gain variation of the inline SOAs is significantly reduced. This principle has been demonstrated in 32 × 2.5 Gb/s transmission over 125 km (Sun 1999). Three SOA-amplified spans (without dispersion compensation) of 42 km were used, and bit error rates were measured with and without a reservoir channel. Without it, operation with low error rate was not possible, while operation with no measurable BER and a power penalty of 2 dB was demonstrated with the reservoir channel present.

Further analysis of a system with and without reservoir channel reveals that the optimum launched power is widely different in these two situations. During transmission of 32 × 10 Gb/s, 50-nm spaced channels over four 40-km spans (plus DCF) without a reservoir channel, an optimum total launched power of −7.5 dBm was found (see Fig. 16), while the addition of a CW reservoir channel moved this optimum to a launched power of −3.5 dBm (Spiekman 2000b). The error rate measured in the optimum was almost identical in both situations, however, illustrating that in this experiment the reservoir channel did not lead to an absolute improvement, but rather shifted the most advantageous operating point to higher powers. The curve for no-reservoir operation

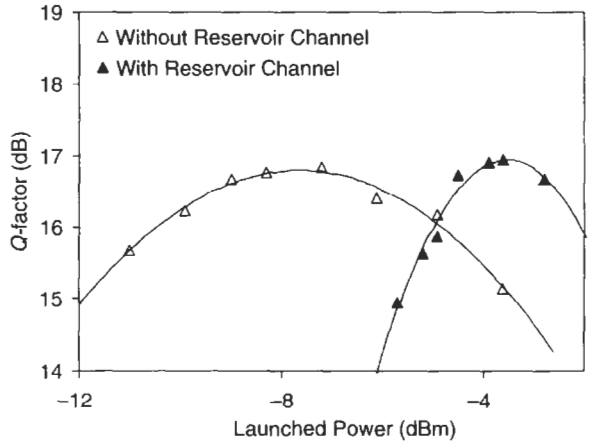


Fig. 16 Q-factor versus total launched power (into first inline amplifier) for transmission with and without reservoir channel. The optimum for transmission with the reservoir channel is located at significantly higher launch power. However, the optimal Q values are almost identical in both cases.

in Fig. 16 is wider, pointing out that a lower-power, more linear system is more forgiving to variations in launched power.

3.2.3. Amplification of Bursty Data

Before the impression would develop that the fast gain dynamics of the SOA are always a disadvantage in amplification applications, we should take a look at a system transmitting bursty data. It is well known that EDFAs, due to their millisecond gain dynamics, will exhibit transient effects when data channels are switched on or off (Srivastava 1997). When that happens, the EDFA will readjust its gain according to the new average output power, upon which the next amplifier in line will react to both the adding or dropping of the channels and the gain adjustment of the previous amplifier. In a SOA-amplified system, the gain dynamics take place at the speed of the bitrate, and a channel dropped seems to just transmit zeroes from that time onward. Therefore, the transient effects impeding packet- and channel-switched data transmission in an EDFA-amplified system do not affect a system in which SOAs are used.

An experiment demonstrating this was configured as follows (Gnauck 2000). Sixteen 100-nm spaced channels were modulated at 10 Gb/s. The eight even channels were led through a fast optical switch before reaching the data modulator, and were switched on and off with various duty cycles, at frequencies varying from 3 Hz to 50 MHz. The data channels were launched into four spans of 40 km of SSMF (plus the appropriate amount of dispersion compensation) and were received in an EDFA-preamplified receiver. Care was taken that only one—nonswitched—channel at a time was seen by the EDFA, to prevent it from generating transients.

Received spectra of eight channels (even channels off), sixteen channels (even channels on), and eight stationary and eight switched channels (50% duty cycle) are shown in Fig. 17. A small difference of 0.5 dB in average power is observed in the stationary channels between the 8 and 16 channel conditions, related to a residual gain variation versus output power; the SOAs were not rigorously operated in the small-signal domain. This variation led to a small amount of eye closure in the non-AGC receiver, producing a penalty in Q-factor relative to non-switched behavior of 0.5–1 dB. This penalty did not vary significantly over the whole range of switching speeds (3 Hz–50 MHz) and for various duty cycles (see Fig. 18), implying robust behavior under operation with bursty data.

3.3. ALL-OPTICAL SIGNAL PROCESSING

The carrier dynamics of the SOA make it a very rich device that can be applied in various ways for all-optical signal processing. The earliest of those applications was wavelength conversion, in which the data modulated on one wavelength channel are transferred to another signal wavelength. The nonlinear response of the SOA enables regeneration of the data to take place, while

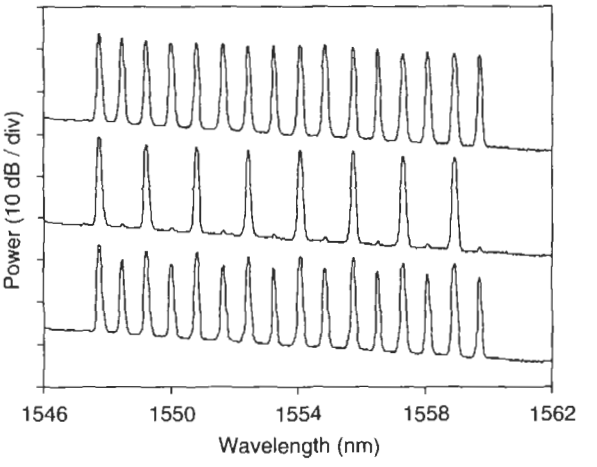


Fig. 17 Received spectra of channel dropping experiment. 1) Even channels on; 2) even channels off; 3) even channels switched in and out at 100 kHz.

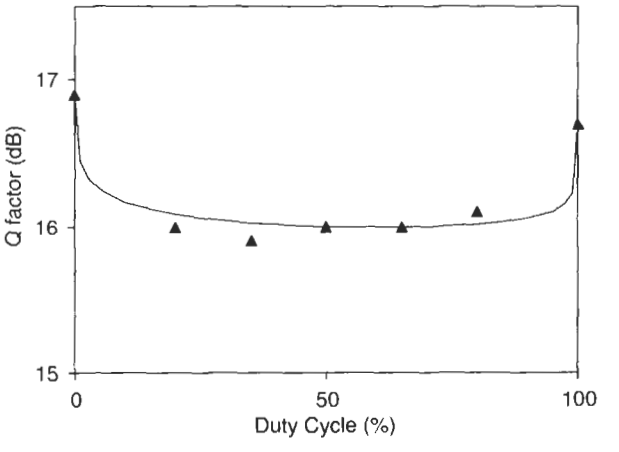


Fig. 18 Effect of varying duty cycle in the channel dropping experiment. A small and constant penalty of 0.7 dB is observed relative to the stationary (8 or 16 channels on) states. The line represents a calculated result.

the fast gain dynamics allow for applications in optical time-division multiplexing. These three applications of the SOA will each be briefly discussed.

3.3.1. Wavelength Conversion

3.3.1.1. Cross-Gain Modulation

Wavelength conversion can help reduce the blocking probability in all-optical WDM networks. This is traditionally accomplished by converting the signal

to the electrical domain and retransmitting it at another wavelength. It was recognized early on that the same could be accomplished by employing the interaction between wavelength channels in a SOA offered by the mechanism of cross-gain modulation. The data on one wavelength (commonly called the *pump*) modulates the carrier density in a SOA, and the resulting gain variations imprint an inverted copy of the data onto a CW signal (the *probe*) injected into the SOA at the same time (Wiesenfeld 1992, Joergensen 1993). Use of this mechanism is attractive from a cost standpoint, especially at very high data rates where electronics would become extremely expensive. The speed may be increased by operating the SOA under high optical intensity in order to reduce the gain recovery time due to stimulated carrier recombinations (Manning 1994). In a wavelength conversion setup, this can be done by using a high power probe (Wiesenfeld 1993). Wavelength conversion based on cross-gain modulation has been demonstrated to operate at bitrates as high as 40 Gb/s (Joergensen 1996a).

Two problems are associated with XGM wavelength conversion. The first is that due to bandfilling, it is more difficult to compress the gain at longer wavelengths than it is at shorter wavelengths (Inoue 1987a). Therefore, there is an extinction ratio penalty associated with wavelength conversion toward longer wavelengths. The second problem is that the gain compression caused by the modulation of the carrier density is accompanied by a phase modulation due to the associated change in refractive index. This results in chirping of the converted signal, which seriously limits the ability to transmit this signal over dispersive fiber (Perino 1994).

3.3.1.2. Cross-Phase Modulation

The chirp imparted on the converted signal in XGM wavelength conversion can be used to advantage by including the SOA in an interferometric structure that converts this cross-phase modulation (XPM) into an intensity modulation. This was first done using discrete components in a Michelson (Mikkelsen 1994a) and a Mach-Zehnder interferometer configuration (Durhuus 1994), but stability requirements prompted the development of integrated versions (Mikkelsen 1994b, Idler 1995, Pan 1995, Ratovelomanana 1995).

A phase shift of only π is needed to obtain complete extinction in an interferometer, which can be achieved with only a few dB of gain compression in the SOA. The phase shift is virtually independent of wavelength, so conversion to longer wavelengths is no problem with XPM. In addition, by using different biasing conditions of the interferometer ("normally off" versus "normally on"), inverting or noninverting modes of operation can be selected that have different signs for the chirp. The noninverting mode of operation produces negative chirp that will initially compress pulses in standard single-mode fiber, which results in an improvement of the transmission distance that can be achieved.

The most important disadvantage of the interferometric structure is that a phase shift of more than π will cause an overshoot, which will impair the extinction ratio. This can be remedied by controlling the bias conditions of the SOAs in the interferometer (Joergensen 1996b), or by adjusting the gain of preamplifiers that can be integrated with the device (Spiekman 1998, Janz 1999a).

The CW beam onto which the data will be converted can be injected either copropagating or counterpropagating with the data signal. Counterpropagating operation has the advantage of not requiring a filter to separate the two wavelengths, but time-of-flight issues limit the bitrate. A solution that offers copropagating operation but does not require a filter is to inject the data into a higher-order waveguide mode and use a mode filter to separate the converted data from the original signal (Leuthold 1999). This dual-order mode (DOMO) Mach-Zehnder device accomplished an extinction of 28 dB between the two wavelengths (Janz 1999b).

Very high conversion speeds can be reached in a push-pull configuration, e.g., by using a Mach-Zehnder interferometer device in which the data is injected in both arms with a time difference equal to the bit period. The slow gain recovery in both arms is canceled out, gating the interferometer for exactly one bit slot (Mikkelsen 1997). An alternative method uses XPM in a single SOA that is subsequently converted to intensity modulation in a passive Mach-Zehnder interferometer that has a one-bit-period delay in one of its arms. The highest bitrate demonstrated using this principle is 100 Gb/s (Leuthold 2000).

3.3.1.3. Four-Wave Mixing

Four-wave mixing (FWM) is an ultrafast all-optical effect that can be used, among other things, for high-bitrate wavelength conversion (Kelly 1998). Since FWM preserves the phase of optical signals in addition to their intensity, it can handle intensity and phase modulated signals as well as frequency shift keying when used for wavelength conversion, offering enhanced transparency in all-optical networks. Its polarization dependence issues have been addressed in dual-pump and polarization diversity schemes of various kinds (Jopson 1993, Schnabel 1994, Hunziker 1996, Lacey 1997, Lin 1998). By careful optimization of the input powers, the power penalties associated with the signal-to-noise ratio degradation in FWM wavelength conversion can be minimized (Summerfield 1996). However, the conversion efficiency is not great and very dependent on the distance between signal and converted wavelength (Zhou 1994), as illustrated in Fig. 19. Therefore, this scheme is not expected to be widely used in future all-optical networks.

Another application of FWM is in dispersion management, by means of *optical phase conjugation*. In the spectral domain, the four-wave mixing product is a mirror image of the original signal which has the opposite chirp. By

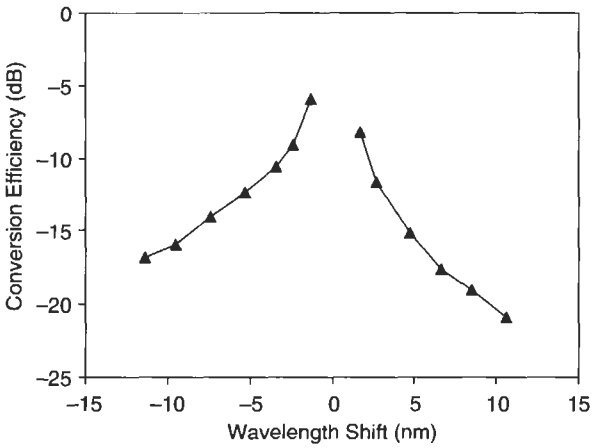


Fig. 19 Typical shape of the dependence of conversion efficiency of four-wave mixing on wavelength shift.

placing the SOA in the middle of a transmission span, all pulse distortion due to dispersion that occurred in the first half of the span will be undone in the second half, yielding undistorted bits at the output (Tatham 1993, 1994, Feiste 1999).

3.3.2. Regeneration

The transfer function of the interferometers used for XPM wavelength conversion is nonlinear enough to be used for partial 2R regeneration (*reamplification* and *reshaping*) of the signal (Mikkelsen 1996, Wolfson 2000). The sinusoidal response function of the interferometer redistributes the noise at the 0 and 1 rails, leading to limited 2R regeneration. Retiming can be accomplished when the interferometer is used to gate a pulse source that is synchronized with the incoming data (Eiselt 1993a). Combination of the two techniques as shown in Fig. 20 yields a 3R regenerated signal (Jepsen 1998). Transmission of 10 Gb/s data over a total distance of 200,000 km comprising over a thousand regenerated spans has been demonstrated using this type of regenerator (Lavigne 1998).

Chapter 15 of this book is dedicated entirely to regeneration of optical signals.

3.3.3. Optical Time-Division Multiplexing

In addition to wavelength conversion, the mechanism of cross-phase modulation in SOAs can also be employed for temporal demultiplexing of optical time-division multiplexed (OTDM) signals. The phase shift triggered by

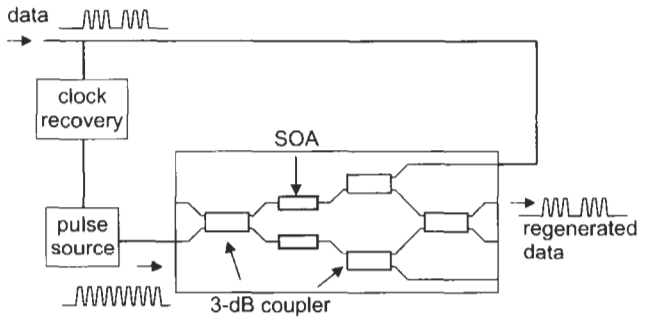


Fig. 20 3R regeneration by using an interferometer to gate a pulse train obtained by clock recovery.

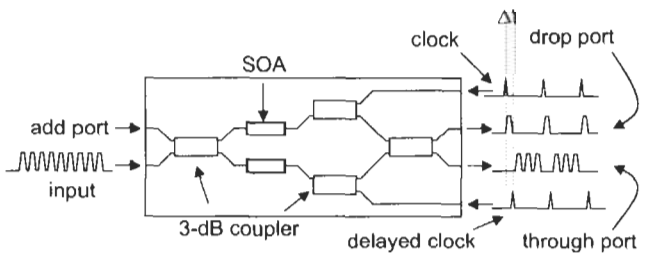


Fig. 21 Mach-Zehnder interferometer used as an OTDM add-drop multiplexer.

a properly timed control pulse train injected into the SOA can be used to select one in every n bits. This was shown first for 9 to 3 Gb/s demultiplexing (Eiselt 1993b) in a SLALOM (Semiconductor Laser Amplifier in a LOop Mirror) interferometer (Eiselt 1992). By using intense control pulses, operation at much higher speed is possible (Ellis 1993, Morioka 1996). Loop interferometers have even been shown to be integrable on a single chip (Jahn 1996). The same principle is applicable to various device configurations. Demultiplexing a 40-Gb/s bitstream to 10 Gb/s has been shown in a single-interferometer-arm configuration that cancels out the effects of long-lived refractive index nonlinearities (Patel 1996). In a push-pull configuration, the same cancellation can be accomplished in a Mach-Zehnder interferometer. In an integrated device, 80 to 10 Gb/s demultiplexing was demonstrated (Hess 1998a). The interferometer in this device is equipped with two input and two output ports (see Fig. 21), allowing simultaneous adding and dropping of TDM channels from a bitstream going through the device (Hess 1998b).

A demultiplexing approach that is highly linear and does not add spontaneous emission noise to the signal is the *gain-transparent switch*, a SLALOM-type demultiplexer with a 1300-nm SOA. The control pulses at 1300 nm present a phase change to the signal at 1550 nm without an associated gain change.

Operation in 40 to 10 Gb/s demultiplexing has been demonstrated (Diez 1999). Gain-transparent switching in a Mach-Zehnder interferometer was done up to 160 Gb/s (Diez 2000).

3.4. *SOAs IN ACCESS NETWORKS*

The WDM experiments described earlier indicate the possible use of the SOA as an inline amplifier in metro or access networks. Network architectures have been proposed, however, in which the utility of the SOA goes much further (Iannone 2000). By directly modulating the injection current of a SOA, it can be used as a polarization independent modulator. An access architecture can make use of this by placing all signal lasers at the central network node, and placing a SOA modulator in the end user nodes (Fig. 22). Because the network sources all wavelengths in this architecture, wavelength registration is considerably simplified. In addition, all end user nodes can be identical, irrespective of which wavelength each user is served by. This architecture has been demonstrated at a bitrate of 622 Mb/s, with electrical preemphasis of the high frequency components to compensate for the rolloff of the SOA response

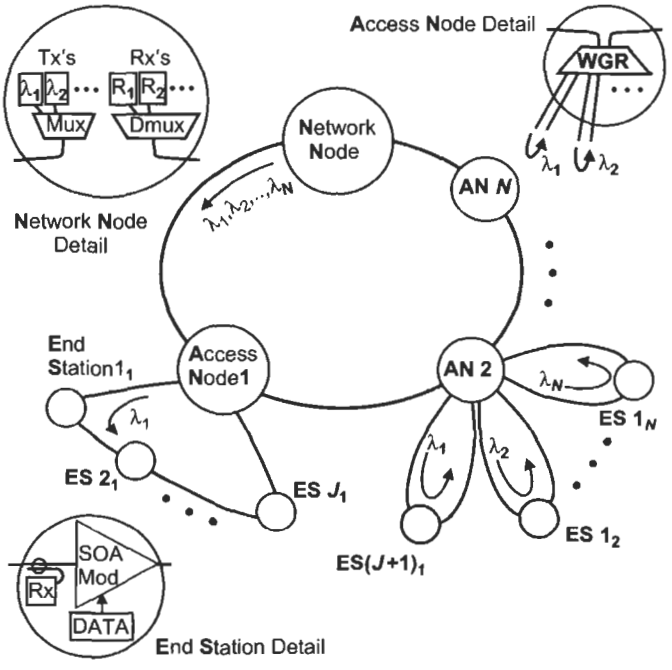


Fig. 22 Access architecture with SOAs in the end user nodes. The SOAs provide the amplification needed in the metro ring, and double as modulators for the upstream data.

under direct modulation. As many as four SOA/modulators have been cascaded in the demonstration. The WDM experiments described earlier suggest that it is reasonable to expect that at least eight cascaded SOAs can be supported, and probably many more because of the lower OSNR requirement associated with the relatively low bitrate.

4. Summary and Outlook

The SOA is a rich device with many properties that are potentially useful in optical telecommunication. Used as an amplifier, its main properties are gain, saturation output power, and noise figure. Typical values for the gain are 10 dB to as high as 30 dB. Output power is typically +10 dBm, but a device with a P_{sat} of +17 dBm has been demonstrated. Noise figure is usually between 8 and 10 dB. A number of dynamic effects, such as cross-gain and cross-phase modulation and four-wave mixing, enable the operation of a SOA as an all-optical processing device. Interband carrier dynamics with lifetimes of the order of 100 ps allow operation at bitrates up to 10 Gb/s, but the stimulated lifetimes are much shorter, allowing considerably faster operation when the optical power in the SOA is high.

The many forms of all-optical processing possible with SOAs (wavelength conversion, optical time-division multiplexing, all-optical regeneration, optical phase conjugation) are what kept interest in these devices high over the years, and new applications, such as all-optical bitwise logic, are coming up all the time. The bitrates supported have been rising steadily. Doubling of the speed of cross-gain and cross-phase modulation with respect to the previous state of the art is regularly announced, and the 100 Gb/s mark has already been crossed, where new dynamic mechanisms like carrier heating and spectral hole burning become important. It looks unlikely that these mechanisms will support 200 or 400 Gb/s soon, but of course previous cautious predictions of that kind have been proven wrong time and again.

All the progress in signal processing applications does not mean that SOAs cannot be used simply as amplifiers. They can, as many single- and multichannel transmission experiments have demonstrated. Operation in the nearly-linear regime has even allowed adding and dropping of wavelengths, which is still a stumbling block in EDFA-amplified systems. Without a doubt, linearity is the holy grail when simple amplification is needed. Gain clamping gets us there a long way, but the relaxation oscillations in present-day devices still limit their usefulness.

Recent transmission experiments fit well with predictions of ultimate link capacity based on a SNR-limited system, as shown in Fig. 23. Presently, the state of the art is 320 Gb/s over 160 km. Improvements in device noise figure or output power immediately lead to a higher capacity-distance product, and this can be contributed to both by chip improvements and reduction of the

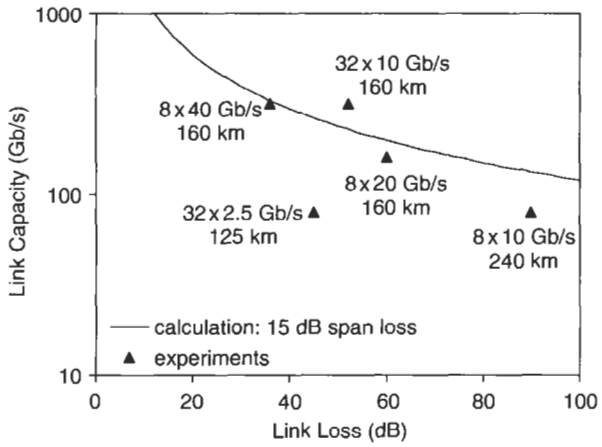


Fig. 23 Expected ultimate transmission distances and capacities using SOAs, assuming OSNR limited transmission. The points are the results discussed in this chapter.

fiber-chip coupling loss. Terabits/s over metro distances or hundreds of Gb/s over 1000 km may soon be within reach.

But undoubtedly, the most attractive feature of the SOA is that it is a compact device that can be fabricated efficiently in high volume using an IC production process. Compactness and low cost go well together to enable use of a component in many places in the network. When SOAs can deliver on these promises, and cheap gain and optical processing power become ubiquitous in optical networks, this will revolutionize the way these networks are designed and built.

References

G.P. Agrawal and N.K. Dutta, 1993. Semiconductor Lasers, 2nd ed., Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

M. Bachmann, P. Doussière, J.Y. Emery, R. N'Go, F. Pommereau, L. Goldstein, G. Soulage, and A. Jourdan, 1996. Polarisation-insensitive clamped-gain SOA with integrated spot-size converter and DBR gratings for WDM applications at 1.55 μm wavelength, Electron. Lett., 32 (22), pp.2076–2078.

B. Bauer, F. Henry, and R. Schimpe, 1994. Gain Stabilization of a Semiconductor Optical Amplifier by Distributed Feedback, IEEE Photon. Technol. Lett., 6 (2), pp.182–185.

T. Brenner and H. Melchior, 1993. Integrated Optical Modeshape Adapters in InGaAsP/InP for Efficient Fiber-to-Waveguide Coupling, IEEE Photon. Technol. Lett., 5 (9), pp.1053–1056.

J.W. Crowe and W.E. Ahearn, 1966. Semiconductor Laser Amplifier, IEEE J. Quantum Electron., 2 (8), pp.283–289.

S. Diez, R. Ludwig, and H.G. Weber, 1999. Gain-Transparent SOA-Switch for High-Bitrate OTDM Add/Drop Multiplexing, IEEE Photon. Technol. Lett., 11 (1), pp.60–62.

S. Diez, C. Schubert, R. Ludwig, H.J. Ehrke, U. Feiste, C. Schmidt, and H.G. Weber, 2000. 160 Gbit/s all-optical demultiplexer using hybrid gain-transparent SOA Mach-Zehnder interferometer, Electron. Lett., 36 (17), pp.1484–1486.

S.P. DiJaili, J.M. Wiesenfeld, G. Raybon, C.A. Burrus, A. Dienes, J.S. Smith, and J.R. Whinnery, 1992. Cross-Phase Modulation in a Semiconductor Laser Amplifier Determined by a Dispersive Technique, IEEE J. Quantum Electron., 28 (1), pp.141–150.

P. Doussière, P. Garabedian, C. Graver, D. Bonnevie, T. Fillion, E. Derouin, M. Monnot, J.G. Provost, D. Leclerc, and M. Klenk, 1994. 1.55 μm Polarisation Independent Semiconductor Optical Amplifier with 25 dB Fiber to Fiber Gain, IEEE Photon. Technol. Lett., 6 (2), pp.170–172.

T. Durhuus, C. Joergensen, B. Mikkelsen, R.J.S. Pedersen, and K.E. Stubkjaer, 1994. All Optical Wavelength Conversion by SOA's in a Mach-Zehnder Configuration, IEEE Photon. Technol. Lett., 6 (1), pp.53–55.

M. Eiselt, 1992. Optical loop mirror with semiconductor laser amplifier, Electron. Lett., 28 (16), pp.1505–1507.

M. Eiselt, W. Pieper, and H.G. Weber, 1993a. Decision gate for all-optical data retiming using a semiconductor laser amplifier in a loop mirror configuration, Electron. Lett., 29 (1), pp.107–109.

M. Eiselt, W. Pieper, and H.G. Weber, 1993b. All-optical high speed demultiplexing with a semiconductor laser amplifier in a loop mirror configuration, Electron. Lett., 29 (13), pp.1167–1168.

A.D. Ellis and D.M. Spirit, 1993. Compact 40 Gbit/s optical demultiplexer using a GaInAsP optical amplifier, Electron. Lett., 29 (24), pp.2115–2116.

J.Y. Emery, T. Ducellier, M. Bachmann, P. Doussière, F. Pommereau, R. Ngo, F. Gaborit, L. Goldstein, G. Laube, and J. Barrau, 1997. High performance 1.55 μm polarisation-insensitive semiconductor optical amplifier based on low-tensile-strained bulk GaInAsP, Electron. Lett., 33 (12), pp.1083–1084.

U. Feiste, R. Ludwig, C. Schmidt, E. Dietrich, S. Diez, H.J. Ehrke, E. Patzak, H.G. Weber, and T. Merker, 1999. 80-Gb/s Transmission over 106-km Standard-Fiber Using Optical Phase Conjugation in a Sagnac-Interferometer, IEEE Photon. Technol. Lett., 11 (8), pp.1063–1065.

D.A. Francis, S.P. DiJaili, and J.D. Walker, 2001. A single-chip linear optical amplifier, Optical Fiber Communication Conference—OFC 2001, Anaheim, California, 17–22 March 2001, Post-deadline papers, paper PD13.

B. Glance, J.M. Wiesenfeld, U. Koren, A.H. Gnauck, H.M. Presby, and A. Jourdan, 1992. High performance optical wavelength shifter, Electron. Lett., 28 (18), pp.1714–1715.

A.H. Gnauck, L.H. Spiekman, J.M. Wiesenfeld, and L.D. Garrett, 2000. Dynamic add/drop of 8-of-16 10-Gb/s channels in 4 × 40 km semiconductor-optical-amplifier-based WDM system, Optical Fiber Communication Conference—OFC 2000,

Baltimore, Maryland, 5–10 March 2000, Post-deadline papers, paper PD39, pp.284–286.

R. Hess, M. Caraccia-Gross, W. Vogt, E. Gamper, P.A. Besse, M. Duelk, E. Gini, H. Melchior, B. Mikkelsen, M. Vaa, K.S. Jepsen, K.E. Stubkjaer, and S. Bouchoule, 1998. All-Optical Demultiplexing of 80 to 10 Gb/s Signals with Monolithic Integrated High-Performance Mach–Zehnder Interferometer, IEEE Photon. Technol. Lett., 10 (1), pp.165–167.

R. Hess, M. Duelk, W. Vogt, E. Gamper, E. Gini, P.A. Besse, H. Melchior, K.S. Jepsen, B. Mikkelsen, M. Vaa, H.N. Poulsen, A.T. Clausen, K.E. Stubkjaer, S. Bouchoule, and F. Devaux, 1998. Simultaneous all-optical add and drop multiplexing of 40 Gbit/s OTDM signals using monolithically integrated Mach-Zehnder interferometer, Electron. Lett., 34 (6), pp.579–580.

G. Hunziker, R. Paiella, D.F. Geraghty, K.J. Vahala, and U. Koren, 1996. Polarization-Independent Wavelength Conversion at 2.5 Gb/s by Dual-Pump Four-Wave Mixing in a Strained Semiconductor Optical Amplifier, IEEE Photon. Technol. Lett., 8 (12), pp.1633–1635.

P.P. Iannone, K.C. Reichmann, A. Smiljanic, N.J. Frigo, A.H. Gnauck, L.H. Spiekman, and R.M. Derosier, 2000. A Transparent WDM Network Featuring Shared Virtual Rings, J. Lightwave Technol., 18 (12), pp.1955–1963.

W. Idler, M. Schilling, K. Daub, D. Baums, U. Körner, E. Lach, G. Laube, and K. Wünstel, 1995. Signal quality and BER performance improvement by wavelength conversion with an integrated three-port Mach-Zehnder interferometer, Electron. Lett., 31 (6), pp.454–455.

K. Inoue, T. Mukai, and T. Saitoh, 1987a. Gain saturation dependence on signal wavelength in a travelling-wave semiconductor laser amplifier, Electron. Lett., 23 (7), pp.328–329.

K. Inoue, T. Mukai, and T. Saitoh, 1987b. Nearly degenerate four-wave mixing in a traveling-wave semiconductor laser amplifier, Appl. Phys. Lett., 51 (14), pp.1051–1053.

K. Inoue, 1989. Crosstalk and Its Power Penalty in Multichannel Transmission due to Gain Saturation in a Semiconductor Laser Amplifier, J. Lightwave Technol., 7 (7), pp.1118–1124.

T. Ito, N. Yoshimoto, K. Magari, and H. Sugiura, 1998. Wide-Band Polarization-Independent Tensile-Strained InGaAs MQW-SOA Gate, IEEE Photon. Technol. Lett., 10 (5), pp.657–659.

E. Jahn, N. Agrawal, W. Pieper, H.J. Ehrke, D. Franke, W. Fürst, and C.M. Weinert, 1996. Monolithically integrated nonlinear Sagnac interferometer and its application as a 20 Gbit/s all-optical demultiplexer, Electron. Lett., 32 (9), pp.782–784.

C. Janz, B. Dagens, A. Bisson, F. Poingt, F. Pommereau, F. Gaborit, I. Guillemot, and M. Renaud, 1999a. Integrated all-active Mach-Zehnder wavelength converter with −10 dBm signal sensitivity and 15 dB dynamic range at 10 Gbit/s, Electron. Lett., 35 (7), pp.588–590.

C. Janz, F. Poingt, F. Pommereau, W. Grieshaber, F. Gaborit, D. Leclerc, I. Guillemot, and M. Renaud, 1999b. All-active dual-order mode (DOMO) Mach-Zehnder wavelength converter for 10 Gbit/s operation, Electron. Lett., 35 (21), pp.1862–1863.

K.S. Jepsen, A. Buxens, A.T. Clausen, H.N. Poulsen, B. Mikkelsen, and K.E. Stubkjaer, 1998. 20 Gbit/s optical 3R regeneration using polarisation-independent monolithically integrated Michelson interferometer, Electron. Lett., 34 (5), pp.472–474.

C. Joergensen, T. Durhuus, C. Braagaard, B. Mikkelsen, and K.E. Stubkjaer, 1993. 4 Gb/s Optical Wavelength Conversion using Semiconductor Optical Amplifiers, IEEE Photon. Technol. Lett., 5 (6), pp.657–660.

C. Joergensen, S.L. Danielsen, M. Vaa, B. Mikkelsen, K.E. Stubkjaer, P. Doussière, F. Pommerau, L. Goldstein, and M. Goix, 1996. 40 Gbit/s all-optical wavelength conversion by semiconductor optical amplifiers, Electron. Lett., 32 (4), pp.367–368.

C. Joergensen, S.L. Danielsen, T. Durhuus, B. Mikkelsen, K.E. Stubkjaer, N. Vodjdani, F. Ratovelomanana, A. Enard, G. Glastre, D. Rondi, and R. Blondeau, 1996. Wavelength Conversion by Optimized Monolithic Integrated Mach-Zehnder Interferometer, IEEE Photon. Technol. Lett., 8 (4), pp.521–523.

M. Joma, H. Horikawa, C.Q. Xu, K. Yamada, Y. Katoh, and T. Kamijoh, 1993. Polarization insensitive semiconductor laser amplifiers with tensile strained InGaAsP/InGaAsP multiple quantum well structure, Appl. Phys. Lett., 62 (2), pp.121–122.

R.M. Jopson and R.E. Tench, 1993. Polarisation-independent phase conjugation of lightwave signals, Electron. Lett., 29 (25), pp.2216–2217.

A.E. Kelly, I.F. Lealman, L.J. Rivers, S.D. Perrin, and M. Silver, 1996. Polarisation insensitive, 25 dB gain semiconductor laser amplifier without antireflection coatings, Electron. Lett., 32 (19), pp.1835–1836.

A.E. Kelly, A.D. Ellis, D. Nesset, R. Kashyap, and D.G. Moodie, 1998. 100 Gbit/s wavelength conversion using FWM in a MQW semiconductor optical amplifier, Electron. Lett., 34 (20), pp.1955–1956.

H.K. Kim and S. Chandrasekhar, 2000. Reduction of Cross-Gain Modulation in the Semiconductor Optical Amplifier by Using Wavelength Modulated Signal, IEEE Photon. Technol. Lett., 12 (10), pp.1412–1414.

S. Kitamura, H. Hatakeyama, K. Hamamoto, T. Sasaki, K. Komatsu, and M. Yamaguchi, 1999. Spot-Size Converter Integrated Semiconductor Optical Amplifiers for Optical Gate Applications, IEEE J. Quantum Electron., 35 (7), pp.1067–1074.

P.I. Kuindersma, G.P.J.M. Cuijpers, J.G.L. Jennen, J.J.E. Reid, L.F. Tiemeijer, H. de Waardt, and A.J. Boot, 1996. 10 Gbit/s RZ transmission at 1309 nm over 420 km using a chain of multiple quantum well semiconductor optical amplifier modules at 38 km intervals, 22nd European Conference on Optical Communication—ECOC '96, Oslo, Norway, 15–19 September 1996, Vol. 2, pp.165–168.

J.P.R. Lacey, S.J. Madden, and M.A. Summerfield, 1997. Four-Channel Polarization-Insensitive Optically Transparent Wavelength Converter, IEEE Photon. Technol. Lett., 9 (10), pp.1355–1357.

B. Lavigne, D. Chiaroni, L. Hamon, C. Janz, and A. Jourdan, 1998. Performance and system margins at 10 Gbit/s of an optical repeater for long haul NRZ transmission, 24th European Conference on Optical Communication—ECOC '98, Madrid, Spain, 20–24 September 1998, Vol. 1, pp.559–560.

J. Leuthold, P.A. Besse, E. Gamper, M. Dülk, S. Fischer, G. Guekos, and H. Melchior, 1999. All-Optical Mach-Zehnder Interferometer Wavelength Converters and Switches with Integrated Data- and Control-Signal Separation Scheme, J. Lightwave Technol., 17 (6), pp.1056–1066.

J. Leuthold, C.H. Joyner, B. Mikkelsen, G. Raybon, J.L. Pleumeekers, B.I. Miller, K. Dreyer, and C.A. Burrus, 2000. 100 Gbit/s all-optical wavelength conversion with integrated SOA delayed-interference configuration, Electron. Lett., 36 (13), pp.1129–1130.

L.Y. Lin, J.M. Wiesenfeld, J.S. Perino, A.H. Gnauck, 1998. Polarization-Insensitive Wavelength Conversion Up to 10 Gb/s Based on Four-Wave Mixing in a Semiconductor Optical Amplifier, IEEE Photon. Technol. Lett., 10 (7), pp.955–957.

R.J. Manning, D.A.O. Davies, D. Cotter, and J.K. Lucek, 1994. Enhanced recovery rates in semiconductor laser amplifiers using optical pumping, Electron. Lett., 30 (10), pp.787–788.

B. Mikkelsen, T. Durhuus, C. Joergensen, R.J.S. Pedersen, C. Braagaard, and K.E. Stubkjaer, 1994a. Polarisation insensitive wavelength conversion of 10 Gbit/s signals with SOAs in a Michelson interferometer, Electron. Lett., 30 (3), pp.260–261.

B. Mikkelsen, T. Durhuus, C. Joergensen, R.J.S. Pedersen, S.L. Danielsen, and K.E. Stubkjaer, 1994b. 10 Gbit/s wavelength converter realised by monolithic integration of semiconductor optical amplifiers and Michelson interferometer, 20th European Conference on Optical Communication—ECOC '94, Firenze, Italy, 25–29 September 1994, Vol. 4 (post-deadline papers), pp.67–70.

B. Mikkelsen, S.L. Danielsen, C. Joergensen, R.J.S. Pedersen, H.N. Poulsen, and K.E. Stubkjaer, 1996. All-optical noise reduction capability of interferometric wavelength converters, Electron. Lett., 32 (6), pp.566–567.

B. Mikkelsen, K.S. Jepsen, M. Vaa, H.N. Poulsen, K.E. Stubkjaer, R. Hess, M. Duelk, W. Vogt, E. Gamper, E. Gini, P.A. Besse, H. Melchior, S. Bouchoule, and F. Deveaux, 1997. All-optical wavelength converter scheme for high speed RZ signal formats, Electron. Lett., 33 (25), pp.2137–2139.

T. Morioka, H. Takara, S. Kawanishi, K. Uchiyama, and M. Saruwatari, 1996. Polarisation-independent all-optical demultiplexing up to 200 Gbit/s using four-wave mixing in a semiconductor laser amplifier, Electron. Lett., 32 (9), pp.840–842.

K. Morito, M. Ekawa, T. Watanabe, T. Fujii, and Y. Kotaki, 2000. High saturation output power (+17 dBm) 1550 nm polarization insensitive semiconductor optical amplifier, 26th European Conference on Optical Communication—ECOC 2000, Munich, Germany, 3–7 September 2000, pp.39–41.

T. Mukai and Y. Yamamoto, 1981. Gain, Frequency Bandwidth, and Saturation Output Power of AlGaAs DH Laser Amplifiers, IEEE J. Quantum Electron., 17 (6), pp.1028–1034.

V.G. Mutalik, G. van den Hoven, and L. Tiemeijer, 1997. Analog performance of 1310-nm gain-clamped semiconductor optical amplifiers, Optical Fiber Communication Conference—OFC '97, Dallas, Texas, 16–21 February 1997, pp.266–267.

M.A. Newkirk, B.I. Miller, U. Koren, M.G. Young, M. Chien, R.M. Jopson, and C.A. Burrus, 1993. 1.5 μm Multiquantum-Well Semiconductor Optical Amplifier with Tensile and Compressively Strained Wells for Polarization-Independent Gain, IEEE Photon. Technol. Lett., 4 (4), pp.406–408.

N.A. Olsson, 1989. Lightwave Systems With Optical Amplifiers, J. Lightwave Technol., 7 (7), pp.1071–1082.

X. Pan, J.M. Wiesenfeld, J.S. Perino, T.L. Koch, G. Raybon, U. Koren, M. Chien, M. Young, B.I. Miller, and C.A. Burrus, 1995. Dynamic Operation of a Three-Port, Integrated Mach-Zehnder Wavelength Converter, IEEE Photon. Technol. Lett., 7 (9), pp.995–997.

N.S. Patel, K.A. Rauschenbach, and K.L. Hall, 1996. 40 Gb/s Demultiplexing Using an Ultrafast Nonlinear Interferometer (UNI), IEEE Photon. Technol. Lett., 8 (12), pp.1695–1697.

J.S. Perino, J.M. Wiesenfeld, and B. Glance, 1994. Fiber transmission of 10 Gbit/s signals following wavelength conversion using a travelling-wave semiconductor optical amplifier, Electron. Lett., 30 (3), 256–258.

J.L. Pleumeekers, T. Hessler, S. Haacke, M.A. Dupertuis, P.E. Selbmann, R.A. Taylor, B. Deveaud, T. Ducellier, M. Bachmann, and J.Y. Emery, 1997. Relaxation Oscillations in the Gain Recovery of Gain-Clamped Semiconductor Optical Amplifiers: Simulation and Experiments, Optical Amplifiers and their Applications—OAA 1997, Victoria, Canada, 21–23 July 1997, pp.224–227.

F. Ratovelomanana, N. Vodjdani, A. Enard, G. Glastre, D. Rondi, R. Blondeau, C. Joergensen, T. Durhuus, B. Mikkelsen, K.E. Stubkjaer, A. Jourdan, and G. Soulage, 1995. An All-Optical Wavelength Converter with Semiconductor Optical Amplifiers Monolithically Integrated in an Asymmetric Passive Mach-Zehnder Interferometer, IEEE Photon. Technol. Lett., 7 (10), pp.992–994.

J.J.E. Reid, L. Cucala, M. Settembre, R.C.J. Smets, M. Ferreira, and H.F. Haunstein, 1998. An international field trial at 1.3 μm using an 800 km cascade of semiconductor optical amplifiers, 24th European Conference on Optical Communication—ECOC '98, Madrid, Spain, 20–24 September 1998, Vol. 1, pp.567–568.

A.A.M. Saleh and I.M.I. Habbab, 1990. Effects of semiconductor-optical-amplifier nonlinearity on the performance of high-speed intensity-modulation lightwave systems, IEEE Trans. Commun., 38 (6), pp.839–846.

R. Schnabel, U. Hilbk, T. Hermes, P. Meissner, C. Helmolt, K. Magari, F. Raub, W. Pieper, F.J. Westphal, R. Ludwig, L. Küller, and H.G. Weber, 1994. Polarization-Insensitive Frequency Conversion of a 10-Channel OFDM Signal Using Four-Wave-Mixing in a Semiconductor Laser Amplifier, IEEE Photon. Technol. Lett., 6 (1), pp.56–58.

P.E. Selbmann, T.P. Hessler, J.L. Pleumeekers, M.A. Dupertuis, B. Deveaud, B. Dagens, and J.Y. Emery, 1999. Observation of dark-pulse formation in gain-clamped semiconductor optical amplifiers by cross-gain modulation, Appl. Phys. Lett., 75 (24), pp.3760–3762.

J.C. Simon, P. Doussière, P. Lamoulier, I. Valiente, and F. Riou, 1994. Travelling wave semiconductor optical amplifier with reduced nonlinear distortions, Electron. Lett., 30 (1), pp.49–50.

L.H. Spiekman, J.M. Wiesenfeld, U. Koren, B.I. Miller, and M.D. Chien, 1998. All-Optical Mach-Zehnder Wavelength Converter with Monolithically Integrated Preamplifiers, IEEE Photon. Technol. Lett., 10 (8), pp.1115–1117.

L.H. Spiekman, J.M. Wiesenfeld, A.H. Gnauck, L.D. Garrett, G.N. van den Hoven, T. van Dongen, M.J.H. Sander-Jochem, and J.J.M. Binsma, 2000a. Transmission of 8 DWDM Channels at 20 Gb/s over 160 km of Standard Fiber Using a Cascade of Semiconductor Optical Amplifiers, IEEE Photon. Technol. Lett., 12 (6), pp.717–719.

L.H. Spiekman, A.H. Gnauck, J.M. Wiesenfeld, and L.D. Garrett, 2000b. DWDM transmission of thirty two 10 Gbit/s channels through 160 km link using semiconductor optical amplifiers, Electron. Lett., 36 (12), pp.1046–1047.

L.H. Spiekman, G.N. van den Hoven, T. van Dongen, M.J.H. Sander-Jochem, J.H.H.M. Kemperman, J.J.M. Binsma, J.M. Wiesenfeld, A.H. Gnauck, and L.D. Garrett, 2000c. Recent advances in WDM applications of semiconductor optical amplifiers, 26th European Conference on Optical Communication—ECOC 2000, Munich, Germany, 3–7 September 2000, pp.35–38.

L.H. Spiekman, J.M. Wiesenfeld, A.H. Gnauck, L.D. Garrett, G.N. van den Hoven, T. van Dongen, M.J.H. Sander-Jochem, and J.J.M. Binsma, 2000d. 8×10 Gb/s DWDM Transmission over 240 km of Standard Fiber Using a Cascade of Semiconductor Optical Amplifiers, IEEE Photon. Technol. Lett., 12 (8), pp.1082–1084.

A.K. Srivastava, Y. Sun, J.L. Zyskind, and J.W. Sulhoff, 1997. EDFA Transient Response to Channel Loss in WDM Transmission System, IEEE Photon. Technol. Lett., 9 (3), pp.386–388.

A.K. Srivastava, S. Banerjee, B.R. Eichenbaum, C. Wolf, Y. Sun, J.W. Sulhoff, and A.R. Chraplyvy, 2000. A Polarization Multiplexing Technique to Mitigate WDM Crosstalk in SOAs, IEEE Photon. Technol. Lett., 12 (10), pp.1415–1416.

M.A. Summerfield and R.S. Tucker, 1996. Optimization of Pump and Signal Powers for Wavelength Converters based on FWM in Semiconductor Optical Amplifiers, IEEE Photon. Technol. Lett., 8 (10), pp.1316–1318.

Y. Sun, A.K. Srivastava, S. Banerjee, J.W. Sulhoff, R. Pan, K. Kantor, R.M. Jopson, and A.R. Chraplyvy, 1999. Error-free transmission of 32×2.5 Gbit/s DWDM channels over 125 km using cascaded in-line semiconductor optical amplifiers, Electron. Lett., 35 (21), pp.1863–1865.

M.C. Tatham, G. Sherlock, and L.D. Westbrook, 1993. Compensation fiber chromatic dispersion by optical phase conjugation in a semiconductor laser amplifier, Electron. Lett., 29 (21), pp.1851–1852.

M.C. Tatham, X. Gu, L.D. Westbrook, G. Sherlock, and D.M. Spirit, 1994. Transmission of 10 Gbit/s directly modulated DFB signals over 200 km standard fiber using mid-span spectral inversion, Electron. Lett., 30 (16), pp.1335–1336.

P.J.A. Thijs, L.F. Tiemeijer, P.I. Kuindersma, J.J.M. Binsma, and T. van Dongen, 1991. High-Performance 1.5 μm Wavelength InGaAs-InGaAsP Strained Quantum Well Lasers and Amplifiers, IEEE J. Quantum Electron., 27 (6), pp.1426–1439.

L.F. Tiemeijer, P.J.A. Thijs, T. van Dongen, R.W.M. Slootweg, J.M.M. van der Heijden, J.J.M. Binsma, and M.P.C.M. Krijn, 1993. Polarization insensitive multiple quantum well laser amplifiers for the 1300 nm window, Appl. Phys. Lett., 62 (8), pp.826–828.

L.F. Tiemeijer, P.J.A. Thijs, T. van Dongen, J.J.M. Binsma, and E.J. Jansen, 1996a. Polarization Resolved, Complete Characterization of 1310 nm Fiber Pigtailed Multiple-Quantum-Well Optical Amplifiers, J. Lightwave Technol., 14 (6), pp.1524–1533.

L.F. Tiemeijer, P.J.A. Thijs, T. van Dongen, J.J.M. Binsma, and E.J. Jansen, 1996b. Noise Figure of Saturated 1310-nm Polarization Insensitive Multiple-Quantum-Well Optical Amplifiers, IEEE Photon. Technol. Lett., 8 (7), pp.873–875.

L.F. Tiemeijer, G.N. van den Hoven, P.J.A. Thijs, T. van Dongen, J.J.M. Binsma, and E.J. Jansen, 1996c. 1310-nm DBR-Type MQW Gain-Clamped Semiconductor Optical Amplifiers with AM-CATV-Grade Linearity, IEEE Photon. Technol. Lett., 8 (11), pp.1453–1455.

G.R. Walker, R.C. Steele, and N.G. Walker, 1989. Measurement of semiconductor laser amplifier noise figure in coherent optical transmission systems, Electron. Lett., 25 (25), pp.1681–1682.

J.M. Wiesenfeld and B. Glance, 1992. Cascadability and Fanout of Semiconductor Optical Amplifier Wavelength Shifter, IEEE Photon. Technol. Lett., 4 (10), pp.1168–1171.

J.M. Wiesenfeld, B. Glance, J.S. Perino, and A.H. Gnauck, 1993. Wavelength Conversion at 10 Gb/s using a Semiconductor Optical Amplifier, IEEE Photon. Technol. Lett., 5 (11), pp.1300–1303.

J.M. Wiesenfeld, 1996. Gain dynamics and associated nonlinearities in semiconductor optical amplifiers, Int. J. High-Speed Electronics and Systems, 7 (1), pp.179–222.

D. Wolfson, A. Kloch, T. Fjelde, C. Janz, B. Dagens, and M. Renaud, 2000. 40-Gb/s All-Optical Wavelength Conversion, Regeneration, and Demultiplexing in a SOA-Based All-Active Mach–Zehnder Interferometer, IEEE Photon. Technol. Lett., 12 (3), pp.332–334.

J. Yu, X. Zheng, and P. Jeppesen, 2000. Experimental demonstration of Cascaded Semiconductor Amplifier Based Gates Using Polarization Multiplexing Technique, Optical Amplifiers and their Applications—OAA 2000, Québec, Canada, 9–12 July 2000, pp.117–119.

J. Zhou, N. Park, K.J. Vahala, M.A. Newkirk, and B.I. Miller, 1994. Four-Wave Mixing Wavelength Conversion Efficiency in Semiconductor Traveling-Wave Amplifiers Measured to 65 nm of Wavelength Shift, IEEE Photon. Technol. Lett., 6 (8), pp.984–987.

Chapter 15 | All-Optical Regeneration: Principles and WDM Implementation

Olivier Leclerc, Bruno Lavigne, and Dominique Chiaroni

Alcatel Research & Innovation, Marcoussis, France

Emmanuel Desurvire

Alcatel Technical Academy, Nozay, France

The breakthrough of optical fiber amplification in the late 1980s combined with the techniques of wavelength-division multiplexing (WDM) and dispersion management in the 1990s have made possible to exploit a sizable fraction of the fiber bandwidth (several terahertz). Systems with error-free capacities of several terabit/s over hundreds of kilometers have reached the commercial area. With the introduction of powerful error-correction coding, the same performance is now achieved in transoceanic systems.

Current optical communications technologies rest upon the principle of repeaterless optical amplification, a passive form of inline signal processing. This is in sharp contrast with the earlier technology phase, where signals were periodically passed through electronic repeaters. A key question is to what extent active inline signal processing, such as all-optical signal regeneration, could prove beneficial in future developments, both in terms of system performance and economic return. Such a question is irrelevant as long as the passive approach meets capacity (1–5 Tbit/s), spectral efficiency (0.5–1 bit/s/Hz) and granularity (10–40 Gbit/s) specifications and needs. While substantially greater capacities and spectral efficiencies are unlikely to be reached with the current on/off format and amplification windows, there is potential economic interest in reducing the number of wavelength channels by increasing the channel rate (160–1280 Gbit/s). However, such a development would require breakthroughs in practically all aspects of system design, from terminals to fiber transport and, most importantly, would require inline optical signal processing or inline regeneration. It is safe to say that if the economic returns of optical regeneration are not proven (inasmuch as those of optical amplification were not in the early 1980s), its potential for revolutionizing optical communications networks through new high-bandwidth functionalities remains very promising.

This chapter reviews current technology alternatives for optical regeneration, considering both theoretical/experimental performance and practical component implementation issues with emphasis on WDM applications. It is divided into four sections. The first section is a brief review of signal

OPTICAL FIBER TELECOMMUNICATIONS, VOLUME IVA

ISBN: 0-12-395172-0

degradation causes and the introduction of the 3R regeneration concept. The second section concerns regeneration techniques using nonlinear gates (such as based upon semiconductor optical amplifiers, or SOA, and saturable absorbers). The third section concerns regeneration techniques based on synchronous modulation. The last section discusses the issue of electronic vs. optical regeneration and the engineering/economic implications of the two approaches for high-capacity WDM systems.

I. Signal Degradation Causes and Restoration by 3R Regeneration

It is well known that system transmission limits stem from a combined effect of amplifier noise accumulation, fiber dispersion, fiber nonlinearity, and inter/intrachannel interactions. Regardless of the transmission formats (RZ, NRZ, or CRZ), these various impairments result in three main types of signal degradation: intensity noise, timing jitter, and pulse-envelope distorsion.

Intensity noise might be more accurately referred to as the uncertainty in the energy content of a given bit slot. As such, it is not only a peak power fluctuation in the mark or 1 symbol (as sometimes wrongly thought of), but the result of any perturbing cause, linear or nonlinear, that randomly modifies the bit energy and causes decision errors. Fiber chromatic dispersion coherently mixes the contents of adjacent bits, optical amplification causes beat noises with spontaneous emission, fiber nonlinearities introduce information-dependent power transfer between WDM channels, all resulting in irreversible bit-energy fluctuations. Timing jitter is the uncertainty in the pulse-mark arrival time, or a synchronization default with respect to the bit stream. The main causes for timing jitter are nonlinearities self-phase modulation (SPM), cross-phase modulation (XPM), polarization-mode dispersion (PMD), and for RZ formats, the Gordon-Haus and electrostriction effects. Finally, pulse distorsion can be viewed as an irreversible change in the pulse envelope which increases the probability of symbol detection error. A most obvious pulse distorsion effect is the fill-up of 0-symbol spaces by amplified spontaneous emission (ASE), thus reducing the on/off extinction ratio. Fiber nonlinearities (SPM, XPM, four-wave mixing (FWM), stimulated raman scattering (SRS), and PMD are the essential factors of pulse-envelope distorsion.

So far, the strategy for limiting the above impairments has consisted in improving the transmission format (e.g., CRZ vs. NRZ in submarine systems), reducing power levels, or increasing the fiber mode area. From the terminal side, the introduction of error-correcting codes (ECCs) has made possible high levels of received signal quality (BER $< 10^{-7}$), while allowing relatively substantial signal degradation through the transmission line. Since ECC has become generic in both point-to-point and networking system applications,

we shall not consider it as a substitute for actual signal regeneration, but only as a built-in complement.

Given system impairments from amplifier noise and fiber nonlinearities after some transmission distance, two solutions remain. The first consists in segmenting the system into independent trunks, with full electronic repeater/transceivers at interfaces (we shall refer to this as "electronic regeneration" in the rest of this chapter). The second solution, inline optical regeneration, is not an optical version of the first. As developed here, optical regeneration performs the same signal-restoring functions as the electronic approach, but with far reduced complexity and enhanced capabilities. In both electronic and optical regeneration, the three basic signal-processing functions are reamplifying, reshaping and retiming, hence the generic acronym 3R (Fig. 1). Thus optical amplification (such as with SRS or EDFAs) provides a mere 1R signal-processing function. When retiming is absent, one usually refers to the regenerator as a 2R device, with only reamplifying and reshaping capabilities. Full 3R regeneration with retiming capability thus requires clock extraction, which can be done either electronically or all-optically, as shown below. It should be noted that optical 3R techniques are not necessarily void of any electronic functions (e.g., when using electronic clock recovery and O/E modulation), but the main feature is that these electronic functions are narrowband (as opposed to broadband in the case of electronic regeneration).

There are various ways of defining optical regeneration. It can refer to pure optical/optical (O/O) signal processing (signal modulation through an optical clock), as opposed to optoelectronic (O/E) processing. The all-optical label usually refers to the case where the regenerator subcomponents are optically

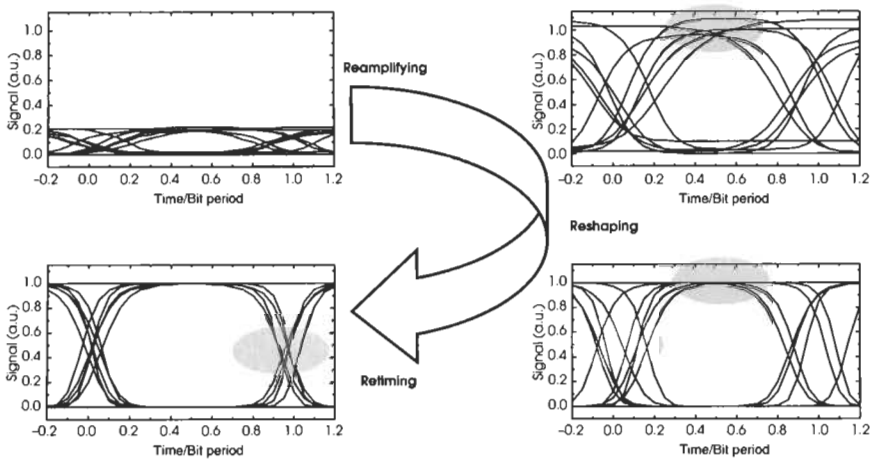


Fig. 1 Principle of 3R regeneration, as applied to NRZ signals (1) reamplifying; (2) reshaping; and (3) retiming.

controlled. This definition also applies when the signal information is transmitted to a physically different optical carrier in the process (including possible wavelength conversion).

II. Optical 3R Regeneration Using Nonlinear Gates

We first consider 3R regeneration based on optical nonlinear gates. A nonlinear gate can be defined as a three-port device where the second input port is used by a decision control, such as in a SOA gate. It can also be a two-port device, where the input signal acts as its own control, such as in the saturable absorber (SA) case.

Figure 2 shows the generic layout of a 3R regenerator based on nonlinear gates. The noisy signals to be regenerated are first launched into the reamplification block. Part of the output signal is extracted in order to recover an unjittered and synchronized clock signal, while the remainder is fed to the decision block, where reshaping and retiming functions are simultaneously performed. Clearly, the optimized design of such a regenerator requires the adaptation of optical device properties between all function block features. Under these conditions, the reamplification stage may not only represent a mere passive input power adaptation, but may also perform a reduction of input intensity fluctuations (as caused by propagation or crossing routing/switching nodes), so as to provide the regenerator with fixed input conditions. In this case, the input reamplification block acts as an adaptation interface, suppressing input fluctuations in power, wavelength, and polarization state, or adapting possible changes in data format. In practice, this results in the addition in the reamplification block of a control circuit (either optical or electrical) whose complexity directly depends on the actual system environment (ultrafast power equalization for packet switching applications and compensation of slow signal power fluctuations in transmission applications).

The clock-recovery block provides a low-jitter clock signal, as synchronized to the incoming data, in an optical form. This clock must be free of skew or random phase drift. Whatever recovery technique is used, the clock-pulse temporal width should match the decision window of the switching element. Hence, control of the optical clock pulse width is required for optimized operation of the second block. In this view, sinusoidal clocks are not necessarily suitable.

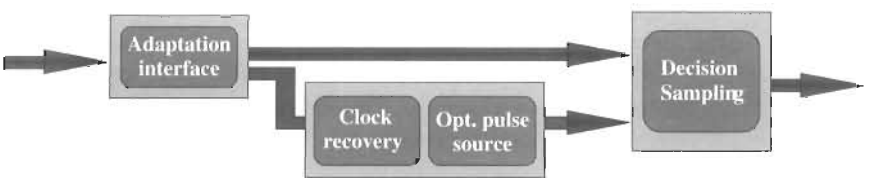


Fig. 2 Generic layout of 3R regenerator.

Clock recovery should also be adapted to system constraints. For instance, the introduction of optical packet switching techniques will impose new synchronization schemes with the appearance of asynchronism at the bit level.

The end block, i.e., the decision gate, performs reshaping and retiming. Ideally, it should also act as a transmitter capable of renewing or restoring optical pulse shapes. The chirp possibly introduced by the decision element in the resulting signal should carefully and adequately match line transmission requirements.

In the following subsections, we describe current solutions to realize the different blocks by means of all-optical techniques. The issue of the regenerator characterization is then addressed. In particular, we detail the impact of loop experiments to assess the effectiveness of the 3R regenerator to noise and timing jitter accumulation. Finally, we focus on a more exhaustive description of the implementation of a 3R regenerator using semiconductor optical amplifier (SOA) based devices.

A. BUILDING BLOCKS AND IMPLEMENTATION OF AN OPTICAL 3R REGENERATOR

1. Decision Block

The nonlinear response of the decision element is a relevant parameter for its regenerative properties. The theoretical evolution of bit-error rate (BER) with number of regenerator concatenation, for regenerators having different nonlinear responses, illustrates this feature in Fig. 3 [1]. The calculation however does not account for dynamic effects or jitter impact in the regenerator cascade. The nonlinearity response is characterized through the factor γ, which changes the shape of the nonlinearity from a step function ($\gamma = 0$) to a linear function ($\gamma = 1$). The noise level is adjusted so that the output BER is fixed to 10^{-12} after a single regenerator. As seen in the figure, the highest regeneration efficiency (as obtained when considering the less-penalizing concatenation of regenerators) is achieved with the ideal step function. In the physical domain, decision elements having such an ideal step function do not exist. Different nonlinear optical transfer functions approaching more or less the ideal response can be realized in various media: fiber, SOA, electroabsorption modulator (EAM), and laser.

A first class of nonlinear gates, extensively investigated and reported in the literature, concern optical interferometers. This type of regenerator uses either fiber-based or monolithic SOA-based devices. The principle of operation is the following. Consider an optical interferometer including one arm incorporating a nonlinear medium, as shown in Fig. 4a. Assuming the presence of two signals: an input signal carried by λ_1 wavelength acting as a pump signal and a local CW signal carried by λ_2 wavelength acting as a probe signal, we can express

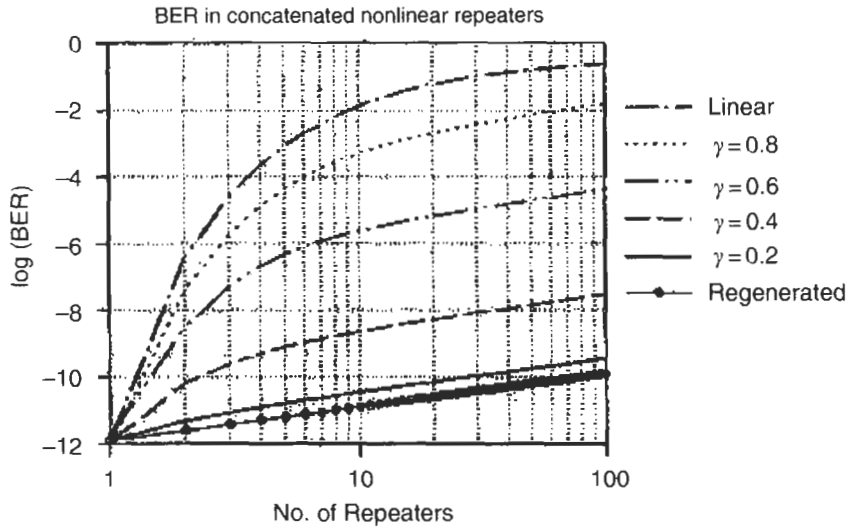


Fig. 3 Evolution of the BER with concatenated regenerators for different nonlinear transfer function (i.e., γ parameter).

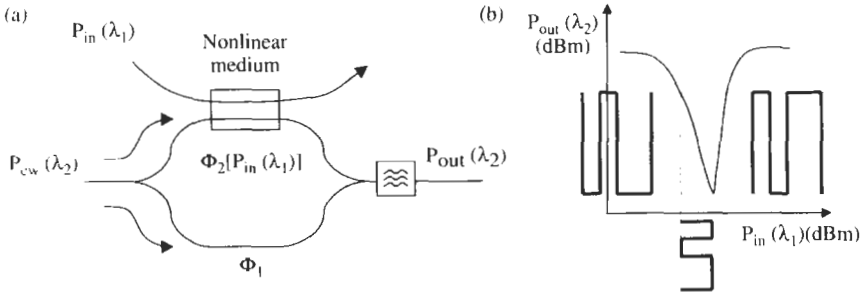


Fig. 4 (a) Schematic of an optical interferometer with a nonlinear medium on one arm; (b) nonlinear transfer function of the interferometer and application to incident coded signals.

the output power as follows:

$$P_{out}(\lambda_2) \propto P_{cw}(\lambda_2)[1 + \cos(\Delta\Phi(P_{in}(\lambda_1)))]$$

where $\Delta\Phi(P_{in}(\lambda_1)) = \Phi_2(P_{in}(\lambda_1)) - \Phi_1$. Injection of the signal at λ_1 induces a phase shift through XPM, the amount of which depends upon the power level $P_{in}(\lambda_1)$. In turn, the phase modulation induces amplitude modulation on the signal at λ_2. Figure 4b plots the nonlinear transfer function of the interferometer. The fast nonlinearity induces noise redistribution, resulting in a narrower distribution for marks and spaces and in an improvement of the

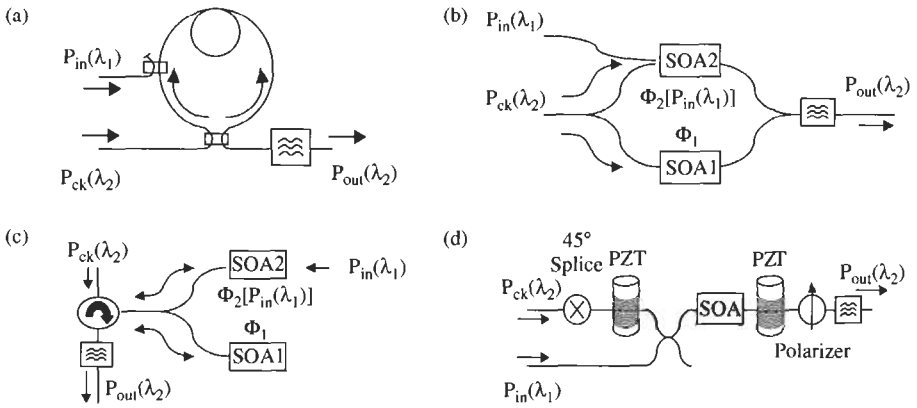


Fig. 5 Several kind of interferometers: (a) nonlinear loop mirror (NOLM), (b) SOA-based Mach-Zehnder interferometer (SOA-MZI), (c) SOA-based Michelson interferometer (SOA-MI), (d) SOA-based ultrafast nonlinear interferometer (UNI).

data extinction ratio. Under these conditions, the interferometer performs as a 2R regenerator. Depending on the selected function slope, data polarity is preserved (positive slope) or switched (negative slope). It should be noted that if a pulsed-clock signal is substituted for the CW signal at λ_2, reshaping and retiming are simultaneously performed by the sampling effect, resulting in full 3R regeneration.

The different types of interferometers can be classified according to the nature of the nonlinearity exploited to achieve a π phase shift.

In the case of fiber-based devices such as the nonlinear optical loop mirror (NOLM) (see Fig. 5a), the phase shift is induced through Kerr effect in an optical fiber. Dispersion-shifted fibers (DSF) are used to cancel the walkoff effect between the pump (λ_1) and probe (λ_2) signals. In absence of fiber loss and walkoff, and assuming near-complete polarization scrambling, the power level required to achieve a π phase shift is given by:

$$P_\pi = \frac{\lambda_1 A_{eff}}{(4n_2 L)}$$

where A_{eff} is the fiber effective area, n_2 the nonlinear Kerr index, and L the fiber length.

The key advantage of fiber-based devices such as NOLM lies in the near-instantaneous (fs) response of the Kerr nonlinearity, making them very attractive for ultrahigh bitrate operation ($\geq$160 Gbit/s). On the negative side, major limitations come from environmental instability, strong polarization dependence, and reduced integrability, mainly due to the requirement of relatively long fiber lengths. Polarization-insensitive NOLMs have been realized,

although with the same drawbacks concerning integrability. A comprehensive overview of NOLM possibilities for all-optical signal processing and 3R regeneration can be found in [2].

In the case of SOA-based interferometers, the phase shift is due to the effect of photoinduced carrier depletion in the gain saturation regime. It is related to phase and gain modulation through the linewidth enhancement factor α. As a result, no pure PM occurs in optically driven SOAs. In practice, a modulation depth of the gain near 5 dB results in a phase shift of π in such SOA-based interferometers. Such devices have been studied due to their high integration potential and open the possibility to integrate complex optical functions for optical signal processing.

Two basic structures have been explored to realize monolithic SOA-based interferometers, namely Mach-Zehnder and Michelson structures as shown in Fig. 5b and 5c, respectively. In Mach-Zehnder interferometers (MZI), cross-phase modulation is induced in the second SOA (labeled SOA2 in Fig. 5b). Referred SOA1 is used to balance the interferometer and also to select output pulse data polarity that can be either preserved or switched, depending on SOA1-induced static phase shift. Pump and probe signals can be launched in counter- or codirectional ways. In the first case, no optical filter is required at device output for rejecting the signal at λ_1 wavelength, but operation of the MZI is limited in speed. In Michelson interferometers (MI) (see Fig. 5c), the probe signal is coupled to the interferometer via a circulator and is back-reflected in two SOAs. As in the previous case, only SOA2 induces cross-phase modulation. The MI type has higher speed potential compared to the MZI type, mainly due to the fact that back reflections double the interaction length.

Optical 3R regeneration based on SOA-MI has been demonstrated at 20 Gbit/s [3], while its operation at up to 40 Gbit/s is assessed in [4]. In the case of SOA-MZI devices, loop transmission results have been obtained at 10 Gbit/s [5, 6].

Remarkable results have also been achieved by means of a SOA-based ultrafast nonlinear interferometer (UNI) [7], whose structure is shown in Fig. 5d. At the UNI input, the probe signal is split into 2 orthogonally polarized pulses by launching it at 45° to the axes of a polarization-maintaining fiber (PMF). After passing through the fiber, the probe pulse pairs are launched into the SOA where they experience a phase shift induced by the pump signal. The pulse pairs recombine before passing through a polarizer but after crossing a second PMF fiber, inducing a reversed delay between them. Both PMF fibers are wound onto piezoelectric drums to precisely control their lengths and the induced delay between probe polarizations. Operation speed at up to 80 Gbit/s was recently demonstrated with such devices [8], but the clear assessment of the UNI regeneration capability was reported in [9] where 20 UNI-type regenerators have been cascaded at 40 Gbit/s.

At this point, one should mention that the photo-induced modulation effects in SOAs are intrinsically limited in speed by the gain recovery time,

which is a function of the carrier lifetime and the injection current. The ultimate speed of operation of such devices is still under investigation, but is generally admitted to be near 100 Gbit/s, though record operation speed of 100 Gbit/s has already been demonstrated through cross-gain modulation in [10]. In order to increase operating speed in SOA-based devices, another approach taking benefit of the interferometric feature of these devices has been investigated. The differential operation mode (DOM) allows artificial increase of the speed of a component by injecting the same signal on both arms of the interferometer while shifted in time. Figure 6 illustrates the principle of operation. Suppose that a pulse is injected into the first arm of a SOA-MZI. This pulse induces a phase variation that is suppressed after a time τ by the same pulse injected into the second arm. Note that the temporal width of the resulting output pulses is directly governed by the delay τ.

Using the DOM principle, optical 3R regeneration has been demonstrated at 20 Gbit/s [11] and at 40 Gbit/s using SOA-MZI [12]. Figure 7a shows the complete regenerator structure. The input signal is split and launched into

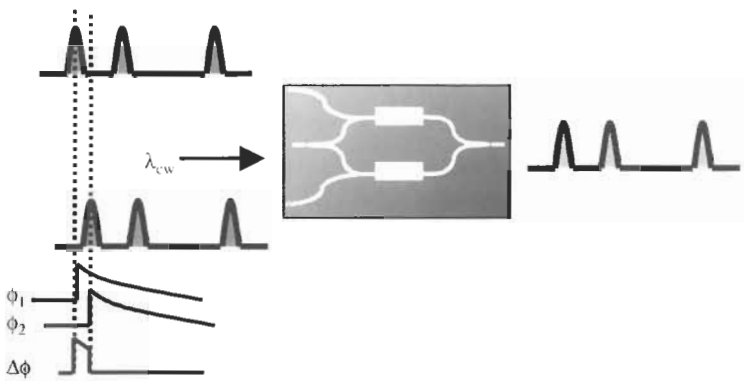


Fig. 6 Operation of a SOA-MZI in differential mode.

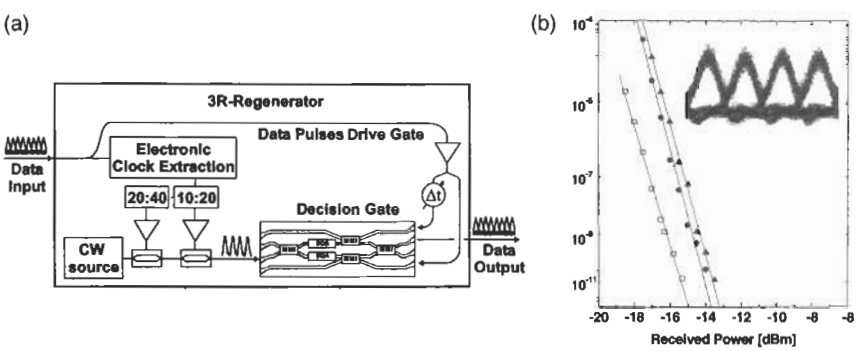


Fig. 7 Regeneration at 40 Gbit/s: (a) regenerator structure, (b) BER measurements.

the 2 arms of the interferometer. The delay between them is controlled by an optical delay line. The 40 GHz clock pulses propagate in a counterdirectional way to be encoded by the incoming data. Bit-error rate (BER) measurements are reported in Fig. 7b in the case where data and clock are carried out by the same wavelength (triangles) or carried out by different wavelengths (circles). In all cases, low penalties with respect to the baseline are observed. In the first case, a penalty of 1.8 dB is measured, while in the second case the penalty is reduced to 1.5 dB. A 40 Gbit/s eye diagram recorded at the output of the 3R regenerator is shown in the inset.

As seen earlier, nonlinear response is the key parameter that governs regeneration efficiency. The combination of two interferometers is a straightforward means to improve nonlinearity. This approach was studied and experimentally validated at 10 and 20 Gbit/s using a cascade of two SOA-MZI [13, 14], and at 20 Gbit/s using two NOLMs [15, 16]. Such a scheme also presents the advantage to restore both data polarity and wavelength of incoming data. Finally, the second conversion stage can be used as an adaptation interface to the transmission link achieved through chirp tuning in this second device as studied and reported in [17].

Another approach for performing optical decisions exploits the potential of laser-based responses. A first example is a Q-switched laser that has been demonstrated at 10 Gbit/s [18]. The device consists of three sections: a lasing DFB section, a passive phase tuning, and a second lasing section pumped at transparency and used as a reflector section. The combination of passive phase tuning and reflector sections allows the control of the back-reflected signal in amplitude and phase. A second example is a DFB laser integrated with a SOA [19]. In this case, lasing is turned off due to the gain saturation induced by the injection of a high-power signal. Used as a wavelength converter, operation at 40 Gbit/s has been shown [20]. Three-R regeneration experiments have been done at 2.5–10 Gbit/s in both cases with direct modulation of the lasing section by an electrical clock, but similar operation could be achieved in principle by simultaneously injecting the signal data and the optical clock. A last approach is based on the cross-absorption/modulation effect in electroabsorption modulators [21], but such a solution clearly suffers from high input power level requirements.

2. Clock Recovery Block

Converging toward fully integrated, all-optical 3R regenerators requires practical and sturdy optical clock recovery blocks. In general, recovering a clock signal in an incoming data stream at bitrate F consists in the extraction of a low-jitter signal at frequency F, the shape of which is not necessarily purely sinusoidal. All clock recovery techniques (electrical and all-optical) are in principle equivalent to filter the tone at F frequency or harmonics.

Restricting our view on monolithic all-optical devices, we describe two emerging realizations.

The first implementation uses the self-pulsating effect in DFB laser (SPL). Major achievements have been obtained with a three-section active device (two active DFB sections and one passive tuning section). The effect is chirping of a rising pulse in the DFB section that modulates the effective reflectivity from the second DFB section and switches the laser on and off [22]. Results demonstrated by such a device are discussed in [23, 24]. These references highlight high bitrate potentials, broad dynamic ranges, broad frequency tuning (6 to 46 GHz), polarization insensitivity, and quite relatively short locking times (1 ns). This last feature makes these devices good candidates for operation in asynchronous packet regime. A complementary 40 Gbit/s loop experiment was also recently conducted for further assessing both the quality of the recovered optical clock signal and its robustness to long series of consecutive 0s in transmission system environment [25]. We should finally mention a loop demonstration of all-optical 10 Gbit/s 3R regeneration [26].

The second uses a mode-locked laser diode (MLLD). The MLLD is a three-section device composed of independent saturable absorber, gain, and DBR sections. The last section allows control of the effective length of the cavity through an electronic control of the absorption coefficient. With such a device, clock recovery at 40 GHz was recently demonstrated on RZ 40 Gbit/s data stream with very short pulses (3 ps) [27]. To assess the quality of the recovered clock, an all-optical 3R regenerator was tested at 20 Gbit/s, combining a MLLD for the clock recovery and a saturable absorber waveguide as the decision element [28]. This result will be detailed in the next section.

3. Input Adaptation Interface

As previously mentioned, the reamplifying block of the regenerator should not only perform signal amplification, but also provide the following blocks with fixed conditions. In the case of synchronous transmission applications, several options have been studied. The first one exploits the use of an electronic control for the compensation of the power fluctuations by piloting the gain of a SOA, as for the case of 2 SOA stages [29] or SOA-MZI [14]. The second uses only optical techniques as the cascade of two EA modulators [30]. In the case of packet switching applications where fast power equalization is required, the exploitation of saturation properties of SOA seems one of the most promising approaches as shown in [31, 32].

After considering the different building blocks of an optical 3R regenerator, we consider its actual regenerative properties. Let us consider once more the curves of Fig. 3 which show the BER evolution with concatenated regenerators. It is seen that regardless of the nonlinearity of the function (and even in the case of a linear response function), all curves standing for different regeneration efficiencies pass through a single point after one device. This means

it is clearly not possible to qualify or evaluate the regenerative capability of any device at its output. Indeed, no difference at all in BER measurements are obtained when considering either a 3R regenerator or a mere amplifier if only one element is considered. This comes from the initial overlap between noise distributions associated to mark and space symbols that cannot be suppressed but only minimized by a single decision element through threshold adjustment. From Fig. 3, one observes different behaviors depending on the nonlinear function shape. The best regeneration performance is obtained with an ideal step function ($\gamma = 0$). In that case, BER linearly degrades in the cascade. Conversely, when nonlinearity is reduced (increasing γ), both BER and noise accumulate until the concatenation of nonlinear functions reaches a steady-state pattern, from which BER evolution becomes linear. In these conditions, concatenation of nonlinear devices clearly magnifies differences in nonlinear response shape and hence allows the actual qualification of their specific regeneration capabilities.

From the experimentation side, concatenation of regenerators is emulated by inserting a regenerator into recirculating loops, in which the optical signal propagates during a certain number of laps. After N laps, the loop is cleansed from data and noise, then a new set of data is injected. Loading and cleansing of the loop is performed using acoustooptic switches controlled by a delay generator. This delay generator also provides gating signals for monitoring the test apparatus (digital scope, optical spectrum, BER analyzer) and hence for evaluating transmission performance at a given distance. Figure 8 shows an example of such a loop test-bed.

As mentioned before, the key role of a nonlinear function is to change the noise characteristics, hence modifying the associated probability density

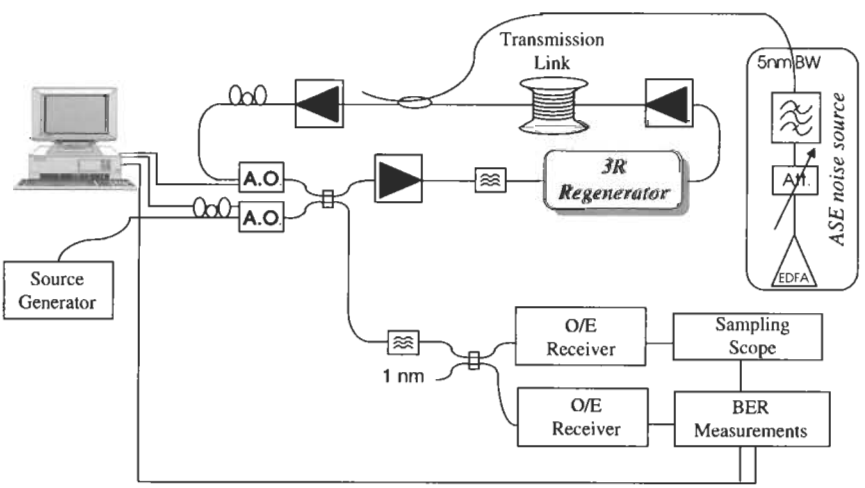


Fig. 8 Example of experimental loop test bed.

function (PDF), which then significantly differs from a Gaussian distribution. In this condition, the Q factor is no longer a valid parameter to qualify ultimate BERs and possible BER floors. How to measure the effect of regeneration then? Instead of measuring BER floors, which is quite difficult due to the long sampling times required in loop experiments, one might qualify regeneration performance through the evolution of the receiver sensitivity penalty at fixed BER with respect to key relevant parameters, such as optical signal-to-noise ratio (OSNR). Evaluation and comparison of regenerative properties then amounts to measuring the minimum OSNR ($OSNR_{min}$), corresponding for example to 1 dB penalty on the reception sensitivity with respect to the first lap, which is tolerated by the regenerative apparatus throughout the cascade. In such measurements, it should be noted that the OSNR is fixed so as to preserve a constant noise level injected into the regenerator whatever the lap. Since the technique is based on a measurement of relative penalties, it has the major advantage of casting off the contribution of the optoelectronic receiver used to carry out BER measurements from the overall system performance. This measurement then reflects the accumulated effects of all degrading factors occurring in the transmission, such as chromatic dispersion, timing jitter accumulation, and amplifier noise, and hence enables the extraction of the actual regenerative performance of the tested device.

B. IMPLEMENTATION OF SOA-MZI-BASED 3R REGENERATOR

We focus next on the design and the characterization of SOA-based optical 3R regenerators for 10 Gbit/s operation. After defining the optical decision element, we address the issue of complete structure optimization by taking into account the line transmission properties. In the last part, we address all-optical 3R regeneration through monolithic integrated devices which include both decision and clock recovery functions.

1. Decision Element Definition

Two structures are compared. The first one combines a SOA operating in a cross-gain regime with a SOA-MZI. We used the saturation properties of the first SOA to perform the power equalization, following the previously described principle. Since the SOA device also acts as a nonlinear element, it must be taken into account in the comparison of the overall nonlinear function. The second structure we have tested combines two cascaded SOA-MZIs. In this case, power equalization is achieved through low-speed control electronics that drives the current of the SOA preamplifier in the first SOA-MZI. For each structure, a single clock signal is used as the probe signal for performing the retiming function in the second SOA-MZI.

Figure 9a and 9b shows the nonlinear function measured in static and dynamic regimes, respectively. The dynamic measurements are obtained in a 10 Gbit/s loop experiment for both regenerator structures. In the last case,

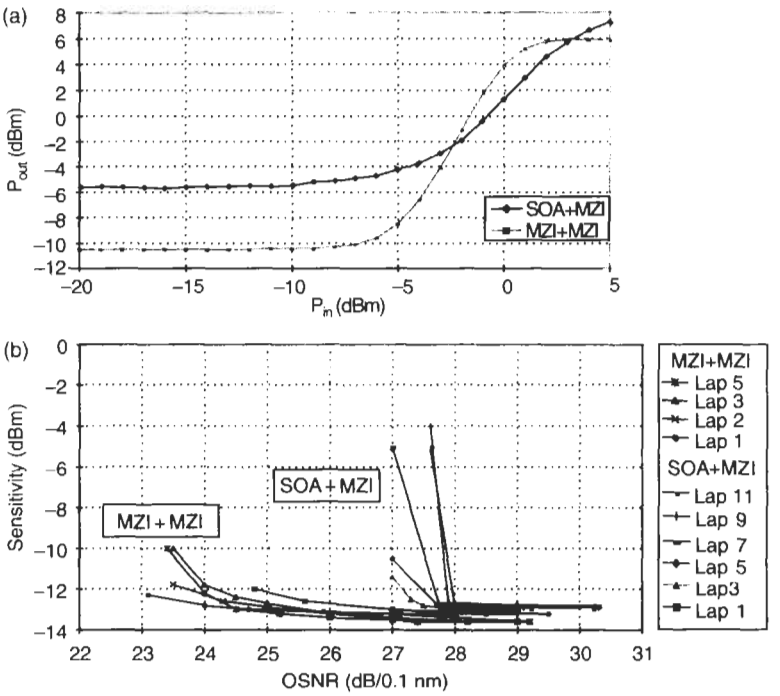


Fig. 9 (a) Nonlinear functions of 2 different regenerators recorded in static regime; (b) corresponding loop experiment results.

the recirculating loop incorporated a single 110 km span of dispersion-shifted fiber. As described before, loop evaluation of the regenerative properties of the two devices consists of the measurement of receiver sensitivity evolution (at fixed 10^{-10} BER) in the cascade given the OSNR. A comparison of the results obtained with the two structures shows a 4 dB gain in the minimum OSNR when combining two SOA-MZI ($OSNR_{min}$ = 24 dB/0.1 nm with NRZ pulses). Such a regenerator configuration exhibits a decision function with nonlinearity greater than in SOA/SOA-MZI devices, which further shows its importance in efficient optical 3R regeneration.

2. Transmission Properties

Another issue is the correct adaptation of the regenerated pulses to enable actual transmission of the renewed signal. In this view, spectral alteration (chirp) of regenerated signals induced by the SOA-MZI and directly depending upon its operation regime, as shown in Fig. 10 [33], plays a key role and should be optimized on the transmission line characteristics. Indeed, mismatch between regenerator-induced chirp characteristics and line dispersion causes significant pulse broadening and severe transmission penalties. In the general

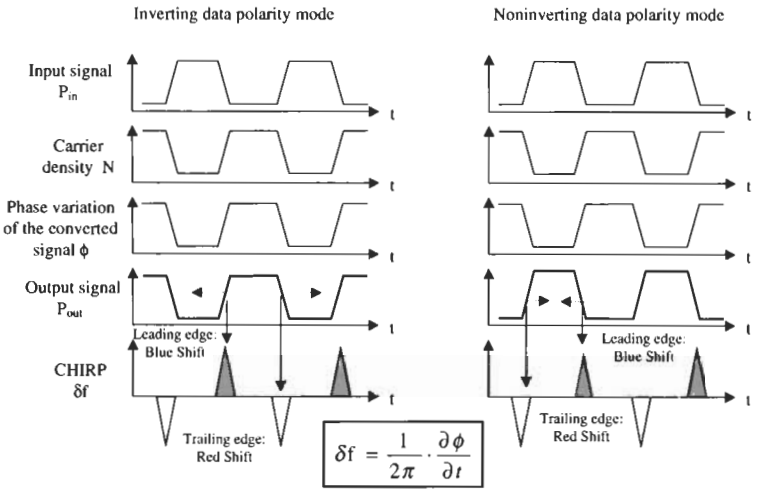


Fig. 10 Schematic showing carrier density, refractive index, and phase variation with respect to the input signal, and the converted signal with the associated chirp.

Table 1 **Comparative Study of Transmission Properties of a SOA-MZI used in Different Configurations. IPM and NIPM stand for inverting and noninverting polarity modes, respectively.**

Data Format	*Ref. Transmitter*	*MZI NIPM + CW*	*MZI IPM + CK*	*MZI NIPM + CK*
RZ	30 km 0.3 dB penalty @ 10^{-9}	70 km 1.7 dB penalty @10^{-9}	30 km 2 dB penalty @10^{-9}	30 km 2 dB penalty @10^{-9}

case of transmission over anomalous-dispersion fiber (e.g., SMF), regenerated data should then be redshifted at the leading edge and blueshifted at the trailing edge, as is the case with noninverting polarity mode (NIPM) operation of SOA-MZI.

This issue is experimentally investigated through a 10 Gbit/s loop experiment with SMF line fiber (dispersion of 17 $ps \cdot nm^{-1} \cdot km^{-1}$) using long pattern PRBS sequences of $2^{23} - 1$ [17]. Two different configurations of the 3R regenerator were considered depending on the type of probe input in the SOA-MZI (CW or clock signal). Results are summarized in Table 1.

When considering a CW probe, transmission is possible only when operating the SOA-MZI in NIPM mode as a (expected) consequence of the adequate matching of chirp and dispersion conditions. By contrast when using with a clock probe, performances obtained in both NIPM and IPM operation are identical. This results mainly from attenuated chirp effects due to

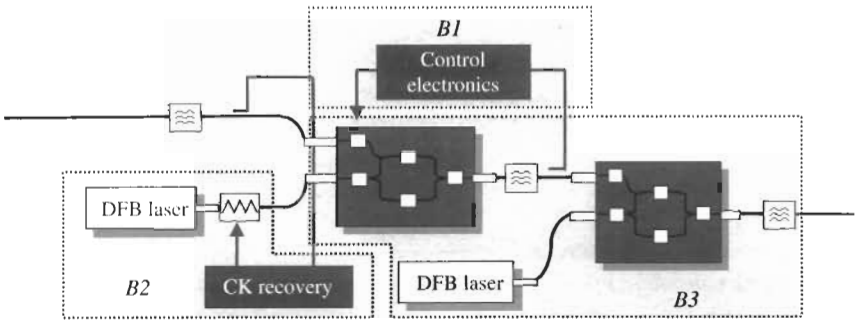


Fig. 11 Optimized structure of a SOA-based 3R regenerator.

data sampling with narrow pulses. When a clock signal is injected into the SOA-MZI, incoming data are sampled; and since clock pulses are narrower than data pulses, the chirp effects associated with the leading and trailing edges of incoming data are largely attenuated. In this case (which corresponds to an actual 3R regeneration operation regime), operation mode of the interferometer becomes less important.

Taking into account both the regeneration and transmission properties, we have implemented an optical 3R regenerator the structure of which is shown in Fig. 11.

This regenerator is composed of 3 blocks:

B1: adaptation interface: low-speed control electronics is used to suppress input power signal fluctuations by monitoring the current of the SOA preamplifier of the first stage;

B2: clock recovery: delivers low jitter and synchronized optical clock pulses;

B3: reshaping and retiming stage: reshaping is performed by cascading two SOA-MZIs. The retiming is achieved in the first stage, through the sampling of the incoming data by the clock pulses. The second stage is used to provide the output wavelength and to adapt the chirp of the regenerated data to the downstream transmission fiber.

Actual evaluation of the regeneration capabilities of the structure was performed through a 10 Gbit/s loop experiment incorporating 50 km of SMF fiber without any dispersion compensation. Figure 12 shows the evolution of receiver sensitivity (as measured at 10^{-10} BER) with the number of laps for different OSNR levels. These measurements were made with RZ pulses and $2^{23}-1$ long pseudorandom sequences. The minimum OSNR that can be tolerated by the 3R regenerator is measured to be 24 dB/0.1 nm. This corresponds to a 1 dB penalty (on the minimum OSNR) compared to the result obtained with DSF (see above). This penalty clearly originates from higher chromatic

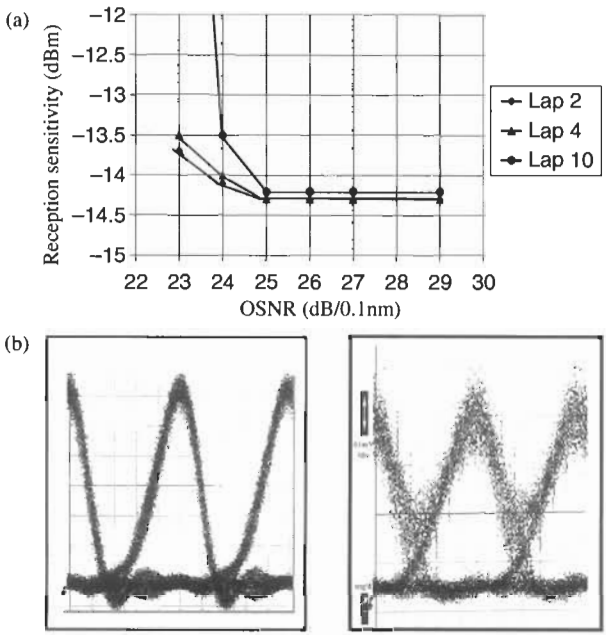


Fig. 12 (a) Evolution of reception sensitivity with OSNR for a different number of laps. (b) Eye diagrams recorded in a back-to-back configuration (left) and after the concatenation of 100 regenerators (right).

dispersion, as illustrated in Fig. 12, where 10 Gbit/s eye diagrams recorded at the input of the loop and after the cascade of 100 regenerators are displayed.

C. ALL-OPTICAL REGENERATION

In order to converge to a cost-effective solution, integration of the decision unit and the clock recovery through exclusively optical techniques is of key importance. A first step toward the realization of such an all-optical regenerator has been the demonstration of 3R regeneration with 10 Gbit/s RZ signals using the combination of two SOA-MZIs for the decision on one hand and a self-pulsating laser for the clock recovery on the other hand. The self-pulsating laser is a three-section InGaAsP/InP ridge waveguide laser operating at 1550 nm mentioned in the previous section with the following length characteristics (DFB laser, 300 μm; phase section 200 μm; dispersive reflector section, 150 μm).

Figure 13 shows the structure of the complete all-optical 3R regenerator. The first stage carries out a conversion on two carriers, λ_1 (1534 nm) and λ_2 (1538 nm). Part of the 1534 nm signal is extracted and launched into the self-pulsating laser (SP-DFB) for clock recovery, while the rest passes through a time-broadening stage for more efficient jitter suppression. The optically

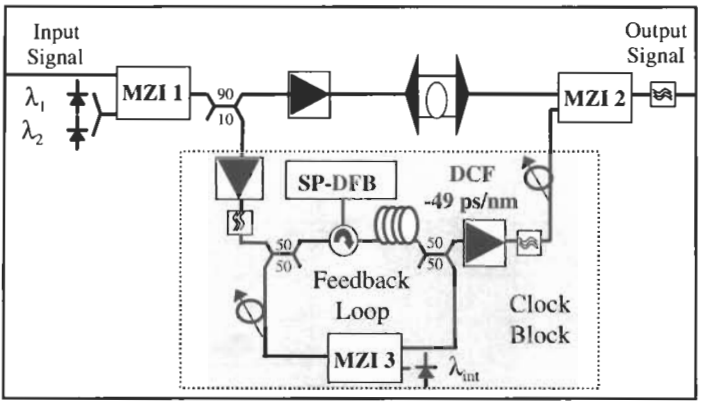


Fig. 13 Structure of the full all-optical regenerator.

recovered clock is launched into the second SOA-MZI conversion stage as a probe signal for performing both data sampling and jitter removal. As seen from Fig. 13, the optical clock recovery stage incorporates a feedback loop the role of which is to reinject a clock signal coming from the SP-DFB synchronized with the input RZ signal. Such a scheme is required for both stabilizing and improving of the clock quality mainly regarding rms jitter. Note that this loop comprises a wavelength converter since SP-DFB lasers cannot recover an optical clock on an input signal of the same wavelength as their own.

For validation purposes, the all-optical structure was inserted into a recirculation loop incorporating 50 km of DSF fiber. Up to 300 regenerators (then corresponding to a 15,000 km long transmission) were traversed without any measurable sensitivity penalty, except for the first lap as seen in Fig. 14a. Evolution of the recorded 10 Gbit/s eye diagram with number of laps is shown in Fig. 14b. No degradation is observed throughout the cascade, thus attesting the excellent regeneration capability of fully all-optical approaches.

D. OPTICAL REGENERATION BY SATURABLE ABSORBERS

In this section, we consider saturable absorbers (SA) as the nonlinear element thus enabling optical regeneration. Such devices have been extensively studied over the past few years for various applications, including passive laser mode-locking for ultrashort pulse generation [34, 35]. Figure 15 shows the typical transfer function of a SA and illustrates its principle of operation. When illuminated with an optical signal with instantaneous peak power below some threshold P_{sat}, the photonic absorption of the SA is high and the device is opaque to the signal (low transmittance). Above P_{sat}, the SA transmittance rapidly increases and saturates to an asymptotic value near unity (passive loss being overlooked). As previously mentioned, such a nonlinear transfer function of the SA makes it possible to use it for 2R optical regeneration.

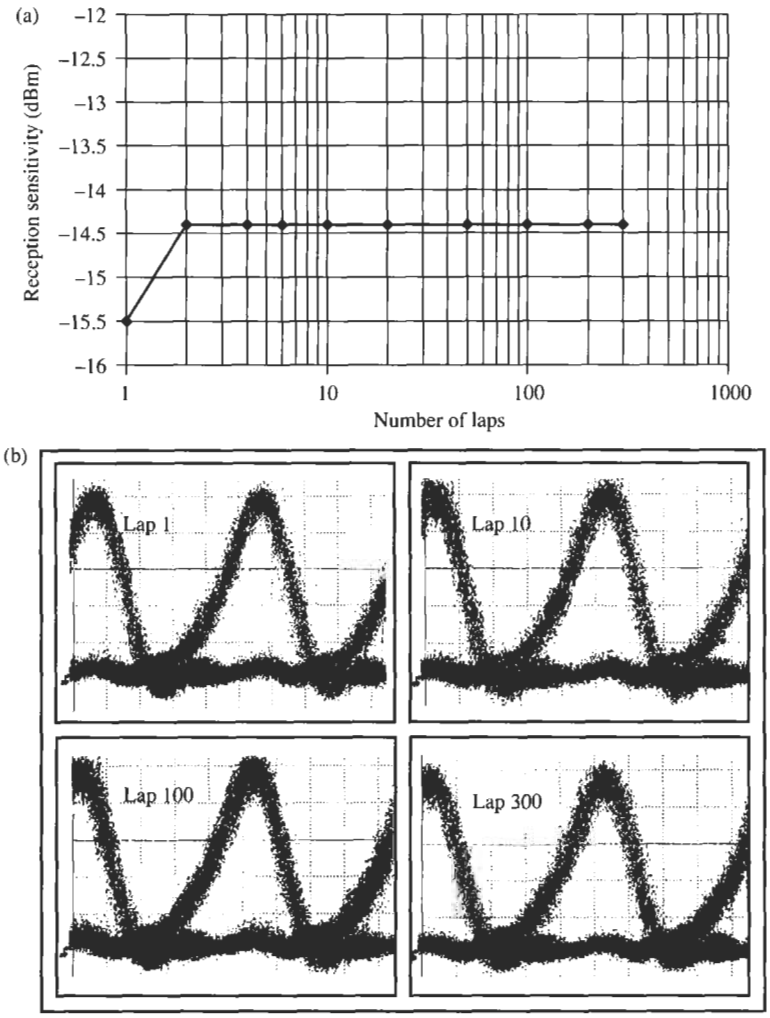


Fig. 14 (a) Evolution of reception sensitivity with number of laps; (b) eye diagram evolution in the cascade.

Prior to considering the regenerative properties of SAs, we review different technologies for implementing an all-optical function. A special attention will be dedicated to semiconductor devices, in which the saturable absorption effect relies on the control of carrier dynamics by means of the recombination center density inside the material. Key parameters such as on-off contrast (ratio of transmittance at high and low incident power), recovery time ($1/e$), and saturation energy are to be evaluated and optimized depending on the desired applications of the saturable absorber. A last but not least parameter

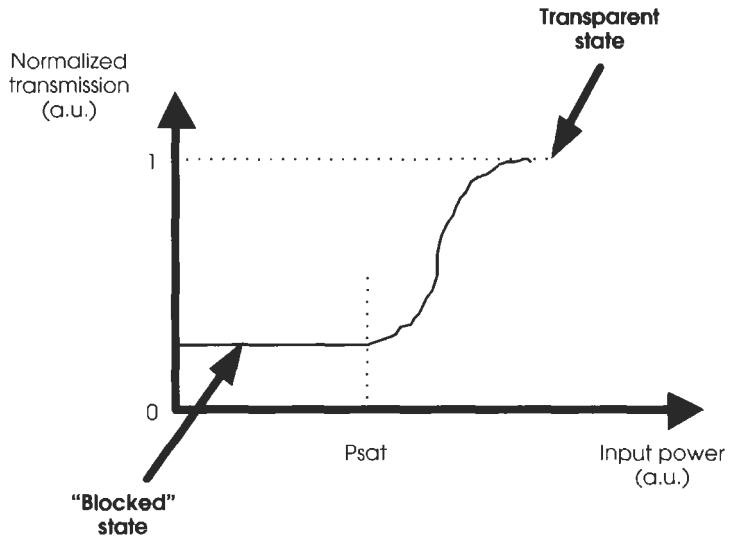


Fig. 15 Transfer function of saturable absorber.

to be closely considered is the thermal dissipation induced by high incident optical power. The main consequence of temperature elevation is a shift of the excitonic peak, resulting in reduced achievable on-off contrast.

A first possible realization uses semiconductors. The materials are prepared from low-temperature-grown (LTG) strained multiquantum wells (MQW) [36]. In this case, carrier lifetime in the SA is directly controlled by the combination of LTG [37] and Be doping (creation of recombination centers). Introduction of compressive strain on MQW and distributed Bragg Reflector (DBR) mirrors with low reflectivity considerably enhances optical nonlinearities and hence on-off contrast. In [36] was reported an optical gate of temporal width of 1.5 ps with an on-off ratio of more than 13 dB (obtained with incident pulses of 2 pJ). Another possibility to achieve SA in semiconductors is based on the technique of ion implantation. This technique was used for implementing the saturation absorption function in the input facet of a semiconductor optical amplifier (SOA) [38]. One of the main drawbacks of ion implantation is the resulting change of nature of the SA material. An alternative to this is brought by the ion irradiation technique using heavy ions such as O^+ or Ni^+ of sufficient energy to pass through the active layer, hence creating point-defect clusters and leading to a large number of recombination channels. Such an irradiation technique was shown to reduce the carrier lifetime from 1.6 ns down to 1.7 ps without significantly altering the excitonic absorption properties, thus making it possible to obtain efficient fast SA for the 1.3–1.5 μm wavelength range [39]. A complementary advantage of the technique is that

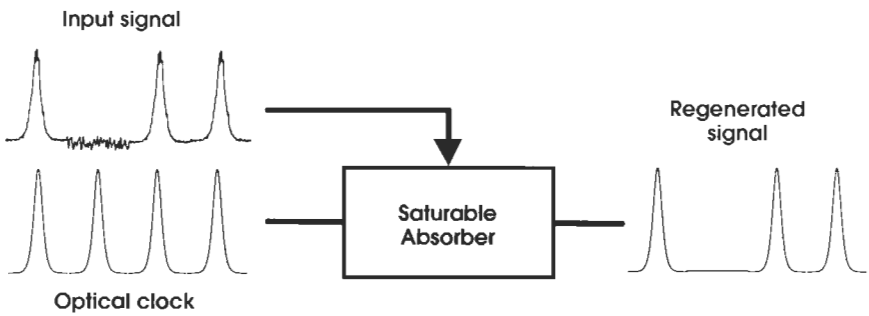


Fig. 16 Basic setup for saturable absorber used as optical gate.

irradiation of the active layers (MQW) can be performed at different stages of the process, thus calling for more practical fabrication of saturable absorbers.

Nonlinear polarization rotation (NPR) combined with a polarizer was proposed and experimentally validated at 10 Gbit/s [40]. Extension of the technique to dispersion-managed transmission was numerically investigated [41] and experimentally assessed in 10 Gbit/s loop transmission over 200,000 km [42]. One should finally mention nonlinear optical loop mirrors (NOLM) [43] or nonlinear amplifying loop mirrors (NALM) [44] as a possible mode of realization of the saturable absorption function. These fiber-based interferometers are of high interest for ultrahigh speed operation, but polarization dependence and intrinsic stability remain to be investigated and improved.

Consider the regenerative properties of saturable absorbers. The first (and straightforward) mode of application of SA for performing optical 3R regeneration revolves around their use as optical discriminator [45]. The incoming data signal triggers the switching of high-quality clock pulses through the SA nonlinear gate acting as the decision element (instead of the SOA-MZI cascade), as illustrated in Fig. 16. Output signal quality clearly depends on both the extinction ratio and the SA recovery time. A SA-based regenerator of this type was experimentally demonstrated in back-to-back configuration with 10 Gbit/s RZ signals [46]. A refined version, showing 10 dB on/off dynamic contrast and near 10 dB insertion loss made it possible to successfully switch 6 ps/10 Gbit/s data pulses with only 0.8 dB penalty. Another similar experiment, using a MLLD to provide high-quality pulses and using the SA as decision element, was recently made at 20 Gbit/s [28]. It is noteworthy that no theoretical or experimental investigations of SA regenerator cascades has been reported to date.

A second SA-based optical regenerator uses the ability of the device to reduce ASE noise level in the 0 bits, resulting in higher contrast between marks and spaces and hence a possible increase in system performance. However, since SAs do not provide effective control of amplitude fluctuations, it is therefore mandatory to associate them with a means of control for marks.

When associated with an optical amplifier, this scheme completes an optical 2R regenerator.

One possibility for controlling marks' amplitude involves the use of a gain-clamped SOA. In such a device, regeneration of spaces is achieved by the saturation absorber made out of the ion-implanted facet of a semiconductor optical amplifier (SOA) and control of marks is directly performed by means of an additional but gain-clamped SOA (GC-SOA) [38]. The apparatus was experimentally investigated through a loop experiment, demonstrating an error-free 10 Gbit/s RZ transmission over 5000 km [47]. Its regenerative capabilities were found to be mainly limited by the slow recovery gain of the SA-SOA.

Another technique for regenerating pulse marks in association with SA associates is optical filtering with nonlinear (soliton) propagation. This can be explained as follows. In absence of chirp, the soliton temporal width scales as the reciprocal of its spectral width (Fourier-transform limit) times its intensity (fundamental soliton relation). Thus, an increase in pulse intensity corresponds to both time narrowing and spectral broadening. Conversely, a decrease in pulse intensity corresponds to time broadening and spectral narrowing. Thus, the filter causes higher loss when the intensity increases, and less loss when the intensity decreases. The filter thus acts as an automatic power control (APC) in feed-forward mode, which leads to eventual power stabilization. The resulting regenerator (composed of the SA and an optical filter) is fully passive, which is of very high interest especially for submarine transmission systems where the overall power consumption is to be minimized, but does not induce any control in the time domain (no retiming effect), which could reduce the maximum achievable transmission distance. The properties of such SA-based regenerator devices have been extensively studied through analytical and numerical simulations and were predicted to significantly increase soliton systems performance. A first analytical model theoretically showed that the association of a MQW saturable absorber with a narrowband filter enabled an error-free 10 Gbit/s soliton transmission over 9000 km with 100 km amplifier spacing [48]. A complementary study reported in [49] considers a NALM as the SA device and incorporates its effect every 2000 km in the nonlinear Schrödinger equation (NLSE), predicting both single-channel 10 or 20 Gbit/s and 2×10 Gbit/s WDM (2×10 Gbit/s) transoceanic transmission. The implementation of ultrafast saturable absorbers with narrow bandpass filters in WDM long-haul transmissions has been investigated [50] with two different modes of application. The first (Fig. 17a) is referred to as *asynchronous* and uses one SA per WDM channel. The second (Fig. 17b) is *synchronous* and requires that all WDM channels be time-coincident at the input of a single SA. Under such conditions, simultaneous processing of all WDM channels within a single device is possible, which entails significant economy in regenerator complexity. Numerical simulations using split-step Fourier resolution of the NLSE show that both asynchronous and synchronous 4×10 Gbit/s

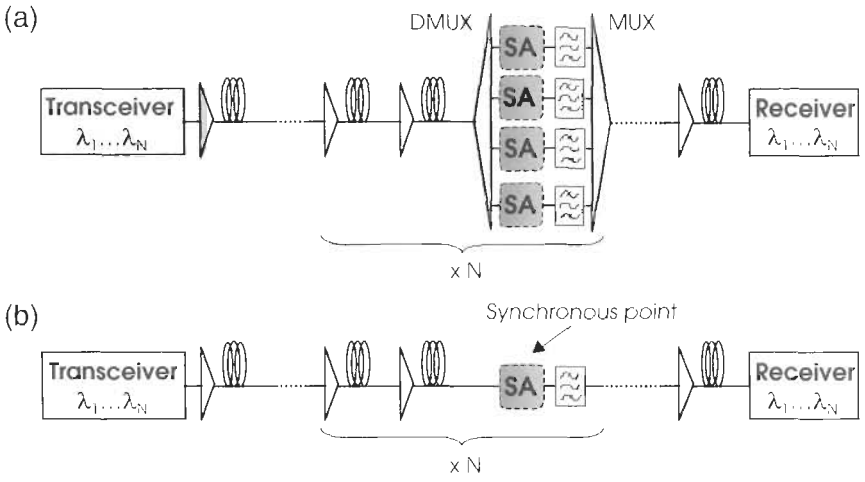


Fig. 17 Two WDM system configurations for applying SA function: (a) asynchronous, with inline demultiplexing, individual saturable absorption (SA), and filtering; (b) synchronous, with single SA common to all WDM channels and filtering with periodic transfer function.

transmission over 10,000 km was feasible with the periodic insertion of SAs every 200 km (50 km amplifier spacing). In this study, the effect of SA time dynamics was included and its impact on system performance evaluated. The study concludes that SA recovery times should be shorter than $T_{bit}/5$. Finally, the potential of SA-based 2R regeneration with filtering for high bitrate transmission was further assessed by numerical predictions concerning 80 Gbit/s soliton transmission over 9000 km [51].

As mentioned before, nonlinear gates such as saturable absorbers significantly modify the nature of both noise and pulse timing position distributions (PDF). Under these conditions, Gaussian approximations used in models could lead to biased results in the prediction of system performance and should not therefore be taken as absolute results. For this reason, complete investigation and assessment of the regenerative capabilities of saturable absorbers clearly require experimental implementation and cascadability measurements with the device.

We now consider a passive 2R regenerator based on the MQW microcavity SA operating in association with both optical filter and nonlinear propagation. The first apparatus was tested in 20 Gbit/s transmission loop experiments [52]. Figure 18a shows the SA structure [53]. The MQW device operates as a reflection-mode vertical cavity, providing both a high on/off extinction ratio by canceling reflection at low intensity and a low saturation energy of 2 pJ. It is also intrinsically polarization-insensitive. Heavy ion irradiation of the SA ensures recovery times shorter than 10 ps. Prior to implementing the device in

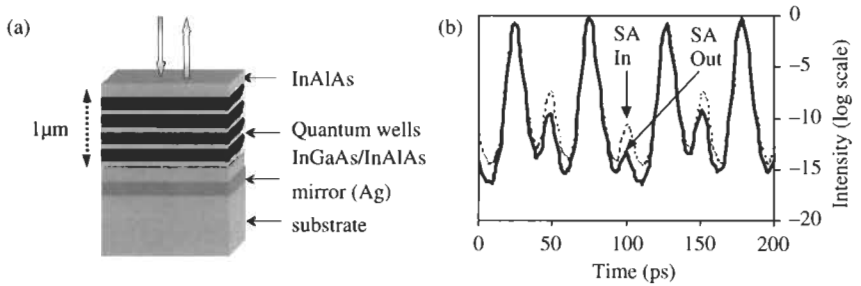


Fig. 18 (a) Structure of multiquantum well saturable absorber and (b) dynamic contrast characterization of saturable absorber with 20 GHz RZ pulses incident on the SA (dotted line) and after passing through the SA (bold line).

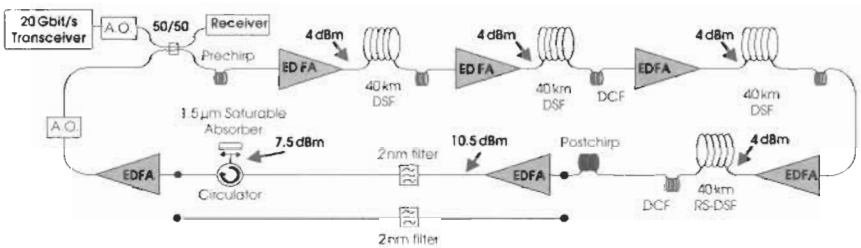


Fig. 19 Loop configuration SA+ filter configuration in 20 Gbit/s experiment.

transmission experiments, a full characterization at 20 Gbit/s and over a broad wavelength range spanning the entire EDFA C-band (1530–1560 nm) was conducted with incident 8 dBm RZ signal made up of 6 ps pulses. Figure 18b shows the 20 Gbit/s pulse train before and after reflection on the SA sample. A clear improvement in the extinction ratio of the 20 Gbit/s pulses train is observed with an average increase of the on-off contrast of 2.5 dB over the entire C-band, showing possible WDM operation. These measurements also show that the recovery time of 10 ps is short enough to make the SA operate at 20 Gbit/s (thus confirming previous numerical predictions [50]), but also indicate possible mark/space discrimination at 40 Gbit/s line rates.

Figure 19 shows the experimental setup of the 20 Gbit/s recirculating loop experiment, as described in [52]. Only when operating with the SA, an additional erbium-doped fiber amplifier (EDFA) is placed in the loop before the device to provide 7.5 dBm incident power on the SA sample and also to compensate for its 7 dB insertion loss.

Figure 20 shows the experimental BER vs. propagation distance for two different ASE noise levels in loop. Without SA and associated EDFA, the maximum error-free distance (BER $< 10^{-9}$) is measured to be 6200 km at 20 Gbit/s (10 Gbit/s) line rate. When inserting the SA with the loss-compensating EDFA,

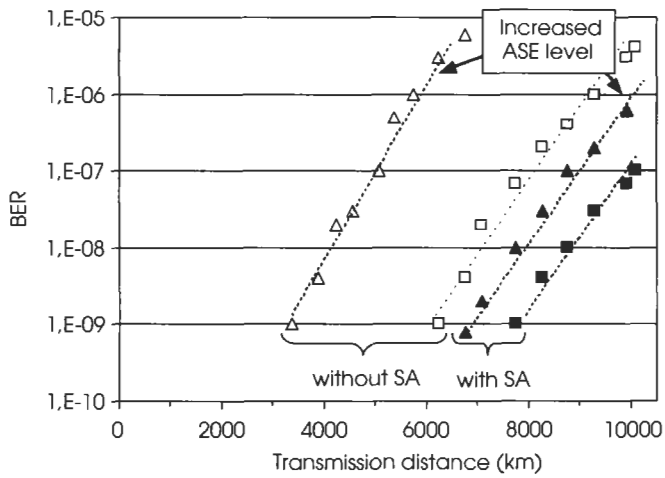


Fig. 20 Experimental BERs vs. propagation distance with SA (full symbols) and without SA (open symbols) for 2 different ASE levels in 20 Gbit/s loop.

a clear improvement in the system performance is observed, yielding an increased error-free distance of 7800 km. In this case, the main system limitation stems from timing jitter accumulation rather than amplitude fluctuations. In order to further investigate SA discrimination efficiency, the ASE noise was purposefully increased by introducing 3 dB attenuation before two-line EDFA. As expected, the system performance is significantly degraded and the error-free distance is reduced to 3400 km without SA. When implementing the SA (with additional EDFA), the error-free distance increases to 6800 km. Comparison of this result with the previous one obtained without any degradation of line EDFA noise figure (error-free distance of 7800 km instead of 6200 km at 20 Gbit/s) further illustrates the SA capability to reduce ASE at 20 Gbit/s line rates. It also demonstrates its regenerative capability in SNR-limited transmission, when associated with nonlinear transmission and optical filtering.

Future improvements for SA-regenerators toward 40 Gbit/s implementation will address the reduction of saturation energy, recovery times, and insertion loss, with an increase in the dynamic extinction ratio. In WDM applications, a key interest of SA regeneration is the possibility of processing several wavelength channels within the same device. Simultaneous WDM implementation raises the issues of dynamic extinction ratio behavior, need for resynchronization of the channels, and packaging with MUX/DMUX stages. Some properties of SA wavelength selectivity, such as found in quantum dots [54], could also be advantageously exploited.

III. Optical 3R regeneration by Synchronous Modulation

A. PRINCIPLES AND DEVICES

All-optical 3R regeneration can be achieved through inline synchronous modulation (SM) associated with narrowband filtering (NF) [55]. Modulation can be made on either signal intensity (IM) or phase (PM), or a combination of both [56]. As shown in Fig. 21, the layout of the SM regenerator is straightforward. It includes a clock recovery block, a modulator, and an optical filter [55, 57]. Regeneration through SM-NF intrinsically requires nonlinear (soliton) propagation in the trunk fiber following the SM block. Therefore, one can refer to the approach as distributed in contrast with lumped optical regeneration, where 3R signal processing is completed within the regenerator block and is independent of the line transmission characteristics. As will be discussed later, it is, however, possible to make the regeneration and transmission functions work independently in such a way that *any* type of RZ signals (soliton or nonsoliton) can actually be transmitted through the system. We refer to this approach as black box optical regeneration (BBOR) [58].

The effects of IM and PM on optical pulses are illustrated in Fig. 22. IM and PM have different qualitative effects. Both act on the timing position of the pulse (jitter suppression), but IM also removes a fraction of the noise background. Such a background results from the ASE accumulated from line amplifiers, the dispersive waves (continuum) generated by nonsoliton pulse components, and other deleterious perturbations. This noise background manifests itself by unwanted energy in the 0 slots and pulse distorsion in the 1 slots. Suppressing this noise background with adequate periodicity leads to SNR stabilization throughout the system, as explained hereafter. Figure 22a shows the effect of associated periodic synchronous IM with nonlinear propagation. Upon passing through the first IM modulator, the noise in the 0 slots is shaped into dark pulses having about 50% less energy, while the 1 slots' energy

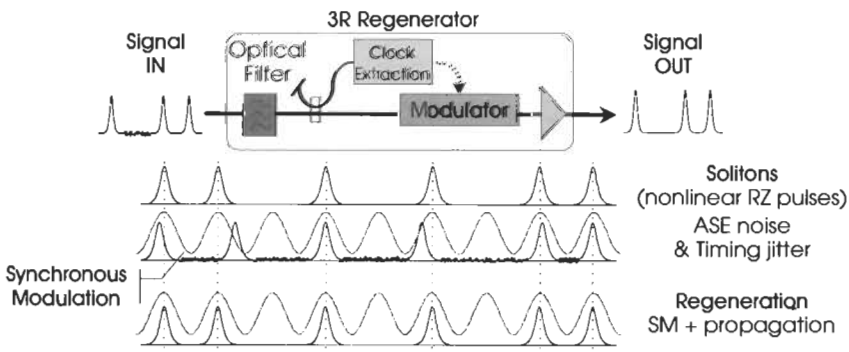


Fig. 21 Basic layout of the all-optical regenerator by synchronous modulation and narrowband filtering and illustration of the principle of operation.

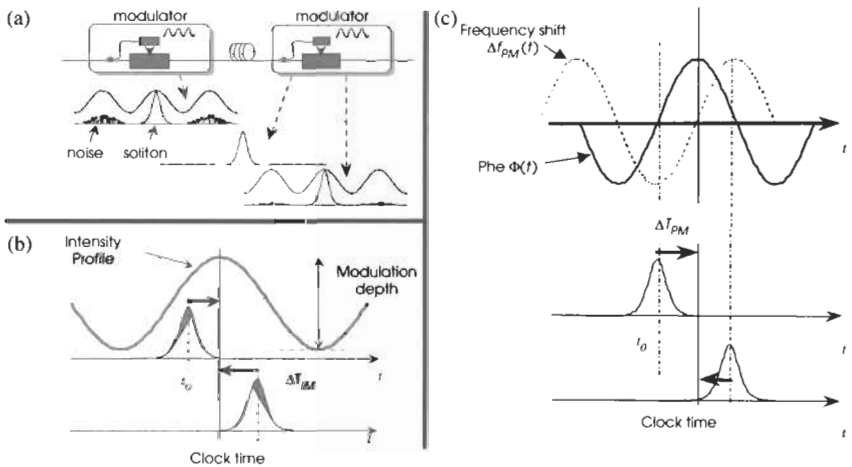


Fig. 22 Principle of operation of a) IM on ASE noise and dispersive waves, b) IM on timing jitter, and c) PM on timing jitter.

remains unchanged. Upon dispersive propagation, the 1 pulses keep a constant shape (soliton effect), while the 0 pulses spread in the time domain. The noise energy fills up the gaps left by the previous modulation stage. After several iterations of this lossy process, the noise background reaches an average and stable level; its rate of growth with distance matches its rate of loss from IM and dispersion. This feature causes asymptotic convergence toward constant SNR, which enables transmission over virtually unlimited distances [55].

Next, we consider the effect of IM and PM in the time domain, showing that both modulations act as a time-reference pull-back force, resulting in jitter suppression. Figure 22b shows the effect of IM on timing jitter. When unsynchronized or jittered pulses undergo IM, asymmetric pulse loss results. The loss is higher on the edge of the pulse that is farthest from the modulation center. This causes the center of gravity of the pulse energy to move closer to the modulation center. The effect of soliton propagation causes the pulse to reshape about this new time position. With iteration of this process, the timing position converges toward the modulation center, which is an effect of resynchronization. On the other hand, the asymmetric loss associated with this resynchronization causes random intensity fluctuations [56]. This can be viewed as a classical PM to IM noise conversion effect. In turn, IM noise can be suppressed using adequate narrowband filtering in accordance with the previously described principle (see Section II D).

Figure 22c shows the effect of phase modulation (PM) on timing jitter [57]. Jitter suppression is based on the property that the frequency shift induced by PM depends on the timing position. Since different pulse positions correspond to different frequency shifts, PM noise is converted to FM noise. With dispersion, FM noise is converted back into PM noise. Depending upon the

dispersion sign this second conversion either enhances or cancels the initial PM noise. Such a jitter cancellation works as follows. Incident pulses that come behind the center of the modulation (Fig. 22c, bottom axis) experience a positive frequency shift ($\Delta f(t) > 0$). Soliton propagation causes this shift to become equally distributed within the pulse. In the positive-dispersion regime, blueshifted pulses travel faster. Thus, the pulse delay causes an increase in group velocity. The pulse travels faster, which cancels the delay after some propagation distance. The opposite happens to pulses arriving ahead of time (Fig. 22c, middle axis). PM causes a redshift, which in turn corresponds to group velocity decrease and pulse delay. To summarize, PM accelerates delayed pulses and slows down advanced pulses. One should note that PM with the opposite sign causes the opposite effect or jitter aggravation (advanced pulses being accelerated and delayed pulses being slowed down). Thus, proper PM sign must be applied in order to suppress timing jitter. Unlike IM, PM does not cause intensity noise; but it has no effect on ASE noise, which is insensitive to the modulation. This points to the advantage of combining IM and PM for most efficient jitter suppression with ASE cancellation and asymptotic SNR stabilization, with modulation-induced intensity noise being kept minimal.

The optimization of SM-based regenerator has several requirements: the adequate tuning of IM/PM depths, the use of proper PM phase (the derivative being positive at the reference clock time), and the inclusion of an optical filter with adequate filter bandwidth. Concerning this last point, it is clear that the filter bandwidth must be optimized according to IM depth. A filter that was too broad would have no effect in intensity noise reduction, while a filter that was too narrow would cause excessive signal loss. Compensation of this loss means excess gain for the ASE noise, which results in exponential ASE accumulation and rapid SNR degradation. The key of the filter optimization is to ensure that PM to IM conversion is adequately reduced, while the resulting excess loss does not cause significant SNR degradation by excess ASE buildup. Such an optimization was experimentally realized [59]. It should be noted that this regeneration technique requires soliton-like pulses, without which time and intensity noises could not be controlled. In turn, this requirement is of major import in transmission line design: regeneration and transmission cannot be designed independently from each other. As discussed below, however, it is possible to separate these two functions with the BBOR approach, in such a way that the transmission line characteristics can be chosen (or designed) independently from the SM-NF regenerator.

1. Single-Channel, 40 Gbit/s Regenerated Transmission

Inline optical regeneration is a key solution when conventional or unrepeatered techniques fail to provide high capacities and low BERs at given transmission distance. This statement is particularly relevant when considering transmissions at channel rates from 40 Gbit/s and above (160–1000 Gbit/s), which are

highly vulnerable to nonlinear propagation/interaction effects, self-induced electrostriction, and overall, PMD. These limitations, which are exclusively due to timing jitter, result in a dramatic reduction of low-BER transmission distances. Optical 3R regeneration by SM cleanses timing jitter, regardless of its physical origin and hence enables virtually unlimited transmission distances. Distance is thus no longer a limiting parameter in regenerated systems, unlike in the conventional case. The received BER can even be significantly improved, which alleviates the need of ECCs and allows greater utilization of bandwidth. The BER levels, however, cannot be made arbitrarily small. There are intrinsic limits to the signal quality resulting from 3R regeneration [60] as will be explained in the next subsection.

Experimental implementation of SM-NF regeneration at 40 Gbit/s, using different components (as discussed below) have led to transmissions over more than 10,000 km distances [61–64]. The approach would become truly attractive if greater line rates and overall capacities could be achieved, to the exclusion of any other technique. Thus, 40 Gbit/s single-channel investigations can be seen as a first step toward this goal.

The first components used for applying SM to 10 Gbit/s soliton pulses have been $LiNbO_3$ Mach-Zehnder (MZ) modulators [55]. Such devices have low insertion loss (typically <5 dB) and can easily be operated at rates up to 40 GHz (SM 3R regeneration requires only narrowband elcctronics). However, $LiNbO_3$ MZ are plagued by strong polarization dependence. This dependence can be cancelled out using two orthogonally butt-coupled modulators [65], resulting in higher insertion loss and a more complex regenerator (including two modulator/drivers). Further, the instability of the operating point with respect to temperature needs to be suppressed, introducing extra complexity. As an alternative to $LiNbO_3$ devices, electroabsorption modulators (EAM) have been investigated for SM-based regeneration applications [66–68]. Indeed, EAM exhibit several advantages with respect to $LiNbO_3$ devices, such as lower driving voltages for similar IM depths, negligible polarization sensitivity (<1 dB) and no time drift of the operating point. However, the EAM modulator has higher fiber-to-fiber insertion loss (typically 10–12 dB). For high bitrate operation (e.g., 40 Gbit/s), the major drawback of EAMs is the absence of independent control of IM and PM depths; such a control is required to finely tune the regeneration effect [59]. Indeed, PM characteristics are directly linked to those of IM through the component characteristics (Henry or α factor). Another drawback is the strong WDM crosstalk when several WDM channels are input to the EAM, thus forbidding implementation in simultaneously WDM regeneration.

The key advantages of $LiNbO_3$ interferometers and EAMs were combined in a novel 40 GHz InP Mach-Zehnder-based regenerator, as recently proposed and demonstrated at 40 Gbit/s [63]. Figure 23 shows an ESM picture of the chip (left) and a picture of the packaged version. The device was shown to perform both polarization-insensitive (<1 dB) and wavelength independent (in the

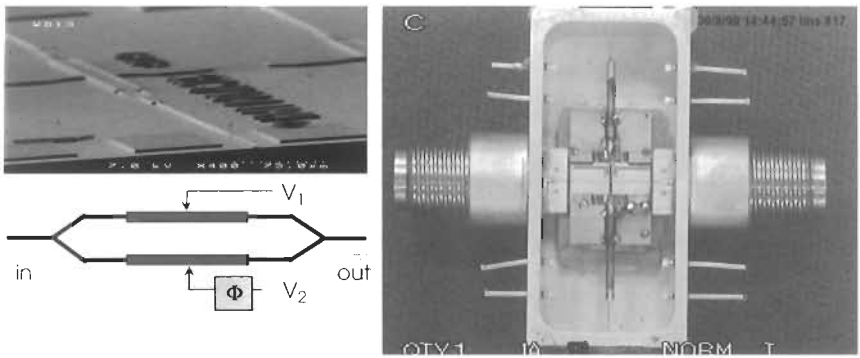


Fig. 23 Top left: ESM picture of InP Mach-Zehnder modulator; right: packaged version.

range of 1545–1560 nm) modulations with beneficial adjustment of IM/PM depths provided through a dual-electrode configuration. It was also demonstrated to exhibit long-term stability of operating point with moderate driving voltages. Moreover, no significant WDM crosstalk was observed or measured with up to 4 channels at 40 Gbit/s, thus making this InP MZ well suited for simultaneous WDM regeneration. Finally, no Peltier cooler is required for satisfactory system operation, an advantage for low-power applications such as submarine cable transmission. One should finally mention other devices such as nonlinear loop mirrors (NOLM) and Kerr fiber modulator (KFM), which have been proposed for applying SM-based optical regeneration [2]; these have been experimentally validated at only 20 Gbit/s so far.

2. Optimizing SM-Regenerated Transmission Systems

While synchronous modulation was experimentally shown to enable single-channel 40 Gbit/s global distance transmission, the analysis and optimization of regenerator characteristics at both physical and system design levels has only slowly progressed. So far, a few models and methods have been proposed.

A first type of model, based upon perturbation theory, is well suited to the analysis of single-channel soliton systems. Soliton propagation was shown to be described through the (perturbative) evolution of four normalized soliton parameters, which are amplitude A, timing position t_0, phase ϕ, and frequency ω and which obey a coupled linear equation system [69, 70]. The effects of synchronous modulation (IM and/or PM) and narrowband filtering can easily be incorporated in such perturbative models, provided induced perturbations remain small [71]. Such models provide both a simple physical description and a qualitative evaluation of the regeneration process, mainly through stability analysis [60]. The main limitation of the perturbative models is that they do not include ASE noise or provide statistical estimates of pulse behavior, which excludes any system BER prediction.

Numerical simulations based on the resolution of the nonlinear Schrödinger equation (NLSE) by split-step Fourier transform represent the most straightforward, complete, and accurate means of investigation and optimization of regenerated transmissions. As in any numerical model, however, physical insight is somewhat lost and multidimensional parameter optimization makes overall conclusions more subjective. NLSE modeling was used to first demonstrate the asymptotic stabilization of Q-factors with distance (or with cascaded regeneration stages) characteristic to optimized regenerated systems [56]. Recent numerical analyses have shown that SM efficiently also cancels electrostriction jitter [72] and reduces PMD impact on transmission [73] when applied to 40 Gbit/s soliton pulses. This powerful tool also makes it possible also to quantitatively predict actual transmission systems performance. Though optical regeneration is known to change statistics for both energy and timing distributions of output signals, evaluation of SM-regenerated system performance using Gaussian statistical approximation in numerical simulations was shown to remain dependable in most regenerated system configurations (provided PM is used at low modulation index ($\Delta\Phi < 10$ deg) in combination with high-index IM) [74]. Further work is required to fully analyze the evolution of non-Gaussian statistics throughout the system so as to evaluate the BER with improved accuracy. Direct numerical predictions of BERs is not possible due to the forbidding computation times required to acquire the corresponding probability distributions over a sufficient dynamic range. Therefore, new statistical models remain to be developed, although they will only remain of indicative value, given the high complexity of the problem.

With a view to simplifying regenerated systems analysis and optimization, a predictive analytical model based on Gaussian pulse/Gaussian statistics was recently developed and reported in [75]. This model permits thorough analysis of IM-regenerated transmissions systems and evaluation of optimum operating regions. It should be noted that since it only incorporates intensity modulation (and not phase modulation), BER estimates based on Gaussian approximations remain valid. This analytical model uses three fundamental postulates, which directly apply to soliton-like pulses: shape invariance (Gaussian), self-retiming about the new center of gravity defined by IM, and frequency width proportional to pulse energy. Hence it conveniently emulates nonlinear signal behavior without introducing complexity. As illustrated in Fig. 24, amplifier noise and timing jitter are accounted for as intrinsic system contributions from line segments separating two regenerators, while there are no other assumptions concerning signal transmission. As an illustration of the accuracy of this analytical model, we investigate both numerically and analytically eye reopening, a key effect induced by SM-based optical regeneration.

Figure 25 shows the evolution of both amplitude and timing Q-factors with distance as computed analytically and numerically in a regenerated 10 Gbit/s system with initial low amplitude Q-factor of 6 (amplitude fluctuations). In

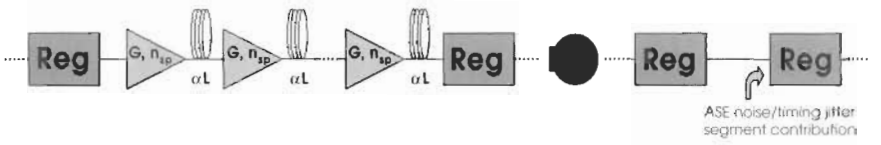


Fig. 24 Illustration of the line description used in the analytical model for predicting regenerated systems performance. Actual line parameters such as amplifier gain G, noise factor n_{sp}, fiber length L, and loss α are evaluated as intrinsic contributions between two regenerators.

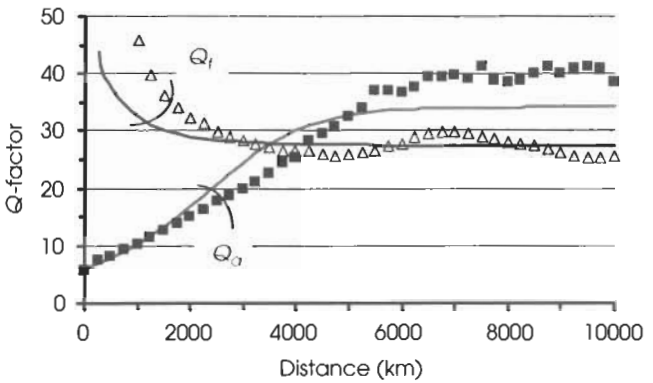


Fig. 25 Evolution of amplitude and timing Q-factors with transmission distance, as computed numerically (marks) and analytically (lines) in 10 Gbit/s system with low amplitude Q-factor ($Q_a = 6$).

this example, initial pulses are of Gaussian shape with 20 ps temporal width. Line fiber chromatic dispersion and attenuation are 1 ps · nm^{-1} · km^{-1} and 0.2 dB/km, respectively. Amplifier ASE noise figure, output power, and spacing are 5 dB, 1.5 dBm, and 50 km, respectively. Regenerators are placed every 5 EDFA (Za = 50 km) and apply a 10 dB intensity modulation associated with 0.42 nm filtering (equivalent to 0.2 rad/s). Resulting timing jitter represents an intrinsic contribution of 4% of the bit duration T_{bit} at the input of every optical 3R regenerator. In the case of numerical simulations, 3 runs of 128 bits (PRBS coding) having different noise seeds and a temporal resolution of 0.78 ps have been concatenated. The amplitude Q-factor (linear) is seen to initially increase from low level $Q_a = 6$, as a consequence of the efficient suppression of amplitude fluctuations by adequate narrowband filtering, causing effective eye re-opening, before stabilizing to an asymptotic level of 33 reached near 6000 km. Meanwhile, timing Q-factor decays until 2500 km, where it stabilizes to $Q_t = 28$. Overall system performance (taken as minimum from Q_a and Q_t) increases and then stabilizes after a transition from intensity-noise-limited ($Q_a < Q_t$) to timing-jitter-limited regimes ($Q_a > Q_t$). Because of the

initial eye-opening effect, the asymptotic values are higher than the transmitter's. This example shows the importance of a proper optimization between IM depth and filtering width. An excellent agreement between numerical and analytical predictions for Q-factors is obtained in this example, further validating the model.

The complementary effect to eye re-opening in the time domain (initial increase of Q_t with distance) was also investigated and first predicted through numerical simulations [76]. An experimental observation of this effect is reported in [63], where poor initial temporal interleaving of a 40 Gbit/s OTDM pulse train was corrected after some loop laps in the presence of synchronous modulation, leading to an increase in measured system BERs with distance. It should be mentioned that such an increase of BER with distance does not represent any recovery of erroneous transmitted data but stands for a better matching of optical pulse characteristics with those of the receiver (in particular the decision threshold). This further points out the need of optimizing optical receivers to take full benefit of optical 3R regeneration in transmission systems.

B. THE BLACK-BOX OPTICAL REGENERATOR (BBOR) APPROACH

In the previous sections, SM-based optical regeneration was shown to be fully efficient on soliton pulses, which makes both the regeneration and the transmission characteristics intrinsically linked. As we will see in this section, it is possible to separate these two functions with the black-box optical regeneration (BBOR) approach. Under this condition, one can take benefit of dispersion management (DM) features mainly for increasing spectral efficiency of ultrahigh capacity long-haul transmission systems, while ensuring high transmission quality thanks to SM-based regeneration.

The BBOR technique consists in the addition of an adaptation stage for incoming RZ pulses in the SM-based optical regenerator as to ensure high regeneration efficiency whatever the data type (linear RZ, DM-soliton, C-RZ, etc.). This is achieved using a local and periodic soliton conversion of RZ pulses by means of adequate input power into a length of fiber with anomalous dispersion. Figure 26 shows the BBOR structure enabling the successful association of DM propagation and optical regeneration by SM, as initially proposed in [58]. The approach is investigated by means of numerical simulations and

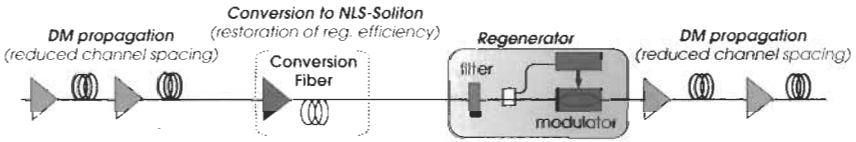


Fig. 26 Schematic of black-box optical regenerator (BBOR).

single-pulse analytical models at 40 Gbit/s line-rate in [77], where local conversion of propagating DM pulses into NLS solitons through the conversion fiber is shown to fully restore filter control of SM-induced amplitude-energy fluctuations resulting from timing jitter correction. Moreover, regeneration efficiency is at a maximum when the BBOR is located at the chirp-free point (where pulse width is minimum).

Such key advantages of the BBOR approach (over the classical SM-based scheme) have been experimentally assessed by means of a 40 Gbit/s DM loop experiment [78]. Figure 27 illustrates the BBOR superiority with both analytical and experimental results. The left side of Fig. 27 shows the results of a numerical single pulse study in which the stabilizing effect of filters is represented for different initial deviations of pulse average power with either the classical SM-based regenerator (top) or BBOR (bottom). Comparison of the two graphs shows the dramatically beneficial impact of conversion on filter efficiency. Eye diagrams from the 40 Gbit/s DM transmission experiment (Fig. 27, right) incorporating either classical SM-based regenerators (top, 1280 km) or BBOR (bottom, 10,000 km) further illustrates the reduction of amplitude fluctuations (increase in transmission distance) when using the second regenerator configuration. It should finally be noted that the analytical model [75] for predicting IM-regenerated systems performance remains fully valid when considering BBOR regenerator, as the model only requires soliton-like behavior at the regenerator location.

One intrinsic property of BBOR-regenerated systems is their tolerance to various initial pulse shape/formats, as illustrated in Fig. 28, which shows the theoretical evolution with distance of the worst Q-factor in a 8 × 40 Gbit/s

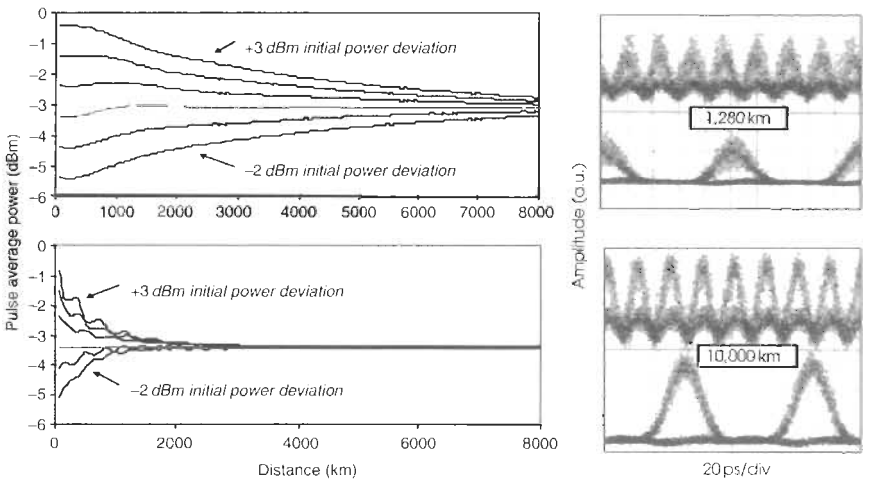


Fig. 27 Numerical (left) and experimental (right) demonstration of recovered filters (regeneration) efficiency owing to DM to NLS soliton conversion.

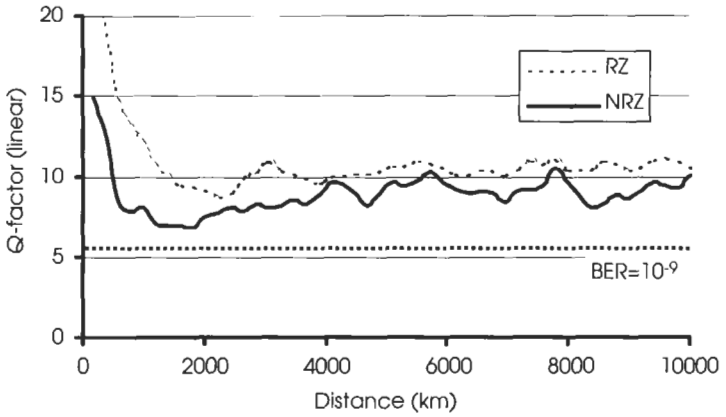


Fig. 28 Theoretical evolution of Q-factors with distance for a regenerated 8×40 Gbit/s transmission with NRZ or RZ initial coding.

regenerated system with either NRZ or RZ input pulses. In both cases, Q-factors are seen to asymptotically stabilize with distance, as expected from efficient optical regeneration. When initial NRZ coding is employed, a stronger initial degradation of the Q-factor can be observed, a result of the adaptation of pulses to the stable state of the system. However, only a slight difference in the asymptotic value of Q-factors ($Q = 10$ and 9 for systems with initial RZ or NRZ pulses, respectively) is ultimately obtained, showing that initial format or pulse shape has only a limited impact on regenerated system asymptotic performance. This result could translate into a cost-effective simplification of system transceivers with the removal of RZ-shaping stages.

As mentioned before, another feature of BBOR is the possible increase in spectral efficiency in regenerated long-haul transmission systems. To investigate such a feature, we have conducted a numerical study of the optimal symmetric dispersion maps (based on anomalous dispersion fiber D+ and dispersion-compensated fiber (DCF)) for passive 40 Gbit/s-based WDM transmission systems. The choice of $+8\,\text{ps} \cdot \text{nm}^{-1} \cdot \text{km}^{-1}$ fiber (TeraLight^{TM})-based dispersion map with near 100% line compensation was found to be optimal for WDM systems with spectral efficiency greater than 0.2 bit/s/Hz, as a result of a tradeoff between intra- and interchannel interactions [79]. A complementary numerical investigation was carried out for the compensating or negative dispersion fibers (NDF) to be associated to TeraLight^{TM} D+ fiber, leading to an optimal NDF with $-16\,\text{ps/(nm} \cdot \text{km)}$ chromatic dispersion enabling full dispersion compensation over the entire C-band [80]. This newly optimized symmetric map based on $+8/-16\,\text{ps} \cdot \text{nm}^{-1} \cdot \text{km}^{-1}$ fibers was finally implemented in a 32×40 Gbit/s regenerated transmission system with 150 GHz channel spacing (0.3 bit/s/Hz spectral efficiency). Figure 29 shows the theoretical evolution of Q-factors with distance in such a 1.28 Tbit/s

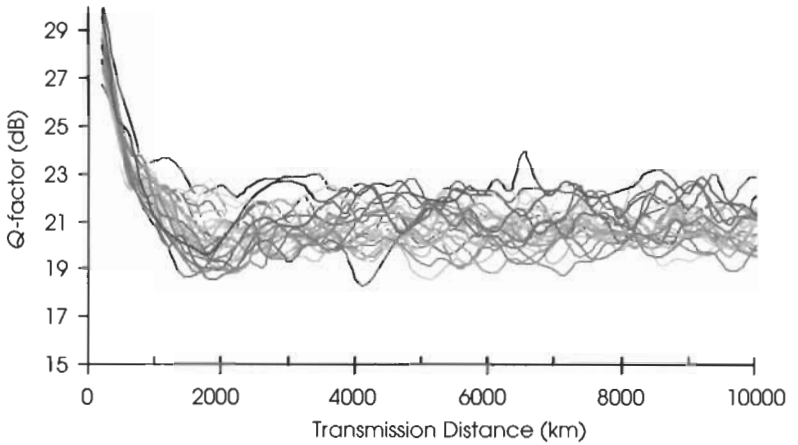


Fig. 29 Theoretical evolution of Q-factor as a function of distance in a regenerated 32×40 Gbit/s transmission using optimized D+/NDF dispersion map.

transmission. All Q-factors are seen to decay with distance but reach an asymptotic value above 19 dB (linear $Q > 8.9$) at 10,000 km distance. This numerical result assesses the BBOR potential for addressing high spectral efficiencies in $N \times 40$ Gbit/s long-haul transmission systems in combination with appropriate dispersion management. It also demonstrates the feasibility of Tbit/s 40 Gbit/s-based transoceanic systems with up to 0.3 bit/s/Hz spectral efficiency.

The issue of SM-based regeneration (BBOR) periodicity along the transmission line must also be addressed since it directly impacts on system complexity, cost, and power consumption and also on the overall performance. This was first studied through 20 Gbit/s loop experiments, where the maximum regeneration spacing ensuring error-free performance (BER $< 10^{-9}$) was found to be 140 km [81] and 200 km when associated with inline remote pumping of Er-doped line fibers [82]. In the case of regenerated $N \times 40$ Gbit/s transmission systems, current numerical investigations find an achievable regeneration periodicity of up to 600 km (amplifier spacing of 40 km) when considering aforementioned optimized symmetrical dispersion maps.

C. WDM COMPATIBILITY & APPLICATION TO 4 × 40 Gbit/s LOOP EXPERIMENT OVER 10,000 km

Since high-capacity systems are all based on wavelength-division multiplexing (WDM), multiwavelength optical regeneration represents a key issue. In particular, simultaneous regeneration of WDM channels through a single 3R device would represent a high added value for regenerated systems.

In this section, we review three possible implementation schemes of optical 3R regeneration with WDM propagation. The first (Fig. 30a) involves allocating a regenerator to each WDM channel. With such a system configuration,

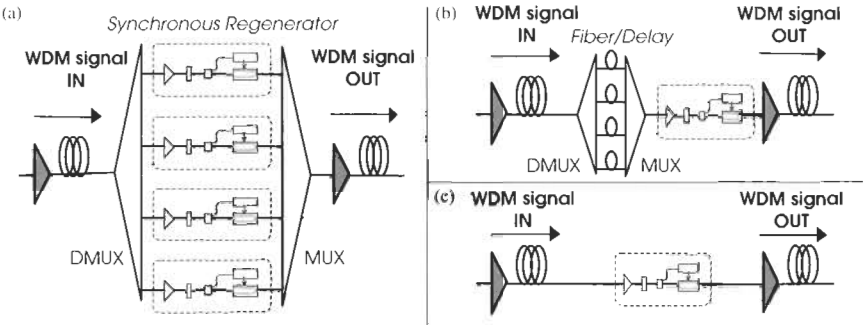


Fig. 30 Basic implementation schemes for WDM all-optical regeneration: (a) parallel asynchronous, (b) serial resynchronized, and (c) serial self-synchronized.

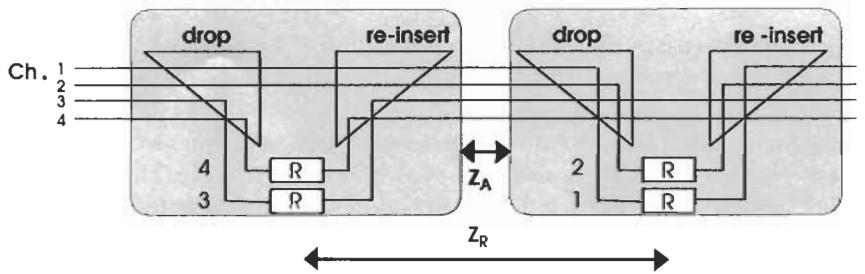


Fig. 31 Distributed regeneration principle.

the penalty-free scalability of regenerated long-haul transmissions was first shown in [83], where a 25 × 40 Gbit/s NLS soliton transmission was numerically demonstrated over 10,000 km. The second scheme involves sharing a single modulator, thus processing the WDM channels at once in serial fashion. This approach, first proposed in [84] and numerically studied in [85], requires WDM synchronicity, meaning that all bits are synchronous with the modulation. Such WDM synchronicity can be achieved either by use of appropriate time-delay lines located within a Dmux/Mux apparatus (Fig. 30b), or by making the WDM channels inherently time-coincident at specific regenerator locations (Fig. 30c). This WDM regeneration scheme also requires the regenerator to be WDM crosstalk-free, as is the case of the newly developed InP Mach-Zehnder interferometer [63]. Clearly, serial regeneration scheme is far simpler and cost-effective than the parallel version; however, optimized resynchronization schemes still remain to be developed for industrial applications. The third regenerator configuration, referred to as the distributed WDM regenerators, is shown in Fig. 31 and was initially proposed and studied through numerical simulations [86]. The principle of this novel approach consists in periodic, drop-and-reinsert channel-selective regeneration. Under these conditions, the WDM regenerators are more closely spaced but include

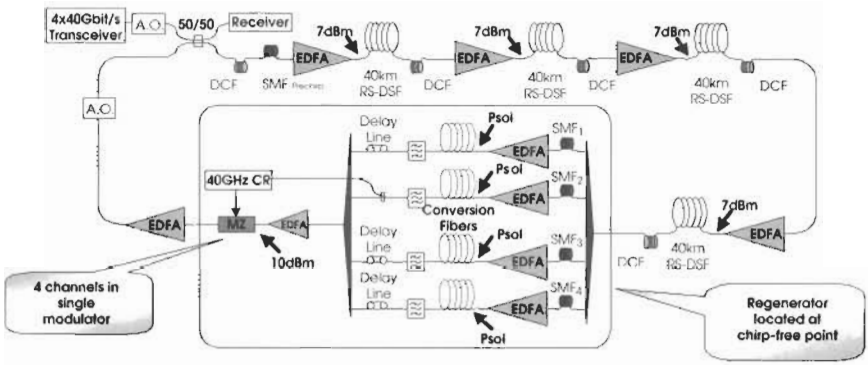


Fig. 32 Experimental loop configuration for 4 × 40 Gbit/s regenerated transmission over 10,000 km using a single all-optical regenerator.

a limited number of parallel stages, leading to possible major simplification in the repeater/regenerator design. Moreover, since this distributed regeneration technique requires the development of a single repeater type, it appears very well suited for submarine applications.

To experimentally demonstrate the BBOR capabilities for WDM regeneration, a 4 × 40 Gbit/s dispersion-managed loop experiment with 1.6 nm channel spacing (0.2 bit/s/Hz) was conducted. Figure 32 shows the 160 km loop configuration incorporating the WDM-compatible BBOR, whose full technical description is reported in [87]. Prior to being optically regenerated, the channels are spectrally demultiplexed. For each channel, a soliton conversion stage changes DM pulses into NLS solitons. After passing through optical filters with optimized bandwidths, all four WDM channels are then recombined through a 200 GHz multiplexer and simultaneously input in a packaged InP Mach-Zehnder modulator with 10 dBm optical power. The modulator is electrically driven by a 40 GHz RF clock extracted from channel 2 by means of high-Q (≈1000) RF filtering. Optical delay lines are inserted to control the optical paths of channels 1, 3, and 4 to ensure that pulses from all four channel remain time-coincident in the modulator. At 10,000 km distance, asymptotic Q-factors of all 4 channels were measured to be greater than 14.5 dB (5×10^{-8} BER) for $2^{31} - 1$ coding length. This experimental result clearly shows the absence of transmission impairments due to long coding sequences in SM-regenerated transmission systems. Moreover, it fully validates the principle of simultaneous regeneration of several WDM channels in a single modulator and confirms the absence of WDM crosstalk in the InP Mach-Zehnder modulator.

This 160 Gbit/s transmission experiment (yielding a capacity-distance product of 1.5 Pbit/s · km) also further assesses the compatibility of 3R regeneration with dispersion management through the BBOR approach, as shown

through high system performance at 10,000 km. Finally, a comparison of system performance obtained at 10,000 km in this 4 × 40 Gbit/s experiment with that obtained in the single-channel 40 Gbit/s reported in [78] shows a limited penalty of 2 dB in system performance arising from WDM propagation (with 4 channels), thus experimentally confirming the WDM scalability offered by WDM compatible BBOR in high bitrate long-haul transmissions.

D. OPTICAL REGENERATION BEYOND 40 Gbit/s

The technique of inline optical regeneration becomes attractive when considering transmissions at very high channel bitrates, 40 Gbit/s and above. In this view, all-optical solutions (combining clock recovery and synchronous modulator) are currently being investigated for the implementation of high-speed SM-based optical regeneration. It is emphasized that the SM technique is well suited for high-bandwidth operation since it only requires a narrowband sinusoidal modulation with fairly moderate depth (e.g., 6 dB for 40 Gbit/s operation).

As a starting point, we have recently proposed and demonstrated a novel BBOR scheme using a 40 Gbit/s SOA-based Mach-Zehnder interferometer (SOA-MZI) [88]. Figure 33 shows the experimental setup for this optically driven synchronous modulator. The principle of operation is similar to that of the InP Mach-Zehnder modulator except that phase shift now results from gain saturation in the SOA located in the two arms of the interferometer. Such an effect is directly controlled by injecting an optical clock signal in one or the two arms (push-pull drive mode) of the SOA-MZI. Experimental loop implementation of this device makes it possible to demonstrate first a loss-free operation of the all-optical regenerator and second an error-free 40 Gbit/s transmission over 10,000 km with $2^{23} - 1$ PRBS sequence. No system penalty was measured when changing PRBS coding sequence length, further confirming the absence of any patterning effect within this new optically controlled BBOR regenerator. Loss-free operation makes it an attractive candidate for long-haul transmission

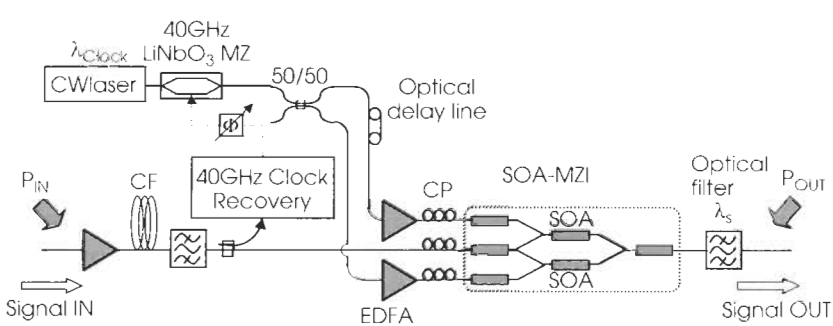


Fig. 33 Experimental setup of the 40 Gbit/s all-optical regenerator based on SOA-MZI-induced synchronous modulation.

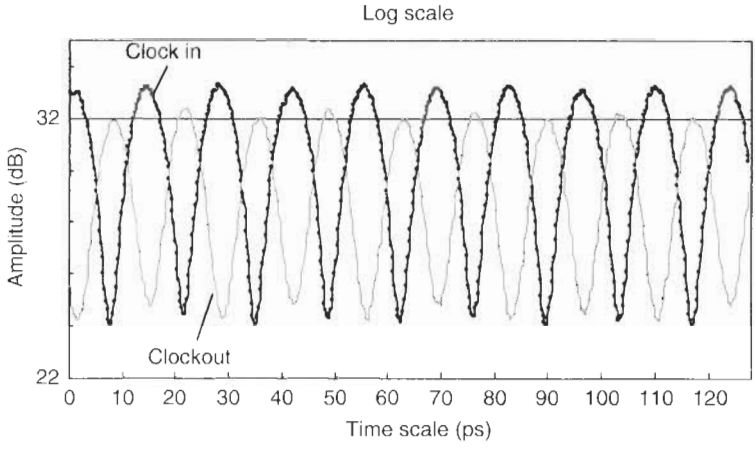


Fig. 34 80 GHz clock conversion through all-optical SOA-MZI.

systems, though WDM operation of the device remains to be fully assessed. Moreover, potential ultrahigh operation of this improved BBOR scheme was recently experimentally demonstrated by means of an 80 GHz clock conversion with appropriate characteristics through the SOA-MZI SM regenerator. Figure 34 shows both the initial (1540 nm) and converted (1552.5 nm) 80 GHz optical clock signals as monitored on a 2 ps resolution streak camera. Extinction ratio of the converted clock (i.e., IM depth) is measured to be 8 dB, which is theoretically suitable for efficient SM-based regeneration. Further work will consist of the experimental implementation of such an 80 Gbit/s optical regenerator.

We should finally mention fiber-based devices such as NOLM [2] and NALM for addressing ultrahigh speed SM-based optical regeneration, though no successful experimental demonstrations have been reported yet in this field. The issue of integrability of such devices also remains to be addressed.

IV. Electrooptical vs. All-Optical 3R regeneration

After having described the current status of all-optical regeneration, two key questions remain. First, is optical 3R regeneration the only solution for capacity enhancement or extension of transmission distance? Second, at which data rate and WDM granularity would signal regeneration be required and in what system or network type?

If complexity increase was allowed in future system evolutions (overlooking cost-effectiveness aspects), one should also consider classical optoelectronic 3R regeneration, which has been used in early system implementations. In such electronic repeater/regenerators (see Fig. 35a), the incoming optical signal

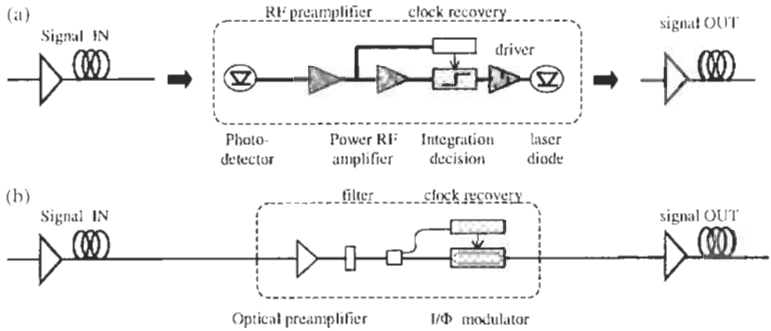


Fig. 35 Basic setup of electrooptical (a) and all-optical regenerator (b).

is converted in the electrical domain by means of a high-speed photodiode. Signal regeneration/processing is then achieved using broadband RF circuitry, after which the regenerated electrical signal is converted back in the optical domain using an electrooptical modulator coupled with a laser diode. As seen, complete integration of such an architecture is rather complex. On the other hand, both optical 3R regenerator layouts are significantly simpler, as shown in Fig. 35b.

As a tentative answer to the above questions, let us draw a parallel between the two alternative approaches in transmission applications. For a 10 Gbit/s line rate, the best candidate for signal regeneration is with no doubt the optoelectronic solution. Indeed, it clearly provides the best regenerative performance, while taking benefit of the maturity of electronic technology, leading to both high-level integration and very low cost (compared to optical technologies). The picture changes somewhat when considering 40 Gbit/s (and higher) applications, for which neither solutions has the lead. On one hand, the electrooptical solution could benefit from the technological background of electronics, leading to possible high-level integration and reduced costs *but only* if satisfactory broadband adaptation of both chips and modules could be achieved. Indeed, parasitic effects such as resonance coupling in housing and group velocity dispersion or conductor losses can no longer be neglected at such high frequencies. Moreover, electrooptical components such as $LiNbO_3$ modulators used for data coding are not mature yet. It should be also mentioned that no electrooptical regenerator has been experimentally tested yet at 40 Gbit/s line rate to our best knowledge. Finally, implementation of electrooptical regeneration in transmission systems is only possible using a channel-by-channel parallel scheme. This amounts to building into the 3R repeater two complete WDM terminals, as assembled in a back-to-back configuration. For long-haul systems in both submarine and terrestrial configurations, such WDM terminal concatenation throughout the trunk is a technical solution but does not make economic sense if the number of channels

exceeded two or four. On the other hand, the optical regeneration technique has been assessed at 40 Gbit/s (and at even higher rates) in many loop experiments. This solution possibly enables simultaneous WDM regeneration and does not suffer from the electronic bandwidth bottleneck. Indeed, only narrowband RF or simple control electronics are required, which reduces RF circuitry complexity and power consumption. As for electronics, complete integration (with either optoelectronic or all-optical clock recovery) of parallel optical regenerators on a single semiconductor chip is technologically possible with optical regeneration. However, the key optical components (SOA-MZI or InP MZ) remain to be further improved with, in particular, the reduction of the residual polarization dependence. The cost of optical packaging is an open issue, though, unlike the case with electronics, it does not depend on the actual operation speed.

Conclusion

We have reviewed the current status of all-optical regeneration, a technique of promising potential in the evolution of high bitrate systems. A variety of solutions exists, as described, including saturable absorbers, nonlinear interferometer/gates, and synchronous modulators. These solutions represent a broad array of possibilities, all having a strong potential for compact integration and all-optical, very-high-bandwidth implementation.

Wherever performance in signal quality, capacity, and transmission distance will reach tangible physical limits, as we are already witnessing, inline regeneration should represent a key alternative for further progress and development. But to reach practical implementation, beyond mere lab demonstrations, several issues remain to be addressed. Some are technical, the others economical.

On the technical side, a workable solution for optical regeneration should include all aspects of system engineering, starting from power consumption issues to packaging and component reuse issues. A 3R regenerator is in itself a subsystem, which implies a high level of reliability, with enhanced supervision and monitoring requirements. As we have described, the regenerator operation should ideally be independent of the type of transmission line and data format. Conceptually, regenerators should also be bitrate transparent and accept any number of WDM channels. This points to the issue of intrinsic regenerator bandwidth (the ultimate rate at which data can be processed) and scalability (the number of WDM channels that can be processed). As highlighted in this chapter, all-optical regeneration offers a unique possibility for simultaneous WDM-channel regeneration, as opposed to the classical, optoelectronic approach. However, this requires WDM channel resynchronization and that all channels have identical transmitter clocks. Assuming these issues to be finally solved, there will remain the problem of power handling capacity

limits of the regenerator. The high powers involved in simultaneous channel regeneration operation could indeed cause accelerated device aging, catastrophic damage, and more basically, signal distorsion and WDM crosstalk. Another attractive and unrivaled application of optical regeneration is single-channel operation at ultrahigh bit rates. Since it is well recognized that all-optical processing has higher bandwidth potential than electronic processing (considering assembled circuits, not only single-transistor gate speeds), it is safe to conclude the same for single-channel regeneration. However, if WDM is excluded, the regenerator must operate at line rates that are competitive with the best systems of the previous generation. The challenges concerning baseband all-optical regeneration at terabit/s rates are manifold. First, clock recovery and driving circuits should be accurate to some 100 fs, and locked to an absolute frequency. Second, the regenerated signal should have a quality at least equivalent to that of the system's transmitter (pulse extinction ratio, intensity noise, and timing jitter), so that regeneration does not introduce its own system impairments. Third, it must work in such a way that transmission distance is effectively unlimited, while the regenerator spacing remains as large as a few kilometers. Thus both simultaneous WDM and single-channel regeneration have their own technical challenges. Clearly, much progress is expected from an intermediate solution combining both characteristics, which simultaneously would regenerate WDM signals at 160 Gbit/s line rate.

On the economic side, the challenges of all-optical regeneration are not specifically different from those concerning new communications technologies: performance enhancements in cost-per-wavelength or cost-per-bandwidth, R&D costs, and time to market. These considerations include significant system reuse, from the fiber-cable plant to the existing terminal platforms and other commercially available subcomponents. One could hardly conceive a commercial success for all-optical regeneration if the technique would require rebuilding systems entirely from scratch. Rather, its best chances remain in the rapid development of 3R units, which could be readily inserted into any kind of existing systems. Regenerators could be placed at every ten amplifiers for instant capacity/performance boost. The cost of such a system upgrade could be competitive with any other solution. The implications of this approach in all-optical networking have been barely considered, probably due to the fact that network nodes already include transceivers and/or regenerators. The economic benefits of all-optical regeneration in the network environment (in contrast with point-to-point, long-haul trunks) remain to be put in perspective.

The main rival of the all-optical regenerator is the classical optoelectronic repeater. The impact of optical amplification in the submarine cable industry and in terrestrial systems has seemingly put electronic regeneration in the backstage. Yet high-speed electronics is rapidly progressing, owing to future demands for 40 Gbit/s systems, operating at the next network/protocol layer. It is then relevant to consider electronic regeneration. As conservative as the approach may sound, the design and implementation of a complete 40 Gbit/s

electronic 3R repeater remains a formidable challenge. This challenge is at least as great as in the all-optical case. The key challenge lies primarily in the integration of the component subblocks, from the front-end receiver to the decision circuit and then to the posttransmitter driver/laser. A secondary but serious challenge is in the fact that some of the processing functions (e.g., error-correction coding) must operate at subharmonic rates, which involves complex MUX/DMUX stages. Finally, there remains for WDM applications a serious cost issue caused by the need of extensive parallel repeater implementation.

The above considerations show that the old debate "electronic vs. all-optical implementation" is still most relevant and inescapable when considering the future of high-speed communications technologies.

From today's status concerning the two alternative approaches for inline regeneration, it is safe to say that the choice between solutions will be dictated primarily by both engineering and economic considerations. It will result from a tradeoff between overall system performance, system complexity and reliability, availability, time-to-market, and rapid returns from the technology investment. The choice in favor of one approach or another may not necessarily be final or be generalizable to all types of communication systems, unless a major technology breakthrough comes into the picture.

Acknowledgments

The authors gratefully acknowledge P. Brindel, B. Dany, P. Guerber, J.-P. Hamaide, A. Jourdan, and D. Rouvillain from Alcatel Research & Innovation; F. Blache, R. Brenot, B. Dagens, F. Devaux, C. Duchet, E. Grard, A. Labrousse, M. Le Pallec, M. Renaud, and A. Shen from OPTO+; and G. Aubin, H. Choumane, J. Mangeney, and J.-L. Oudar from CNRS-LPN for fruitful discussions and for several key contributions to the device and system work described in this chapter.

Abbreviations

APC	Automatic power control
ASE	Amplified spontaneous emission
BBOR	Black box optical regenerator
BER	Bit error rate
CRZ	Chirped return-to-zero
CW	Continuous wave
DFB	Distributed feedback
DSF	Dispersion-shifted fiber
DCF	Dispersion compensating fiber

DM	Dispersion management
DOM	Differential operation mode
EAM	Electroabsorption modulator
ECC	Error-correcting code
EDFA	Erbium-doped fiber amplifier
ESM	Electron scanning microscope
FEC	Forward error correction
FWM	Four-wave mixing
GC-SOA	Gain-clamped SOA
IM	Intensity modulation
IPM	Inverting polarity mode
LTG	Low temperature grown
MI	Michelson interferometer
MLLD	Mode-locked laser diode
MQW	Multiquantum well
MZI	Mach-Zehnder interferometer
NF	Narrowband filtering
NIPM	Noninverting polarity mode
NLSE	Nonlinear Schrödinger equation
NALM	Nonlinear amplifying loop mirror
NOLM	Nonlinear optical loop mirror
NRZ	Nonreturn-to-zero
O/E	Optoelectronic
O/O	Optical/optical
OSNR	Optical signal-to-noise ratio
PDF	Probability density function
PM	Phase modulation
PMD	Polarization mode dispersion
PMF	Polarization-maintaining fiber
PRBS	Pseudo-random bit sequence
RZ	Return-to-zero
SA	Saturable absorber
SM	Synchronous modulation
SMF	Single-mode fiber
SNR	Signal-to-noise ratio
SOA	Semiconductor optical amplifier
SPM	Self-phase modulation
SPL	Self-pulsating laser
SRS	Stimulated Raman scattering
UNI	Ultrafast nonlinear interferometer
WDM	Wavelength-division multiplexing
XGM	Cross-gain modulation
XPM	Cross-phase modulation

References

[1] P. Öhlén, and E. Berglind, "Noise accumulation and BER estimates in concatenated nonlinear optoelectronic repeaters," IEEE Photon. Tech. Lett., Vol. 9, no. 7, pp. 111–222 (1997)

[2] S. Bigo, O. Leclerc, and E. Desurvire, "All-optical fiber signal processing for solitons communications," IEEE Journal of Selected Topics on Quantum Electronics, Vol. 3, no. 5, pp. 1208–1223 (1997)

[3] K.S. Jepsen, A. Buxens, A.T. Clausen, H.N. Poulsen, B. Mikkelsen, and K.E. Stubkjaer, "20 Gbit/s optical 3R regeneration using polarization-independent monolithically integrated Michelson interferometer," Electronics Letters, Vol. 34, no. 5 (1998)

[4] B. Mikkelsen, M. Vaa, H.N. Poulsen, S.L. Danielsen, C. Joergensen, A. Kloch, P.B. Hansen, K.E. Stubkjaer, K. Wünste, K. Daub, E. Lach, G. Laube, W. Idler, M. Schilling, and S. Bouchoule, "40 Gbit/s all-optical wavelength converter and RZ-to-NRZ format adapter realised by monolithic integrated active Michelson interferometer," Electronics Letters, Vol. 33, no. 2 (1997)

[5] D. Chiaroni, B. Lavigne, A. Jourdan, L. Hamon, C. Janz, and M. Renaud, "10 Gbit/s 3R NRZ optical regenerative interface based on semiconductor optical amplifiers for all-optical networks," Post-Deadline paper PD 41 ECOC'97, Edinburgh, 22–25 September (1997)

[6] B. Lavigne, D. Chiaroni, L. Hamon, C. Janz, and A. Jourdan, "Performance and system margins at 10 Gbit/s of an optical repeater for long-haul NRZ transmission," Proc. of European Conference on Optical Communications, ECOC'98, Madrid, 20–24 September (1998)

[7] N.S. Patel, K.A. Rauschenbach, and K.L. Hall, "40 Gbit/s demultiplexing using an ultrafast nonlinear interferometer," IEEE Photon. Tech. Lett., Vol. 8, no. 12, pp. 1695–1697 (1996)

[8] A.E. Kelly, I.D. Phillips, R.J. Manning, A.D. Ellis, D. Nesset, D.G. Moodie, and R. Kashyap, "80 Gbit/s all-optical regenerative wavelength conversion using semiconductor optical amplifier based interferometer." Electronics Letters, Vol. 35, no. 17 (1999)

[9] H.J. Thiele, A.D. Ellis, and I.D. Phillips, "Recirculating loop demonstration all-optical 3R data regeneration using a semiconductor nonlinear interferometer." Electronics Letters, Vol. 35, no. 3 (1999)

[10] A.D. Ellis, A.E. Kelly, D. Nesset, D. Pitcher, D.G. Moodie, and R. Kashyap, "Error-free 100 Gbit/s wavelength conversion using grating assisted cross-gain modulation in 2 mm long seimconductor amplifier," Electronics Letters, Vol. 34, no. 20 (1998)

[11] L. Billès, J.C. Simon, B. Kowalski, M. Henry, G. Michaud, P. Lamouler, and F. Alard, "20 Gbit/s optical 3R regeneration using SOA based Mach-Zehnder interferometer gate," Proc. of European Conference on Optical Communications (ECOC), Edinburgh, 22–25 September (1997)

[12] S. Fischer, M. Dülk, E. Gamper, W. Vogt, E. Gini, H. Melchior, W. Hunziker, D. Nesset, and A.D. Ellis, "Optical 3R regenerator for 40 Gbit/s networks," Electronics Letters, Vol. 35, no. 23 (1999)

[13] B. Lavigne, D. Chiaroni, P. Guerber, L. Hamon, and A. Jourdan, "Improvement of regeneration capabilites in semiconductor optical amplifier-based 3R regenerator," Technical Digest of Conference on Optical Fiber Communications, TuJ3, San Diego, 21–26 February (1999)

[14] B. Lavigne, P. Guerber, C. Janz, A. Jourdan, and M. Renaud, "Full validation of an optical 3R regenerator at 20 Gbit/s," Technical Digest of Conference on Optical Fiber Communications, ThF7, Baltimore, 5–10 March (2000)

[15] S. Watanabe and S. Takeda, "All-optical noise suppression two-stage nonlinear fiber loop interferometers," Electronics Letters, Vol. 36, no. 1 (2000)

[16] S. Watanabe, S. Takeda, K. Hironishi, F. Futami, and T. Chikama, "All-optical 3R regeneration of WDM signals using WDM-OTDM-WDM conversions in two-stage nonlinear fiber loop interferometers," Proc. of European Conference on Optical Communications, Vol. 3, Munich, 3–7 September (2000)

[17] P. Guerber, B. Lavigne, C. Janz, A. Jourdan, D. Wolson, T. Fjelde, and A. Kloch, "Chirp optimised operation of an optical 3R regenerator," Proc. of European Conference on Optical Communications (ECOC), Munich, 3–7 October (2000)

[18] O. Brox, A. Kilk, D. Hoffmann, M. Möhrle, G. Sahin, and B. Sartorius, "Q-switched lasers for wavelength conversion," COST 266/267 Workshop on Optical Signal Processing in Photonic Networks, Berlin, April 5–7 (2000)

[19] M. Owen, M.F.C. Stephens, R.V. Penty, and I.H. White, "All-optical 3R regeneration and format conversion in an integrated SOA/DFB laser," Technical Digest of Conference on Optical Fiber Communications, ThF1, Baltimore, 5–10 March (2000)

[20] M.F.C. Stephens, D. Nesset, K.A. Williams, A.E. Kelly, R.V. Penty, I.H. White, and M.J. Fice, "Wavelength conversion at 40 Gbit/s via cross-gain modulation in distributed feedback laser integrated with semiconductor optical amplifier," Electronics Letters, Vol. 35, no. 20 (1999)

[21] T. Otani, T. Miyazaki, and S. Yamamoto, "40 Gbit/s signal transmission using optical 3R regenerator based on electroabsorption modulators," Technical Digest of Conference on Optical Fiber Communications, ThP3, Baltimore, 5–10 March (2000)

[22] B. Sartorius, M. Möhrle, S. Reichenbacher, H. Preir, H.J. Wünsche, and U. Bandelow, "Dispersive self-Q-switching in self-pulsating DFB lasers," IEEE J. Quantum Electronics, Vol. 33, no. 2 (1997)

[23] C. Bornholdt, B. Sartorius, and M. Möhrle, "All-optical clock recovery at 40 Gbit/s," Post-deadline paper PD 3–5 ECOC'99, Nice, 26–30 September (1999)

[24] D. Chiaroni, B. Lavigne, A. Dupas, P. Guerber, A. Jourdan, C. Bornholdt, S. Bauer, B. Sartorius, and M. Mörhle, "All-optical clock recovery from 10 Gbit/s asynchronous data packets," Proc. of European Conference on Optical Communications (ECOC), Munich, 3–7 October (2000)

[25] B. Sartorius, C. Bornholdt, S. Bauer, M. Möhrle, P. Brindel, and O. Leclerc, "System application of 40 GHz all-optical clock in a 40 Gbit/s optical 3R regenerator," Paper PD11, OFC 2000, Baltimore, 5–10 March (2000)

[26] B. Lavigne, P. Guerber, D. Chiaroni, C. Janz, A. Jourdan, B. Sartorius, C. Bornholdt, and M. Möhrle, "Test at 10 Gbit/s of an optical 3R regenerator using integrated all-optical clock recovery," Proc. of European Conference on Optical Communications (ECOC), Nice, 26–30 September (1999)

[27] H. Yokoyama, Y. Hashimoto, H. Kurita, and I. Ogura, "All-optical sub-harmonic clock recovery and demultiplexing," Technical Digest of Conference on Optical Fiber Communications, ThP5, Baltimore, 5–10 March (2000)

[28] H. Kurita, Y. Hashimoto, I. Ogura, H. Yamada, and H. Yokoyama, "All-optical 3R regeneration based on optical clock recovery with mode-locked LDs," Post-deadline paper PD 3–6 ECOC'99, Nice, 26–30 September (1999)

[29] J.Y. Emery, B. Lavigne, C. Porcheron, C. Janz, F. Dorgeuille, F. Pommereau, F. Gaborit, I. Guillemot-Neubauer, and M. Renaud, "Increased power dynamic range of Mach-Zehnder wavelength converter using a semiconductor optical amplifier power equaliser with 8 dBm output saturation power," Electronics Letters, Vol. 35, no. 12 (1999)

[30] N. Edagawa and M. Suzuki, "Novel all-optical limiter using electroabsorption Modulator," Technical Digest of Conference on Optical Fiber Communications, ThP1, Baltimore, 5–10 March (2000)

[31] Y. Shibata, Y. Yamada, K. Habara, and N. Yoshimoto, "Semiconductor laser diode optical amplifiers/gates in photonic packet switching," J. Lightwave Technol., Vol. 16, no. 12 (1998)

[32] D. Chiaroni, B. Lavigne, A. Jourdan, M. Sotom, L. Hamon, C. Chauzat, J.-C. Jacquinot, A. Barroso, T. Zami, F. Dorgeuille, C. Janz, J.-Y. Emery, E. Grard, and M. Renaud, "Physical and logical validation of a network based on all-optical packet switching systems," J. Lightwave Technol., Vol. 16, no. 12 (1998)

[33] T. Durhuus, B. Mikkelsen, C. Joergensen, S.L. Danielsen, and K.E. Stubkjaer, "All-optical wavelength conversion by semiconductor optical amplifiers," J. Lightwave Technol., Vol. 14, no. 6 (1996)

[34] H.A. Haus and Y. Silberberg, "Theory of mode locking of a laser diode with a multiple-quantum-well structure," J. Opt. Soc. Am. B, Vol. 2, no. 7, pp. 1237–1243 (1985)

[35] F.X. Katner, I.D. Jung, and U. Keller, "Soliton mode-locking with saturable Absorbers," IEEE J. Selected Topics Q. Elec., Vol. 2, no. 3, pp. 540–556 (1996)

[36] R. Takahashi, Y. Kawamura, and H. Iwamura, "1.55 μm-ultrafast surface-reflection all-optical switching using low-temperature-grown Be-doped strained MQWs," Proc. ECOC, pp. 113–115, Firenze, September (1994); R. Takahashi, Y. Kawamura, T. Kagawa, and H. Iwamura, "Ultrafast 1.55 μm photo responses in low-temperature-grown InGaAs/InAlAs quantum wells," Applied Physics Letters, Vol. 65, no. 14, pp. 1790–1792 (1994)

[37] S. Gupta, J.F. Whitaker, and G.A. Mourou, "Ultrafast carrier dynamics in III-V semiconductors grown by molecular-beam epitaxy at very low substrate

temperatures," IEEE J. Quantum Electronics, Vol. 28, no. 10, pp. 2464–2472 (1992)

[38] Z. Bakonyi, G. Onishchukov, C. Knöll, M. Gölles, F. Lederer, and R. Ludwig, "In-line saturable absorber in transmission systems with cascaded semiconductor optical amplifiers," IEEE Photonics Tech. Lett., Vol. 12, no. 5, pp. 570–572 (2000)

[39] E. Lugagne-Delpon, J.-L. Oudar, N. Bouché, R. Raj, A. Shen, N. Stelmakh, and J.-M. Lourtioz, "Ultrafast excitonic saturable absorption in ion-implanted InGaAs/InAlAs multiple quantum wells," Applied Physics Letters, Vol. 72, no. 7, pp. 759–761 (1998)

[40] F.M. Knox, P. Harper, P.N. Kean, N.J. Doran, and I. Bennion, "Low jitter long distance pulse transmission near net fiber dispersion zero wavelength," Electronics Letters, Vol. 31, no. 17, pp. 1467–1468 (1995)

[41] D.S. Govan, N.J. Smith, F.M. Knox, and N.J. Doran, "Stable propagation of solitons of increased energy through the combined action of dispersion management and periodic saturable absorption," J. Opt. Soc. Am. B, Vol. 14, no. 11, pp. 2960–2963 (1997)

[42] P. Harper, I.S. Penketh, S. Alleston, and N.J. Doran, "200000 km, 10 Gbit/s soliton propagation exploiting periodic saturable absorption," Proc. ECOC'98, Madrid, p.107, September 20–24 (1998)

[43] N.J. Doran and D. Wood, "Nonlinear optical loop mirror," Optics Letters, Vol. 13, no. 1, pp. 56–58 (1988)

[44] M.E. Fermann, F. Haberl, M. Hofer, and H. Hochreiter, "Nonlinear amplifying loop mirror," Optics Letters, Vol. 15, no. 13, pp. 752–754 (1990)

[45] A. Hirano, H. Tsuda, K. Hagimoto, R. Takahashi, Y. Kawamura, and H. Iwamura, "10 ps pulse all-optical discrimination using a high speed saturable absorber gate," Electronics Letters, Vol. 31, no. 9, pp. 736–737 (1995)

[46] A. Hirano, H. Kobayashi, H. Tsuda, R. Takahashi, M. Asobe, K. Sato, and K. Hagimoto, "10 Gbit/s RZ discrimination using refined saturable absorber optical gate," Electronics Letters, Vol. 34, no. 2, pp. 198–199 (1998)

[47] Z. Bakonyi, G. Onisshchukov, C. Knöll, M. Gölles, and F. Lederer, "10 Gbit/s RZ transmission over 5000 km with gain-clamped semiconductor optical amplifiers and saturable absorbers," Electronics Letters, Vol. 36, no. 21, p. 1790 (2000)

[48] D. Atkinson, W.H. Loh, V.V. Afanasjev, A.B. Grudinin, A.J. Seeds, and D.N. Payne, "Increased amplifier spacing in soliton system with quantum-well saturable absorbers and spectral filtering," Optics Letters, Vol. 19, no. 19, pp. 1514, 1516 (1994)

[49] M. Matsumoto, H. Ikeda, T. Uda, and A. Hasegawa, "Stable soliton transmission in the system with nonlinear gain," J. Light. Tech., Vol. 13, no. 4, pp. 658–665 (1995)

[50] O. Audouin, E. Pallise, E. Desurvire, and E. Maunand, "Use of inline saturable absorbers in WDM soliton systems," IEEE Photon. Techn. Lett., Vol. 10, pp. 828–829 (1998)

[51] S. Bennett and A.J. Seeds, "Error-free 80 Gbit/s soliton transmission over transoceanic (>8000 km) distance using fast saturable absorbers and dispersion

decreasing fibers," Proc. OFC, San Diego, Vol. 2, pp. 50–52, 22–26 February (1999)

[52] O. Leclerc, G. Aubin, P. Brindel, J. Mangeney, H. Choumane, S. Barré, and J.-L. Oudar, "Demonstration of high robustness to SNR impairment in 20 Gbit/s long-haul transmission using 1.5 μm saturable absorber," Electronics Letters, Vol. 36, no. 23, p. 1944 (2000)

[53] J. Mangeney, G. Aubin, J-L. Oudar, J-C. Harmand, G. Patriache, H. Choumane, N. Stelmakh, and J-M. Lourtioz, "All-optical discrimination at 1.5 μm using ultra-fast saturable absorber vertical cavity device," Electronics Letters, Vol. 36, no. 17, p. 1486 (2000)

[54] T. Akiyama, H. Kuwatsuka, T. Simoyama, Y. Nakata, K. Mukai, M. Suguwara, O. Wada, and H. Ishikawa, "WDM multi-channel nonlinear device utilizing inhomogeneously-broadened gain of self-assembled quantum dots," Paper 9.3.6, Proc. ECOC'00, p. 291, Munich, 1–5 September (2000)

[55] M. Nakazawa, E. Yamada, H. Kubota, and K. Suzuki, "10 Gbit/s soliton data transmission over one million kilometers," Electronics Letters, Vol. 27, no. 14, p. 1270 (1991)

[56] O. Leclerc, E. Desurvire, and O. Audouin, "Assessment of 80 Gbit/s (4 × 20 Gbit/s) regenerated WDM soliton transoceanic transmission," Electronics Letters, Vol. 32, no. 12 (1996)

[57] N.J. Smith, K.J. Blow, W.J. Firth, and K. Smith, "Soliton dynamics in the presence of phase modulators," Optics Comm., no. 102, p. 324 (1993)

[58] P. Brindel, B. Dany, O. Leclerc, and E. Desurvire, "A novel 'black-box' optical regenerator for RZ transmission systems," Electronics Letters, Vol. 35, no. 6, pp. 480–481 (1999)

[59] O. Leclerc et al., "Polarisation-independent InP push-pull Mach-Zehnder modulator for 20 Gbit/s soliton regeneration," Electronics Letters, Vol. 34, no. 10, p. 1011 (1998)

[60] E. Desurvire and O. Leclerc, "All-optical regeneration for global-distance fiber-optic communications," in International Trends in Optics and Photonics, ICOIV, Springer, Berlin (1999)

[61] G. Aubin, T. Montalant, J. Moulu, F. Pirio, J-B. Thomine, and F. Devaux, "40 Gbit/s OTDM soliton transmission over transoceanic distances," Electronics Letters, Vol. 32, no. 24, p. 2188 (1996)

[62] K. Suzuki, H. Kubota, A. Sahara, and M. Nakazawa, "40 Gbit/s single channel optical soliton transmission over 70000 km using inline synchronous modulation and optical filtering," Electronics Letters, Vol. 34, no. 1, p. 98 (1998)

[63] O. Leclerc, P. Brindel, D. Rouvillain, E. Pincemin, B. Dany, E. Desurvire, C. Duchet, E. Boucherez, and S. Bouchoule, "40 Gbit/s polarization-insensitive and wavelength-independent InP Mach-Zehnder modulator for all-optical regeneration," Electronics Letters, Vol. 35, no. 9, pp. 730–731 (1999)

[64] M. Nakazawa, K. Suzuki, and H. Kubota, "160 Gbit/s (80 Gbit/s × 2 channels) WDM soliton transmission over 10,000 km using inline synchronous modulation," Electronics Letters, Vol. 35, no. 16, p. 1358 (1999)

[65] M. Nakazawa, K. Suzuki, E. Yamada, H. Kubota, Y. Kimura, and M. Takaya, "Experimental demonstration of soliton data transmission over unlimited distances with soliton control in time and frequency domains," Electronics Letters, Vol. 29, no. 9, p. 729 (1993)

[66] F. Devaux, F. Dorgeuille, A. Ougazzaden, F. Huet, M. Carre, A. Carenco, M. Henry, Y. Sorel, J-F. Kerdiles, and E. Jeanney, "20 Gbit/s operation of a high-efficiency InGaAS/InGaAsP MQW electroabsorption modulator with 1.2 V driving voltage," IEEE Photon. Tech. Lett., Vol. 5, pp. 1288–1290 (1993)

[67] T. Widdowson and A.D. Ellis, "20 Gbit/s soliton transmission over 125 Mm," Electronics Letters, Vol. 30, no. 22, p. 1866 (1994)

[68] G. Aubin, E. Jeanney, T. Montalant, J. Moulu, F. Pirio, J-B. Thomine, and F. Devaux, "20 Gbit/s soliton transmission over transoceanic distances with a 105 km amplifier span," Electronics Letters, Vol. 31, no. 13, p. 1079 (1995)

[69] V.I. Karpman and E.M. Maslov, "Perturbation theory for solitons," JETP, Vol. 46, p. 281 (1977)

[70] D.J. Kaup, "Perturbation theory for solitons in optical fibers," Phys. Review. A, Vol. 42, p. 5689 (1990)

[71] T. Georges, "Perturbation theory for the assessment of soliton transmission control," Optical Fiber Technology, Vol. 1, no. 2, p. 97 (1995)

[72] O. Audouin et al., "Introduction of electrostriction-induced acoustic interaction in modeling of high capacity long haul transmission systems," IEEE Journal of Selected Topics on Quantum Electronics, Vol. 6, no. 2, pp. 297–307 (2000)

[73] A. Sahara et al., "Ultra-high speed soliton transmission in presence of polarization-mode dispersion using inline synchronous modulation," Electronics Letters, Vol. 35, no. 1, p. 76 (1999)

[74] O. Leclerc, E. Desurvire, P. Brindel, and E. Brun-Maunand, "Analysis of timing jitter statistics and bit-error rate in regenerated soliton systems," Optical Fiber Technology, Vol. 5, no. 3, pp. 301–304 (1999)

[75] E. Desurvire and O. Leclerc, "Fundamental bit-error rate limits of inline optical regeneration by synchronous intensity modulation: an analytical Gaussian-statistics model," Optical Fiber Technology, Vol. 6, no. 3, pp. 230–264 (2000)

[76] O. Leclerc and E. Desurvire, "Investigation into the robustness of 100 Gbit/s (5×20 Gbit/s) regenerated WDM soliton transoceanic transmission to line breaks and repairs," Electronics Letters, Vol. 33, no. 18, p. 1568 (1997)

[77] B. Dany, P. Brindel, E. Pincemin, D. Rouvillain, and O. Leclerc, "Recovered efficiency of filters control on dispersion-managed solitons for optical regeneration applications: analysis and experimental validation," Optics Letters, Vol. 25, no. 11, pp. 793–795 (2000)

[78] P. Brindel, O. Leclerc, D. Rouvillain, B. Dany, E. Desurvire, and P. Nouchi, "Experimental demonstration of new regeneration scheme for 40 Gbit/s dispersion-managed long-haul transmission," Electronics Letters, Vol. 36, no. 1, pp. 61–62 (2000)

[79] B. Dany, O. Leclerc, F. Neddam, and P. Le Lourec, "Optimization of dispersion maps for N × 40 Gbit/s transmission with up to 0.4 bit/s/Hz spectral efficiency," Paper TuN5, OFC'01, Anaheim, March (2001)

[80] O. Leclerc, P. Brindel, B. Dany, F. Devaux, C. Duchet, L. Fleury, P. Nouchi, D. Rouvillain, and A. Shen, "Transoceanic N × 40 Gbit/s regenerated WDM systems," Proc. SubOptic Conference, Kyoto, May 19–24 (2001)

[81] E. Brun-Maunand, P. Brindel, O. Leclerc, F. Pitel, and E. Desurvire, "Parametric study of chromatic dispersion influence in 20 Gbit/s, 20 Mm regenerated soliton systems with up to 140 km amplifier spacing," Electronics Letters, Vol. 32, no. 11, p. 1022 (1996)

[82] G. Aubin, E. Jeanney, T. Montalant, J. Moulu, F. Pirio, J-B. Thomine, F. Devaux, and N. Souli, "Record 20 Gbit/s 200 km repeater span transoceanic soliton transmission using in-line remote pumping," IEEE Photon. Tech. Lett., Vol. 8, no. 9, p. 1267 (1996)

[83] E. Pincemin, O. Leclerc, and E. Desurvire, "Feasibility of 1 Terabit/s (25 × 40 Gbit/s transmission optically-regenerated systems," Optics Letters, Vol. 24, no. 11, pp. 720–722 (1999)

[84] E. Desurvire, O. Leclerc, and O. Audouin, "Synchronous in-line regeneration of wavelength-division multiplexed solitons signals in optical fibers," Optics Letters, Vol. 21, no. 14, pp. 1026–1028 (1996)

[85] O. Leclerc, E. Desurvire, and O. Audouin, "Synchronous WDM soliton regeneration: Towards 80–160 Gbit/s transoceanic systems," Optical Fiber Technology, Vol. 3, no. 2, pp. 97–117 (1997)

[86] B. Dany, P. Brindel, O. Leclerc, and E. Desurvire, "Distributed in-line optical regeneration for massive WDM transmission," Electronics Letters, Vol. 35, no. 18, pp. 1575–1576 (1999)

[87] O. Leclerc, B. Dany, D. Rouvillain, P. Brindel, E. Desurvire, C. Duchet, A. Shen, F. Devaux, A. Coquelin, M. Goix, S. Bouchoule, L. Fleury, and P. Nouchi, "Simultaneously regenerated 4 × 40 Gbit/s dense WDM transmission over 10,000 km using a single 40 GHz InP Mach-Zehnder modulator," Electronics Letters, Vol. 36, no. 18 (2000)

[88] O. Leclerc, P. Brindel, D. Rouvillain, B. Dany, R. Brenot, A. Labrousse, and B. Dagens, "Regenerated 40 Gbit/s long-haul transmission using all-optical MZ-SOA as loss-free synchronous modulator," Paper WF6, OFC'01, Anaheim, March (2001)

Chapter 16 | High Bit-Rate Receivers, Transmitters, and Electronics

Bryon L. Kasper

Agere Systems, Advanced Development Group, Irwindale, California

Osamu Mizuhara

Agere Systems, Optical Systems Research, Breinigsville, Pennsylvania

Young-Kai Chen

Lucent Technologies, High Speed Electronics Research, Murray Hill, New Jersey

1. High Bit-Rate Receivers

1.1 INTRODUCTION

For transmission speeds up to 10 Gb/s, lightwave receivers have deviated little from the standard design approach of a photodiode followed by a transimpedance amplifier. Evolutionary improvements in operating speed have come through reduction in device dimensions to improve transit time and reduce capacitance, and through the use of electronics fabricated in materials systems with higher electron mobility (silicon-germanium heterojunction bipolar transistors, AlGaAs HBTs, InGaP HBTs, AlGaAs high-electron-mobility transistors, and InGaAs/GaAs pseudomorphic high-electron-mobility transistors). Vertically-illuminated InGaAs p-i-n or avalanche photodiode detectors have dominated long wavelength applications, and Si or GaAs p-i-n photodiodes have been used at short wavelengths. Ge APDs have been useful below 1 Gb/s at 1310 nm, and GaAs metal-semiconductor-metal photodiodes have made commercial deployment of optoelectronic integrated circuits (OEICs) a reality for short-wavelength datacom applications.

The advent of optical amplifiers, particularly erbium-doped fiber amplifiers (EDFAs), has revolutionized lightwave systems by making dense-wavelength-division multiplexing (DWDM) the cheapest long-haul transmission technology. Optical preamplifiers using EDFAs have been shown to provide very high receiver sensitivity and are employed extensively in WDM applications, but are not commonly used in more cost-sensitive single-wavelength or shorter-reach systems.

Above 10 Gb/s, the landscape begins to change rapidly. Conventional vertical p-i-n photodiodes lose efficiency, and avalanche photodiodes have

OPTICAL FIBER TELECOMMUNICATIONS,
VOLUME IVA

ISBN: 0-12-395172-0

inadequate bandwidth. New design concepts are required to allow 40 Gb/s+ systems of the future to come to fruition.

The receiver section of this chapter will focus on photodetectors and receivers for bit rates of 10 Gb/s and above. Design concepts and demonstrated performance from 10 Gb/s to 100 Gb/s will be described, and fundamental limits to device operation will be presented.

1.2 ULTRAWIDE-BANDWIDTH PHOTODETECTORS

As described in Kato (1999), there have been two major trends in the improvement of high-frequency photodetectors: an increase in bandwidth-efficiency product and an increase in saturation current capability. Bandwidth-efficiency improvements are necessary, because conventional vertically illuminated photodiodes lose efficiency as bandwidth is increased above 10 GHz. Saturation current improvements are desirable because wideband optical amplification is often used in place of bandwidth-limited electrical amplification. Optical-to-electrical conversion then occurs with much higher optical powers incident on the detector, which can cause space-charge induced overload of conventional photodiodes.

1.2.1 Bandwidth Efficiency Tradeoff for Vertically Illuminated Photodiodes

The structure of a typical vertically illuminated photodiode is shown in Fig. 1.1. Absorption occurs in the low-doped or intrinsic InGaAs absorption layer, which is sandwiched between a highly doped n^+ InP layer and a highly doped p^+ region. Electron hole pairs created when photons are absorbed are separated by a high electric field in the n^- layer and drift across this depletion region to be collected at the n^+ and p^+ regions.

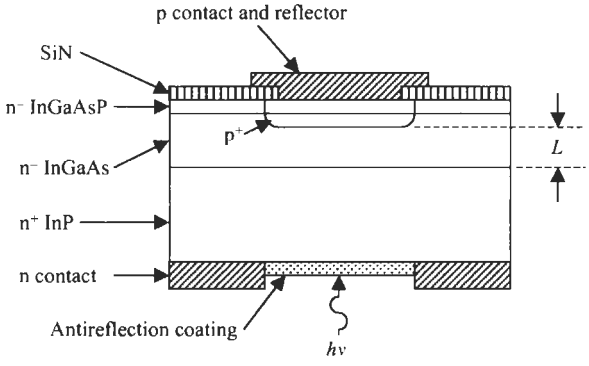


Fig. 1.1 Vertically illuminated p-i-n photodiode.

The general frequency response of a vertically illuminated photodiode has a complex functional dependence (Bowers et al., 1985). However, for the case of a very high speed photodiode with a thin absorbing region ($\alpha L \ll 1$), the transit-time-limited 3-db-down electrical bandwidth f_t is approximately given (Bowers and Burrus, 1987; Alexander, 1997; Kato, 1999) by

$$f_t \approx \frac{0.45v}{d} \tag{1.1}$$

where

v = carrier saturation velocity
d = carrier transit distance
α = optical absorption coefficient
$\alpha = 1.0/\mu m$ for GaAs at 0.85 μm wavelength
$\alpha = 1.16/\mu m$ for InGaAs at 1.3 μm wavelength
$\alpha = 0.68/\mu m$ for InGaAs at 1.55 μm wavelength
L = thickness of photoabsorption layer

Values for the saturated drift velocity v that appear in the literature for InGaAs and/or GaAs range from 7×10^6 cm/s (Bowers et al., 1985) to 5.3×10^6 cm/s (Kato, 1999).

The internal quantum efficiency η_{int} of a vertical photodiode using a double-pass technique where a second pass through the absorbing layer is obtained by using the p-contact electrode as a reflector is given by

$$\eta_{int} = (1 + re^{-\alpha L})(1 - e^{-\alpha L}) \tag{1.2}$$

where

r = reflection coefficient of electrode used as a double-pass reflector.

For the case of a vertical photodiode where $d = L$, and for a very high speed detector where $\alpha L \ll 1$, the bandwidth-efficiency limit for a double-pass vertical photodiode is approximately

$$f_t \cdot \eta_{int} \approx 0.45v(1 + r)\alpha \tag{1.3}$$

Values of these limits for the single-pass case ($r = 0$) and the perfect-reflector double-pass case ($r = 1$) are given in Table 1.1 for some common absorbing materials and wavelengths (Kato, 1999).

In addition to fundamental transit-time and absorption coefficient limits considered above, other factors such as RC time constants and optical coupling losses further limit the attainable bandwidth-efficiency product for vertical photodiodes. The highest reported results range from 20 GHz to 35 GHz (Tucker et al., 1986; Wey et al., 1993; Tan et al., 1995; Wey et al., 1995; Shimizu et al., 1997).

Table 1.1 **Bandwidth-Efficiency Limits for Vertical Photodiodes**

Absorbing Material	*Wavelength*	*Single-Pass Bandwidth-Efficiency Limit*	*Double-Pass Bandwidth-Efficiency Limit*
GaAs	0.85 μm	30 GHz	60 GHz
InGaAs	1.3 μm	34 GHz	68 GHz
InGaAs	1.55 μm	20 GHz	40 GHz

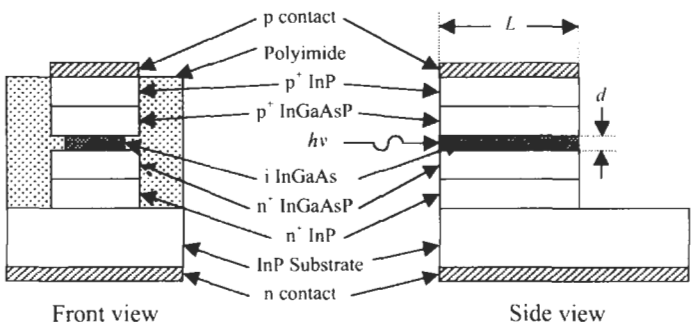


Fig. 1.2 Schematic view of waveguide photodiode structure. Adapted from Kato, 1999.

1.2.2 Waveguide Photodiodes

In order to overcome the transit-time versus absorption efficiency tradeoff inherent in vertical photodiodes, one approach that has been widely studied is the waveguide photodiode shown in Fig. 1.2. In this structure, the photodiode is implemented in a waveguide such that incoming light propagates horizontally and is absorbed over a long distance L in order to obtain high efficiency. Photogenerated carriers, however, are collected vertically across a much shorter distance d in order to obtain short transit time and hence high bandwidth. Transit time and efficiency become decoupled and can be optimized somewhat independently.

The first waveguide photodiode demonstrated had a bandwidth of 28 GHz and efficiency of 25% (Bowers and Burrus, 1986), with the low efficiency resulting from poor optical coupling into the very thin photoabsorption layer that formed the waveguide. A major improvement in coupling efficiency can be obtained by using transparent layers above and below the absorbing layer (as shown in Fig. 1.2) to form a double-core multimode waveguide that can more closely match the optical field distribution of incoming light (Kato et al., 1991; Wake et al., 1991). A second improvement can be made by reduction of the photodiode RC time constant through minimizing waveguide

area and reducing contact resistance using a structure such as a mushroom mesa. With these techniques, a bandwidth of 110 GHz and efficiency of 50% (bandwidth-efficiency product of 55 GHz) have been obtained (Kato et al., 1994).

1.2.3 Distributed Photodetectors

A possibility for further improving speed is to overcome RC bandwidth limitations by making the photodetector capacitance a distributed capacitance along a transmission line. The capacitance becomes incorporated into the transmission line, which when terminated into matching impedances can have very high bandwidth that is relatively independent of the capacitances (Taylor et al., 1990). One realization of this concept is the traveling-wave photodiode (Fig. 1.3), in which photoabsorption occurs in an optical waveguide that continuously parallels an electrical waveguide (transmission line) that collects photogenerated carriers (Goldsmith et al., 1997; Giboney et al., 1992, 1997). A second configuration is the periodic traveling-wave photodiode (Fig. 1.4),

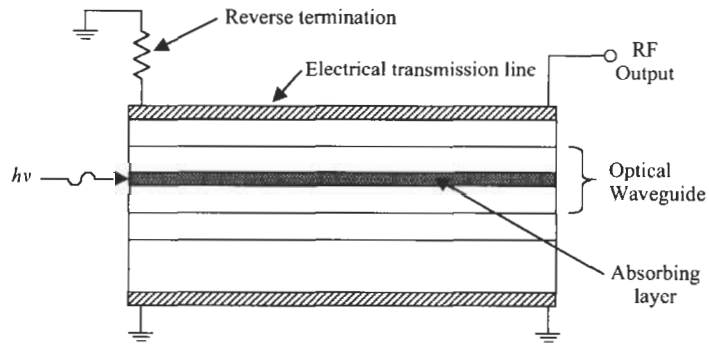


Fig. 1.3 Schematic view of traveling-wave photodiode structure. Adapted from Kato, 1999.

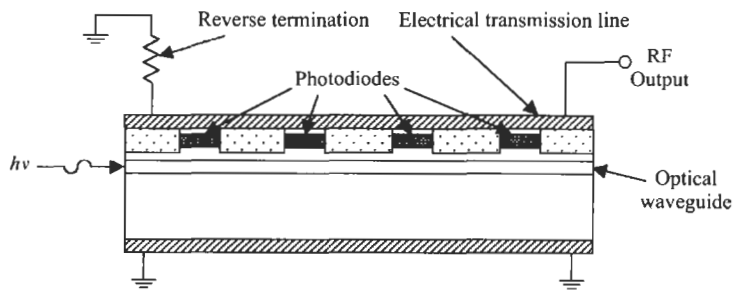


Fig. 1.4 Schematic view of periodic traveling-wave photodiode structure. Adapted from Kato, 1999.

in which absorption occurs in a series of discrete photodiodes distributed periodically along an optical waveguide and connected to a parallel electrical transmission line (Lin et al., 1997a; Giboney et al., 1997). Coupling from the optical waveguide to the photodiodes occurs through the evanescent field that extends outside the waveguide core. The advantage of the periodic traveling-wave photodiode is that the propagation velocities of the optical and electrical waves can be more closely matched by using an electrical slow-wave structure, thereby further increasing the bandwidth capability (Wu and Itoh, 1993).

At short wavelengths, a GaAs/AlGaAs traveling-wave photodiode has achieved 172 GHz bandwidth and 42% efficiency, for a bandwidth-efficiency product of 76 GHz (Giboney et al., 1995).

1.2.4 Resonant Cavity Photodiodes

An alternative approach to improving the efficiency of high-speed photodiodes is to employ a thin photoabsorption layer within an optically resonant cavity. The structure is illustrated in Fig. 1.5, in which the absorbing layer is located between two reflectors. At resonance, incoming photons are reflected back and forth between the two mirrors, making many passes through the absorbing layer, thereby enhancing absorption efficiency.

The bottom mirror typically consists of a quarter-wave-stack dielectric Bragg reflector, similar to that employed in a vertical-cavity surface-emitting laser. The quarter-wave-stack is made from alternating layers such as GaAs/AlAs for short wavelengths (Chin and Chang, 1990; Ünlü et al., 1990), or InP/InGaAsP for long wavelengths (Dentai et al., 1991). Obtaining high reflectivity is more difficult at long wavelengths because the refractive index difference between InP and InGaAsP is at least two times smaller than that between GaAs and AlAs (Campbell, 1995). By using wafer bonding to attach a high-reflectivity GaAs/AlAs mirror to a long-wavelength InGaAs

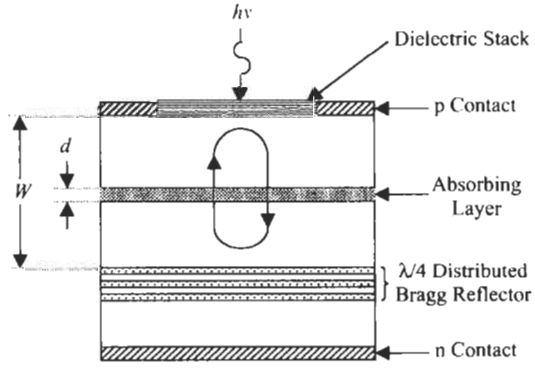


Fig. 1.5 Schematic view of resonant cavity photodiode structure. Reprinted from Campbell, 1995. © 1995 IEEE.

photodiode, Tan et al. (1994) were able to demonstrate an efficiency of 93% (increased by a factor of 16) at 1.3 μm wavelength using a 900-angstrom thick InGaAs absorption layer.

The highest reported bandwidth-efficiency products for resonant-cavity photodiodes have been 17 GHz (17 GHz × 99%) for a short-wavelength GaAs p-i-n (Barron et al., 1994) and 20 GHz (100 GHz × 20%) for a short-wavelength Schottky photodiode (Onat et al., 1998). The technique has also been used at short wavelengths with a GeSi/Si structure to achieve a bandwidth of 15 GHz and efficiency of 67% (Murtaza et al., 1996a). This bandwidth represents a 10× improvement over conventional Si photodiodes.

Because they are resonant structures, the operating wavelength range of resonant-cavity photodiodes is narrower than that of conventional detectors. In fact, the original work focused on increasing wavelength selectivity and efficiency rather than bandwidth. Resonance and hence high responsivity is achieved at wavelengths for which the round-trip phase shift is a multiple of 2π, and as is typical of Fabry–Perot cavities, the spacing between resonances is a function of the cavity length. An expression for the external quantum efficiency η is given (Kishino et al., 1991) by

$$\eta = \left\{ \frac{[1 + R_2 e^{-\alpha d}]}{1 - 2\sqrt{R_1 R_2} e^{-\alpha d} \cos(4\pi n W/\lambda + \phi_1 + \phi_2) + R_1 R_2 e^{-2\alpha d}} \right\} \times (1 - R_1)(1 - e^{-\alpha d}) \tag{1.4}$$

where

R_1, R_2 = power reflectivity of the two mirrors
W = separation between the mirrors
α = absorption coefficient
d = thickness of absorbing layer
n = refractive index
ϕ_1, ϕ_2 = phase shift experienced at the two mirrors
λ = wavelength

The narrower spectral response of resonant-cavity photodiodes can be a disadvantage in some applications, but it can also be useful in applications such as wavelength division multiplexing. Some reported spectral response widths have included values ranging from 14 nm (Tan et al., 1994) to 1.8 nm (Murtaza et al., 1996b).

1.2.5 Avalanche Photodiodes

Avalanche photodiodes (APDs) have been shown to be useful high-sensitivity detectors at bit rates up to 10 Gb/s. For long-wavelength operation, the typical structure involves separate absorption, grading, and multiplication layers to obtain high quantum efficiency, high bandwidth, low noise, and low dark

current (Campbell et al., 1989; Forrest, 1988). For a vertically illuminated APD, the tradeoff between efficiency and transit time in the absorbing layer is similar to that for a p-i-n photodiode. However, APDs have additional layers through which carriers must transit, and during the multiplication process that produces avalanche gain there are more transit time effects added for carriers generated during impact ionization in the multiplication layer. These additional transit delays result in the bandwidth of APDs always being less than the bandwidth of p-i-n photodiodes with comparable photoabsorption layer thicknesses, and thereby limit the bit rates at which APDs can be employed.

Vertically illuminated APDs have been reported with bandwidths up to 17 GHz (Kagawa et al., 1993), and with gain-bandwidth products of 140 GHz (Clark et al., 1999) or 150 GHz (Makita et al., 1995). Waveguide-type APDs have also been fabricated to allow the use of thinner photoabsorption layers, and have demonstrated a bandwidth of 20 GHz along with a gain-bandwidth product of 160 GHz (Cohen-Johnathan et al., 1997), as well as a bandwidth of 32 GHz along with a gain-bandwidth product of 180 GHz (Nakata et al., 2000). The highest gain-bandwidth products have been obtained using resonant-cavity structures which have given a gain-bandwidth product of 290 GHz combined with a bandwidth of 33 GHz at short wavelengths (Nie et al., 1998), and a gain-bandwidth product of 290 GHz combined with a bandwidth of 24 GHz at long wavelengths (Lenox et al., 1999).

1.3 HIGH SATURATION CURRENT PHOTODIODES

With the progress that has occurred in EDFAs and associated high-power pump lasers, it has become common to replace electronic gain with optical gain, particularly in multiple wavelength applications such as DWDM. Also, as electronic amplifiers with the necessary bandwidth to operate at speeds of 40 Gb/s or higher are not readily available, it is often preferable to employ the extremely wide bandwidth possible with optical amplification in place of lower-bandwidth electronic gain. At some point, however, optical-to-electrical conversion is necessary, and with optical amplifiers the optical power level at which O/E conversion takes place is typically much higher than it was in the past. It is readily possible to generate optical signal powers of hundreds of milliwatts, which would saturate or even damage most older-generation detectors. It has therefore become essential to improve the saturation current capability of high-speed photodiodes for use in optically amplified applications.

1.3.1 Design Considerations

The two factors that determine the photocurrent capability of a detector are space-charge effects and thermal considerations. Space-charge effects, or electric field screening, produce nonlinear response because of a reduction in the internal electric field in the photodiode, accompanied by a lowering of carrier velocity. Thermal difficulties can arise with high photocurrents and associated

Joule heating, causing internal photodiode temperature and dark current to increase to the point of thermal runaway and device failure (Paslaski et al., 1996).

Williams and Esman (1999) describe a number of factors that can improve saturation current capability, including minimization of optical power density, optimization of illumination direction, and minimization of intrinsic region doping. Reduction in optical power density decreases the space-charge effect by reducing photogenerated carrier density. The space-charge effect is produced when photogenerated carriers are separated by the external reverse-bias-voltage-induced electric field in the photodiode depletion region. Carrier separation creates an opposing space-charge electric field that counteracts the bias-induced field. If the carrier density goes above a critical point, then the overall depletion region field collapses and RF signal compression occurs (Williams et al., 1994). Linearity can be restored by increasing the external bias voltage at the expense of increased power dissipation and heating. It is preferable to decrease the carrier density at any given point by uniformly illuminating the absorbing region as much as possible rather than focusing the light to a small spot. With a Gaussian beam, one cannot achieve uniform illumination while maintaining high quantum efficiency; hence there is a tradeoff between efficiency and linearity. It has been shown (Davis et al., 1996) that a favorable tradeoff can be made for a vertically illuminated detector by overfilling the absorbing area such that 5–10% of the Gaussian beam is beyond the active region. The reduction in peak carrier density that results allows a doubling of the compression current for a given applied bias voltage.

A second factor that affects saturation of vertical photodiodes is the choice of illumination direction. It has been shown (Williams et al., 1998) that because of differences in the distribution of carriers and because of a decrease in hole velocity at low electric fields, it is possible to substantially increase (up to 10 times) compression current by illumination of the p-side of a detector rather than the n-side.

Other factors that affect saturation include intrinsic region doping and thickness. Williams et al. (1999) show that saturation is improved by low doping of the intrinsic region, and also by decreasing absorbing region length.

Waveguide photodiodes would at first glance appear to offer improved space-charge and thermal performance over vertically illuminated detectors because absorption can be distributed over a longer distance. However, increasing the waveguide length reduces the bandwidth; hence there is a tradeoff between bandwidth and saturation (Lin et al., 1997a). In addition, saturation is limited by the higher carrier density near the input end because of exponential photoabsorption decay (Harari et al., 1995). Also, the thermal impedance advantage of a waveguide photodiode can be lost if thick layers of low thermal conductivity material such as InGaAsP are used around the

absorbing region to form a larger multimode waveguide for improved coupling efficiency (Williams et al., 1999).

Traveling-wave photodiodes offer somewhat improved saturation over waveguide photodiodes, because absorption can be distributed over a longer waveguide without RC time-constant limitation being a factor. However, the best combination of efficiency, saturation, and bandwidth is expected using velocity-matched distributed photodetectors (Lin et al., 1997b; Kato, 1999), as absorption can be distributed over the longest distance without suffering from bandwidth loss caused by optical and electrical velocity mismatch.

1.3.2 Uni-Traveling-Carrier Photodiodes

An alternative approach to improving saturation current exploits the high overshoot velocity of electrons to reduce the space-charge effect, and is called the uni-traveling-carrier (UTC) photodiode (Ishibashi et al., 1997). The band diagram of a UTC photodiode is illustrated in Fig. 1.6. The structure is similar to a conventional p-i-n photodiode, except that absorption does not occur in the i-region, but rather in a thin p-type layer next to a wide-bandgap undoped drift layer. When photons are absorbed to create electron-hole pairs in the p-layer, the holes do not travel but instead quickly join other holes that form the majority carrier population. Electrons, blocked above by a conduction-band step from a diffusion-blocking barrier layer, diffuse toward the depletion layer and rapidly drift across it, thereby generating the photocurrent. If the depletion layer is very thin, the electrons will drift at their overshoot velocity (4×10^7 cm/s) which is five times their saturation velocity (Kato, 1999). As a result, the carrier density and associated space-charge effect are reduced by a factor of five relative to a conventional p-i-n detector.

The bandwidth of a UTC photodiode is limited by electron diffusion time in the p-type absorber. This layer is made very thin ($<0.3\,\mu$m) to obtain high

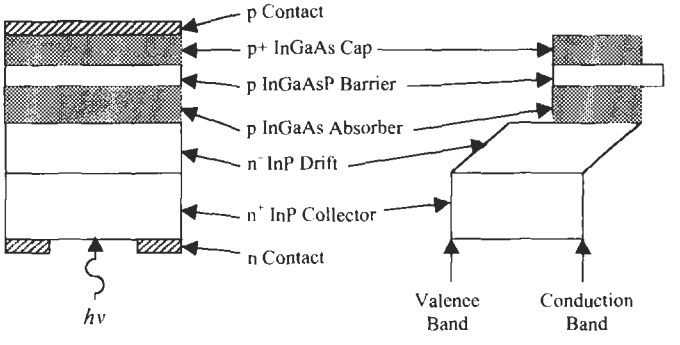


Fig. 1.6 Schematic view of uni-traveling-carrier photodiode structure and energy-band diagram. Adapted from Shimizu et al., 1998.

Table 1.2 **Experimental Saturation Currents of Various Photodiode Structures**

Structure	*Saturation Current*	*Bandwidth*	*Efficiency (Responsivity)*	*Reference*
Vertical p-i-n	15 mA	16 GHz	(0.75 A/W)	Paslaski et al., 1996
Waveguide p-i-n	32 mA	32 GHz	(0.45 A/W)	Welstrand et al., 1997
Evanescently coupled waveguide p-i-n	10 mA	40 GHz	77% (0.96 A/W)	Takeuchi et al., 2001
Distributed GaAs MSM	56 mA	49 GHz	12.3%	Lin et al., 1996
Vertical UTC	76 mA	114 GHz	13%	Shimizu et al., 1997
Waveguide UTC	26 mA	55 GHz	32%	Muramoto et al., 1998

bandwidth, but as a result quantum efficiency is low. Grading of the composition of the p-layer to produce a small conduction-band gradient can speed diffusion and improve the bandwidth to where it is comparable to that of conventional photodiodes (Ishibashi et al., 1997).

The UTC concept has been shown to be very effective in providing high saturation current capability. A vertically illuminated UTC detector with a 0.14-μm-thick graded photoabsorption layer and a 0.2-μm-thick depletion layer demonstrated a bandwidth of 114 GHz, an efficiency of 13%, and a peak photocurrent of 76 mA (Shimizu et al., 1997), allowing an output voltage swing of 1.9 V when operated into a 25 Ω load.

Some of the highest reported saturation currents for various photodiode structures are shown in Table 1.2. It should be noted that there is no standard definition for the saturation current values reported here; hence the comparisons of this parameter may not be exact.

1.4 RECEIVER SENSITIVITY CALCULATIONS

1.4.1 Receiver Sensitivity: p-i-n Detector

Expressions for the sensitivity of an optical receiver using a p-i-n detector are well known and are available from many sources (Smith and Personick, 1980; Kasper, 1988; Alexander, 1997). One form that this expression can take is

$$\overline{P}_S = \left(\frac{hc}{\eta\lambda q}\right) Q(\sigma_{ckt} + qQB_e) \tag{1.5}$$

where

η = detector quantum efficiency
$\overline{P}_S$ = average input optical power required to obtain desired bit-error ratio as specified by signal-to-noise ratio Q
h = Planck's constant $= 6.623 \times 10^{-34}$ joule/sec
c = speed of light $= 3 \times 10^8$ m/sec
λ = wavelength of optical signal
Q = 6.00 for BER $= 1 \times 10^{-9}$
= 7.03 for BER $= 1 \times 10^{-12}$
= 7.94 for BER $= 1 \times 10^{-15}$
σ_{ckt} = rms equivalent input noise current of receiver
q = electronic charge $= 1.6 \times 10^{-19}$ coulombs
B_e = electrical bandwidth of receiver

The second term in Eq. 1.5 represents the shot noise of the received optical signal. If shot noise were the dominant noise source, then the receiver would be operating at or near the quantum limit of 10 photons/bit for a BER of 1×10^{-9} (Kasper, 1988). However, electronic transimpedance amplifiers typically have an equivalent input noise current that is 100 or more times larger; hence the second term in Eq. 1.5 can generally be ignored. The shot-noise term above is included for comparison to the expressions for APD and optically preamplified receiver sensitivity to follow, where shot-noise-related terms become much more significant.

Neglecting shot noise, the sensitivity of a p-i-n receiver can be written as

$$\overline{P}_S\,(\mu\mathrm{W}) = \frac{1.24}{\eta\lambda\,(\mu\mathrm{m})} Q\sigma_{ckt}\,(\mu\mathrm{A_{rms}}) \tag{1.6}$$

where the wavelength has units of μm.

1.4.2 Receiver Sensitivity: APD Receiver

For avalanche photodiode receivers, sensitivity is given by

$$\overline{P}_S = \left(\frac{hc}{\eta\lambda q}\right) Q\left(\frac{\sigma_{ckt}}{M} + qQB_eF(M)\right) \tag{1.7}$$

where

M = average multiplication factor
$F(M)$ = APD excess noise factor.

Comparing Eqs. 1.7 and 1.5, it can be seen that the effect of circuit noise is reduced in inverse proportion to the APD gain M, whereas the effect of signal shot noise is increased by the excess noise factor $F(M)$. Because $F(M)$ increases as M increases, there is an optimum gain value that minimizes $\overline{P}_S$ and provides the best receiver sensitivity. Expressions for the excess noise factor

and optimum gain are well known (Smith and Personick, 1980; Kasper, 1988; Alexander, 1997) and will not be reproduced here. The expression for APD receiver sensitivity in Eq. 1.7 above will be compared to that for an optically preamplified receiver below.

1.4.3 Receiver Sensitivity: Optical Preamplifier

The block diagram of an example erbium-doped fiber optical preamplifier is shown in Fig. 1.7. Parameters that determine the sensitivity of the optically preamplified receiver are indicated in the figure, and are defined as

η_{in} = coupling loss between transmission fiber and erbium-doped fiber

F_m = noise figure of erbium-doped fiber

F_{eff} = effective noise figure referred to amplifier input = F_m/η_{in}

G = gain of erbium-doped fiber

η_{out} = coupling loss between erbium-doped fiber output and detector input including loss of optical filter

G_{tot} = total gain of optical preamplifier from transmission fiber to detector = $\eta_{in}\eta_{out}G$

B_o = bandwidth of optical filter, Hz

B_e = electrical bandwidth of the receiver, Hz

η = detector efficiency

P_S = input optical power

ν = input optical frequency

λ = input optical wavelength

σ_{ckt} = equivalent rms input noise current of p-i-n receiver

The example shows an amplifier configuration in which the optical pump power copropagates with the signal, and leaves out additional elements such as optical isolators that are often employed in actual EDFAs. However, the parameters that determine the sensitivity can readily be generalized to alternative EDFA configurations, and can also be applied to semiconductor optical amplifiers and Raman amplifiers.

1.4.3.1 Photocurrents

Using parameters as defined above, the signal photocurrent is given by

$$I_S = \left(\frac{\eta q}{h\nu}\right) G_{tot} P_S \tag{1.8}$$

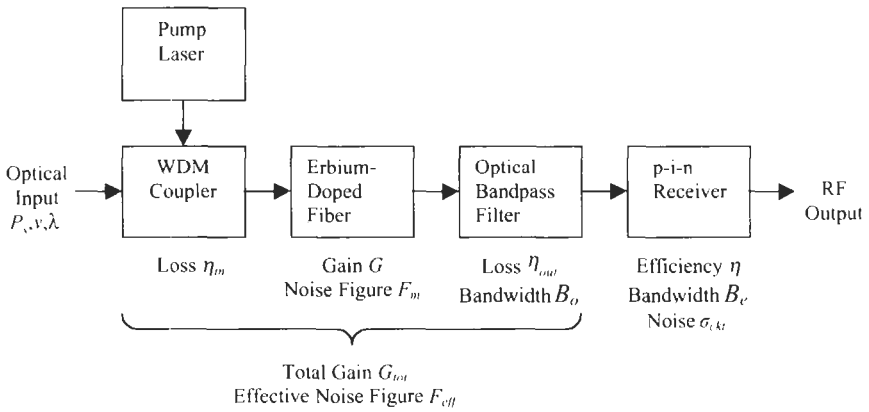


Fig. 1.7 Block diagram of erbium-doped-fiber optically preamplified receiver.

and the photocurrent generated by amplified spontaneous emission is given by

$$I_{ASE} = F_m q \eta_{out} \eta (G - 1) B_o \tag{1.9}$$

1.4.3.2 Noise Sources

A photodiode acts as a square-law detector; hence its output photocurrent contains beat-frequency or difference-frequency components generated from all of the optical frequency signals incident upon it. As such, the photocurrent contains components proportional to signal * signal (the signal photocurrent I_s above), ASE * ASE (also called spontaneous–spontaneous beat noise), and signal * ASE (signal–spontaneous beat noise). In addition to the signal and two noise sources mentioned above, the signal photocurrent and ASE photocurrent produce regular shot noise at the detector output. For the normal case, in which the optical filter bandwidth is greater than the receiver electrical bandwidth ($B_o > B_e$), the noise components contained in the detector photocurrent are as follows (Olsson, 1989; Smith and Kasper, 1997; Alexander, 1997):

$$\sigma^2_{shot} = 2q(I_S + I_{ASE})B_e \tag{1.10}$$

$$\sigma^2_{spon\text{-}spon} = \tfrac{1}{2} F_m^2 q^2 \eta_{out}^2 \eta^2 (G - 1)^2 (2B_o - B_e) B_e \tag{1.11}$$

$$\sigma^2_{sig\text{-}spon} = 2q^2 F_m \eta_{in} \eta_{out}^2 \eta^2 G(G - 1) \left(\frac{1}{h\nu}\right) P_S B_e \tag{1.12}$$

Assuming an ideal system in which the received optical signal power in the 0 state is zero (perfect extinction ratio), using an analysis similar to that in

Smith and Kasper (1997) the average received optical power required for a specified BER in an NRZ system is given by

$$\overline{P}_S = \left(\frac{hc}{\lambda q}\right) Q \left\{ \frac{\sigma_{ckt}^2}{\eta^2 G_{tot}^2} \right. \qquad \text{Circuit noise}$$

$$+ \frac{2q^2 F_{eff}}{\eta G_{tot}} \left(\frac{G-1}{G}\right) B_o B_e \qquad \text{ASE shot noise}$$

$$\left. + \frac{q^2 F_{eff}^2}{2} \left(\frac{G-1}{G}\right)^2 (2B_o - B_e) B_e \right\}^{1/2} \qquad \text{ASE} * \text{ASE beat noise}$$

$$+ \left(\frac{hc}{\lambda q}\right) Q \left\{ \frac{qQB_e}{\eta G_{tot}} \right\} \qquad \text{Signal shot noise}$$

$$+ \left(\frac{hc}{\lambda q}\right) Q \left\{ qQB_e F_{eff} \left(\frac{G-1}{G}\right) \right\} \qquad \text{Signal} * \text{ASE beat noise} \tag{1.13}$$

where the noise source associated with each term is listed at the right.

Simplification can be achieved by noting that the gain of erbium-doped amplifiers is typically more than 30 dB, hence we can assume that both $G \gg 1$ and $G_{tot} \gg 1$. In this case, it can be seen by inspection that ASE shot noise is negligible relative to ASE * ASE beat noise, and that signal shot noise is negligible relative to signal * ASE beat noise. The receiver sensitivity then becomes

$$\overline{P}_S = \left(\frac{hc}{\lambda q}\right) Q \left\{ \frac{\sigma_{ckt}^2}{\eta^2 G_{tot}^2} \right. \qquad \text{Circuit noise}$$

$$\left. + \frac{q^2 F_{eff}^2}{2} (2B_o - B_e) B_e \right\}^{1/2} \qquad \text{ASE} * \text{ASE beat noise}$$

$$+ \left(\frac{hc}{\lambda q}\right) Q \left\{ qQB_e F_{eff} \right\} \qquad \text{Signal} * \text{ASE beat noise} \tag{1.14}$$

Analogies can be drawn between the expression for APD receiver sensitivity and that for optically preamplified receiver sensitivity. Comparing the first term in Eq. 1.7 to the circuit noise term above, it can be seen that in both cases the importance of circuit noise σ_{ckt} is inversely proportional to either M or the gain of the optical preamplifier G_{tot}. Comparing the last term of Eq. 1.7 to the signal * ASE beat noise term above, it can be seen that in each case the noise arises as shot noise multiplied by an excess noise factor which expresses the degree to which the randomness exceeds the basic shot noise of the signal itself. A fundamental difference between an APD and an optical preamplifier

is that the excess noise factor $F(M)$ of an APD is strongly dependent on gain and becomes larger as gain increases, whereas the noise figure of an optical amplifier remains low and is relatively constant up to high values of optical gain. Hence, an optical preamplifier can be used at much higher gain than an APD and can generally obtain higher sensitivity.

1.4.3.3 Optical Preamplifier Sensitivity Limit

If we assume that circuit noise is small enough to be negligible and that the optical filter bandwidth is small enough that spontaneous–spontaneous beat noise can be ignored, then we can find a limiting sensitivity for an optical preamplifier in which signal–spontaneous beat noise is the only noise source. We can also assume no input coupling loss ($\eta_{in} = 1$) minimum possible noise figure from perfect inversion of the gain medium ($F_m = 2$), minimum electrical bandwidth for a given received bitrate B of $B_e = B/2$, and BER $= 1 \times 10^{-9}$ ($Q = 6$). The limiting sensitivity is then found to be

$$\overline{P}_S = 36(h\nu)B \tag{1.15}$$

which corresponds to 36 photons per bit. In reality, the optical filter bandwidth cannot be made less than the modulated optical signal spectral width, so there is always some amount of spontaneous–spontaneous beat noise present. More detailed calculations by a number of workers have arrived at a limiting sensitivity of 38 photons per bit (Henry, 1989; Li and Teich, 1991; Humblet and Azizoglu, 1991).

1.4.3.4 Effect of Optical Filter Bandwidth

Spontaneous–spontaneous beat noise can be minimized by using a narrow optical filter. However, there are practical limits to filter bandwidth that result from variability of transmitter wavelengths and also from the difficulties associated with fabricating very narrow yet stable optical resonators. It is therefore important to understand what bandwidth the optical filter should have to prevent spontaneous–spontaneous beat noise from limiting receiver performance.

If we assume that circuit noise is negligible, then it can be found from Eq. 1.14 that the optical filter bandwidth $B_{o,critical}$ that makes ASE $*$ ASE beat noise equal to signal $*$ ASE beat noise is given by

$$B_{o,critical} = \left(Q^2 + \tfrac{1}{2}\right) B_e \tag{1.16}$$

Expressing the optical bandwidth in terms of wavelength rather than frequency using the relationship

$$B_\lambda = \frac{\lambda^2}{c} B_o \tag{1.17}$$

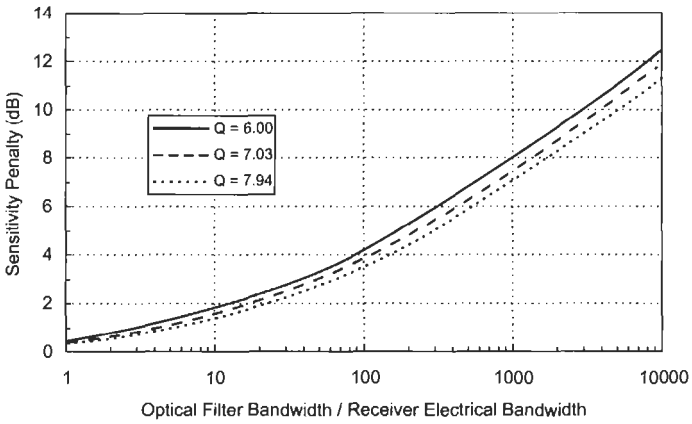


Fig. 1.8 Effect of optical filter bandwidth on optically preamplified receiver sensitivity.

and assuming $\lambda = 1550$ nm, the critical optical filter bandwidth in units of nm is given by

$$\begin{aligned} B_{\lambda,critical}\,(\text{nm}) &= 0.292\,B_e\,(\text{GHz}) \qquad \text{for } Q = 6.00\,(\text{BER} = 1 \times 10^{-9}) \\ &= 0.400\,B_e\,(\text{GHz}) \qquad \text{for } Q = 7.03\,(\text{BER} = 1 \times 10^{-12}) \\ &= 0.509\,B_e\,(\text{GHz}) \qquad \text{for } Q = 7.94\,(\text{BER} = 1 \times 10^{-15}) \end{aligned} \tag{1.18}$$

where B_e is in GHz.

The sensitivity degradation caused by spontaneous–spontaneous beat noise when the optical filter bandwidth is $B_{o,critical}$ is 3 dB. The effect of varying the filter bandwidth above and below $B_{o,critical}$ is shown in Fig. 1.8. The sensitivity penalty in this figure is relative to the ideal value, neglecting spontaneous–spontaneous beat noise, as would be calculated using only the signal $*$ ASE beat noise term in Eq. 1.14.

1.4.3.5 Effect of Circuit Noise

In many calculations of optically preamplified receiver sensitivity, circuit noise is ignored because it is assumed that the amplifier gain is sufficiently large to make its effects negligible. However, as bit rates become very high, electronic amplifier noise can be expected to increase significantly. In addition, photodetector efficiency tends to decrease. From the terms in Eq. 1.14, we can calculate a critical noise level at which circuit noise will be equal to signal–spontaneous beat noise as follows:

$$\sigma_{ckt,critical} = QqB_e\eta G_{tot}F_{eff} \tag{1.19}$$

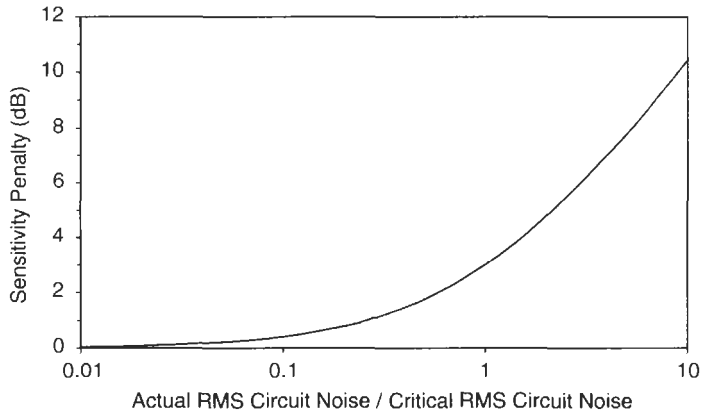


Fig. 1.9 Optical preamplifier sensitivity penalty from circuit noise (neglecting spontaneous–spontaneous beat noise).

At $\sigma_{ckt} = \sigma_{ckt,critical}$, the receiver sensitivity will be reduced by 3 dB from the ideal sensitivity assuming signal–spontaneous beat noise only. Relative to this ideal sensitivity, the effect of circuit noise as σ_{ckt} varies relative to $\sigma_{ckt,critical}$ is given by

$$\frac{\overline{P}_S(\sigma_{ckt})}{\overline{P}_S(\sigma_{ckt}=0)} = \frac{\sigma_{ckt}}{\sigma_{ckt,critical}} + 1 \tag{1.20}$$

and is shown in Fig. 1.9.

From Fig. 1.9, in order for the penalty from circuit noise to be less than 1 dB, the requirement is that σ_{ckt} be less than $0.26\,\sigma_{ckt,critical}$.

1.4.3.6 Effect of Extinction Ratio

If the transmitter extinction ratio is not zero, then in the general case a closed form solution for receiver sensitivity cannot be found and calculation must be done numerically. However, if we make the assumption that signal–spontaneous beat noise is dominant for both ones and zeros, then an expression for the sensitivity penalty as a function of extinction ratio can be found. The result will overestimate the penalty for small values of extinction ratio and as such will represent an upper bound, but will become more accurate as extinction ratio increases.

We define extinction ratio to be

$$r = P_S(0)/P_S(1) \tag{1.21}$$

For NRZ pulse format with 50% mark density the average optical power is then given by

$$\overline{P}_S = P_S(1)\frac{1+r}{2} \tag{1.22}$$

Assuming signal–spontaneous beat noise only, the ratio of receiver sensitivity as a function of r, $\overline{P}_S(r)$, to receiver sensitivity with $r = 0$ can be found to be

$$\frac{\overline{P}_S(r)}{\overline{P}_S(r=0)} = \frac{(1+r)\left(1+\sqrt{r}\right)^2}{(1-r)^2} \tag{1.23}$$

The condition for the above expression is that signal–spontaneous beat noise dominates in the zero state. This can be shown to be true if circuit noise is negligible (i.e., sufficiently high optical gain) and if

$$\frac{r}{1+r} \gg \frac{F_{eff}(2B_o - B_e)}{8(\overline{P}_s/h\nu)} \tag{1.24}$$

For a typical EDFA noise figure of 5 dB ($F_{eff} = 3$) and for typical optically preamplified receiver sensitivities of >100 photons/bit, and also assuming that $2B_o \gg B_e$, then this condition reduces to

$$\frac{r}{1+r} \gg 0.0075\frac{B_o}{B} \tag{1.25}$$

where B is the bitrate. For an optical filter bandwidth of 50 GHz and a bitrate of 10 Gb/s, for example, the approximation assumed in Eq. 1.22 will give a good approximation for $r \geq 0.2$ and may overestimate the penalty for smaller values of r.

A plot of the extinction ratio penalty for an optically preamplified receiver predicted by Eq. 1.22 is shown in Fig. 1.10, along with the extinction ratio penalty for a p-i-n receiver (Smith and Personick, 1980) and an APD receiver (Alexander, 1997) as a comparison. It should be noted that even though the

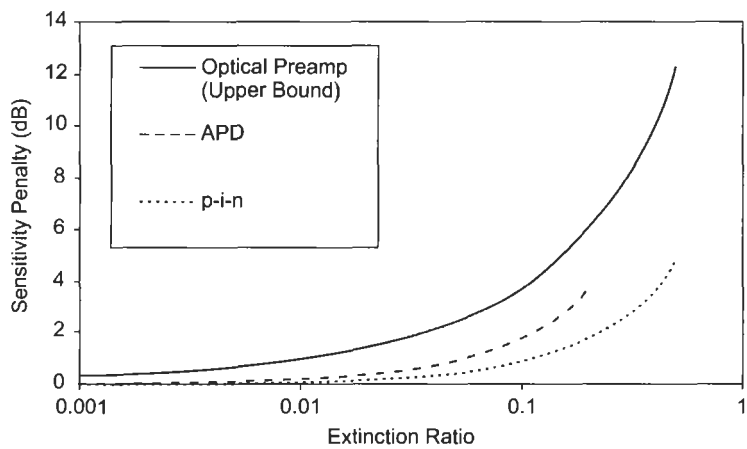


Fig. 1.10 Sensitivity penalty caused by transmitter extinction ratio.

sensitivity penalty from imperfect transmitter extinction ratio for an optically preamplified receiver or an APD receiver will be larger than that for a p-i-n receiver, the actual sensitivity will generally still be significantly better because of the noise advantage provided by the gain of an APD or optical preamplifier.

1.4.3.7 Effect of Intersymbol Interference

One of the factors that can have a significant effect on receiver sensitivity is intersymbol interference (ISI). In general, some amount of ISI is present in every optical transmission system, and causes measured sensitivities to be less than those expected from ideal calculations that assume ISI to be zero. Sources of ISI include nonideal transmitter waveforms, pulse distortion caused by fiber dispersion, and nonideal receiver, postamplifier, and decision circuit pulse response.

Exact modeling of the effects of ISI is mathematically challenging, and as a result calculations of ISI penalties are usually not included in discussions of receiver performance. However, there are some simplified models of ISI that can provide insight into the relative performance of different receiver types in the presence of eye closure.

The simplest and most intuitive measure of ISI uses the concept of eye opening as illustrated in Fig. 1.11. Intersymbol interference caused by pulse distortion produces a peak amount of distortion or eye closure labeled D. Eye opening is a widely accepted criterion of data system performance, and has the advantage that it is not dependent on noise statistics or on the statistical distribution of the data sequence (Lucky, 1965). Calculations of performance penalties based on the peak distortion criterion provide an upper bound on the degradation in BER. Many other methods for calculation of error probability in the presence of ISI are available and provide tighter bounds (Saltzberg, 1968; Lugannani, 1969; Glave, 1972; Matthews, 1973; Shimbo and Celebiler, 1971).

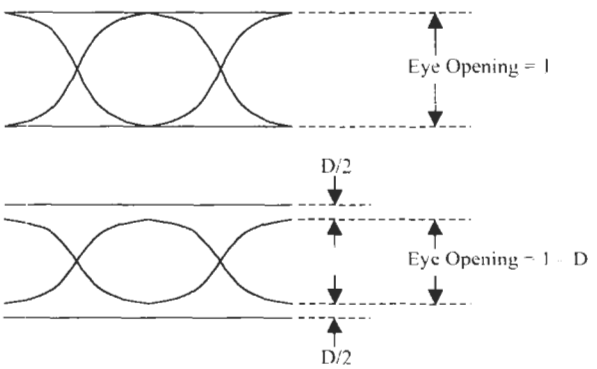


Fig. 1.11 Definition of eye opening.

but we will use the peak distortion criterion to illustrate the relationship between ISI and receiver performance.

Receiver sensitivities in the presence of ISI causing peak distortion D can be calculated by replacing the signal photocurrent amplitude I_s by a reduced effective photocurrent amplitude I_s' given by

$$I_s' = I_s(1 - D) \tag{1.26}$$

In doing so, the expression for p-i-n receiver sensitivity, neglecting shot noise, becomes

$$\overline{P}_S = \left(\frac{hc}{\eta\lambda q}\right) Q \frac{\sigma_{ckt}}{(1 - D)} \tag{1.27}$$

and the sensitivity penalty is then given by

$$\frac{\overline{P}_S(D)}{\overline{P}_S(D = 0)} = \frac{1}{(1 - D)} \tag{1.28}$$

For an optical preamplifier, making the substitution in Eq. 1.26 in the sensitivity calculation can be shown to give the result

$$\overline{P}_S = \left(\frac{hc}{\lambda q}\right) \frac{Q}{(1 - D)} \left\{ \frac{\sigma_{ckt}^2}{\eta^2 G_{tot}^2} + \frac{q^2 F_{eff}^2}{2} (2B_o - B_e)B_e \right\}^{1/2} + \left(\frac{hc}{\lambda q}\right) \frac{Q^2}{(1 - D)^2} \{qB_e F_{eff}\} \tag{1.29}$$

If we assume that signal–spontaneous beat noise dominates, then the sensitivity penalty is given by

$$\frac{\overline{P}_S(D)}{\overline{P}_S(D = 0)} = \frac{1}{(1 - D)^2} \tag{1.30}$$

The sensitivity penalty due to ISI for an optically preamplified receiver, expressed in dB, will be up to twice as large as that for a p-i-n receiver. The penalties found from Eqs. 1.28 and 1.30 are shown graphically as a function of eye closure D in Fig. 1.12. As was the case with extinction ratio previously, it should be noted that although an optically preamplified receiver will have a larger sensitivity penalty from intersymbol interference, it will still generally provide better sensitivity than a p-i-n receiver because of the noise advantage provided by the optical gain.

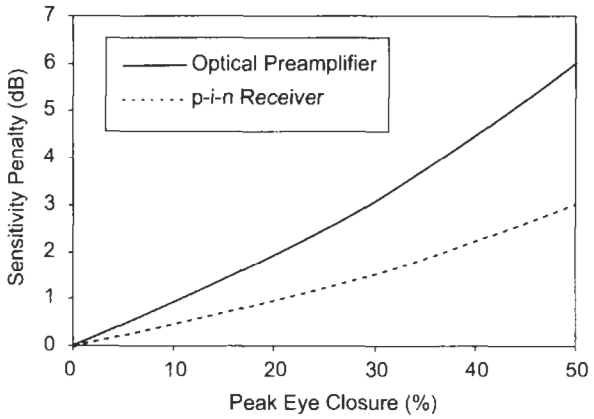


Fig. 1.12 Sensitivity penalty from intersymbol interference.

1.5 EXPERIMENTAL RECEIVER SENSITIVITY RESULTS

Table 1.3 lists a number of experimentally reported receiver sensitivities at bit rates from 5 Gb/s to 100 Gb/s. The same data is shown graphically in Fig. 1.13.

Some comparisons of the sensitivities achieved with different technologies can be made. First of all, the least sensitive results in Fig. 1.13 are for OEIC receivers (optoelectronic integrated circuits). Despite many years of effort to monolithically integrate photodetectors and transistors on the same substrate, sensitivity results at long wavelengths have been inferior to those obtained with hybrid integration of separate devices. In general, the incompatibility of transistor and photodiode structures has resulted in unsatisfactory tradeoffs when monolithic integration has been attempted. The greatest success with commercialization of OEICs has occurred for short-wavelength MSM-MESFET receivers, in which the structure of GaAs MSM (metal-semiconductor-metal) photodiodes is highly compatible with the structure of MESFETs. Such OEIC receivers have become common for short-wavelength, short-reach datacom applications at bit rates up to about 1 Gb/s.

The next receivers listed use PIN detectors and either HBT or HEMT transimpedance amplifiers at 10 Gb/s. As is the case at lower bit rates (Kasper, 1988), FET-based HEMT amplifiers have lower noise at 10 Gb/s than bipolar-based HBT amplifiers, and therefore demonstrate a few dB better sensitivity.

APD-based receivers with either HEMT or HBT transimpedance amplifiers show significantly better sensitivity than PIN-based receivers, with the best APD results being about 6 dB better than the best PIN results.

EDFA optically preamplified receivers have demonstrated the best sensitivities at all bit rates shown. The best result at 5 Gb/s is within 1 dB of the theoretical optical preamplifier limit of 38 photons/bit. The extremely wide

Table 1.3 **Experimental Receiver Sensitivities**

No.	*Bitrate (Gb/s)*	*Sensitivity (dBm)*[1]	*Type of Receiver*	*Reference*
1	10	−16.1	PIN-HEMT OEIC	Takahata et al. (1996)
2	10	−17.3	PIN-HEMT OEIC	Akatsu et al. (1993)
3	10	−19	PIN-HBT	Sieniawski (1998)
4	10	−20.4	PIN-HBT	Yun et al. (1995)
5	10	−22.4	PIN-HEMT	Tzeng et al. (1996)
6	10	−23.5	PIN-HEMT	Yun et al. (1995)
7	10	−26	APD-HEMT	Itzler (2000)
8	10	−27.8	APD	Clark et al. (1999)
9	10	−28.1	APD-HBT	Yamashita et al. (1997)
10	10	−28.7	APD-HEMT	Tzeng et al. (1996)
11	10	−29.2	APD	Nakata et al. (2000)
12	10	−29.4	APD-HEMT	Yun et al. (1996)
13	8	−29.5	SOA[2]	Jopson et al. (1989)
14	10	−33.0	SOA[2], RZ[3]	Smets et al. (1997)
15	10	−36.8	Raman amplifier	Nielsen et al. (1998)
16	10	−37.2	EDFA	Park and Granlund (1994)
17	10	−38.5	EDFA	Nakagawa et al. (1996)
18	10	−40.1	EDFA	Livas (1996)
19	5	−40.5	EDFA	Park and Granlund (1994)
20	5	−45.6	EDFA	Caplan and Atia (2001)
21	20	−29.9	EDFA	Fukuchi et al. (1995)
22	40	−26.6	EDFA	Kuwano et al. (1996)
23	40	−27.7	EDFA	Yonenaga et al. (1998)
24	40	−28.2	EDFA	Ohhira et al. (1998)
25	40	−30.5	EDFA	Ludwig et al. (1997)
26	100	−24.6	EDFA	Takara et al. (1998)

[1] $\bar{P}$ for 1×10^{-9} BER.
[2] Semiconductor optical amplifier.
[3] Return-to-zero pulse format.

bandwidth of optical preamplifiers also makes them capable of operation at bit rates far higher than the maximum of 100 Gb/s shown in Fig. 1.13.

Two results are shown in Fig. 1.13 for semiconductor optical amplifiers (SOA). In general, semiconductor optical amplifiers have higher noise figures and higher input coupling loss than EDFAs, and as a result receiver sensitivities are lower. For example, the SOAs reported in Table 1.3 had noise figures of 7.7 dB and 8 dB, respectively, compared to a typical EDFA effective noise figure of 4–5 dB. The input coupling loss of the SOA in line 12 of Table 1.3 was reported to be 3.5 dB. Despite lower sensitivity, SOAs have some advantages

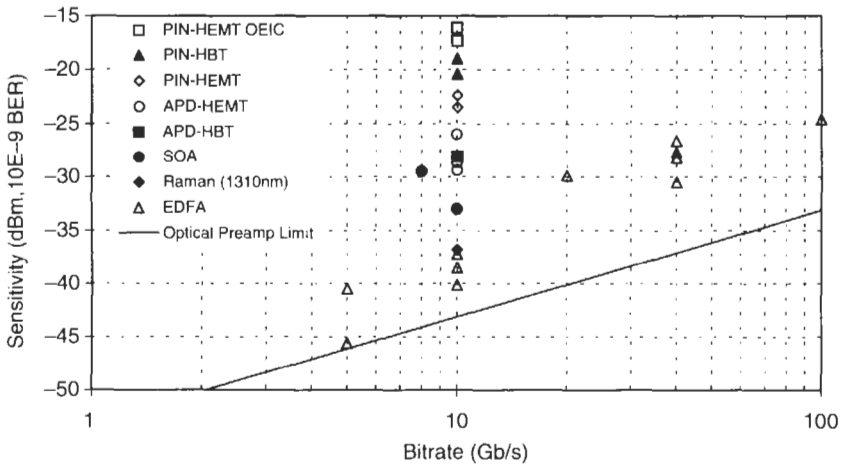


Fig. 1.13 Experimental receiver sensitivity results.

over EDFAs, including operation over a wider range of wavelengths, small size, and potential for monolithic integration that could make them increasingly important in the future.

Finally, there is one result included for a Raman preamplifier. This result at 10 Gb/s is comparable to that for EDFAs, showing the potential that Raman amplification has for low-noise operation. In fact, distributed Raman gain in transmission fiber can be used to significantly increase span lengths and improve system signal-to-noise margins. For a detailed discussion of Raman amplification, the reader is referred to Chapter 5 of this volume.

2. High Bit-Rate Transmitters

2.1 INTRODUCTION

The explosive advancement in optical transmission in recent years has created an unlimited appetite for high-speed, high-capacity systems. In order to address the need, numerous studies have been presented in the industry for achieving 40 Gb/s and beyond, high bit-rate, dense-DWM system capability, exceeding multi-Terabit/s. In addition to long-haul transmission needs, very-short-reach (VSR) and short-haul application needs have been increasing quickly. The need for speed has also increased tremendously in shorter-distance categories, reaching 10 Gb/s and beyond. Typical long-haul requirements, including chirp, extinction ratio, and wavelength stability, are not always considered to be important factors in these applications, but size, power dissipation and cost become the top priorities. In order to meet these very different needs, several different transmitter configurations are

required. In addition to the typical long-haul modulators, including electroabsorption (EA), electroabsorption modulator integrated laser (EML), Mach-Zehnder (MZ), and lithium niobate (LN) types, directly modulated 10 Gb/s Fabry–Perot and distributed feedback (DFB) lasers are being considered for low-cost VSR and short-haul needs.

In this chapter, different style transmitter designs are introduced. They are categorized in terms of features and applications. Transmitter examples are shown, including 10 Gb/s laser transmitters, integrated electroabsorption modulator integrated lasers (EML), and 40 Gb/s optical time-division multiplex (OTDM) and electronic time-division multiplex (ETDM) prototype systems, which consist of transmitters, erbium-doped fiber amplifiers (EDFAs), and receivers. Much attention is paid to stable circuit design. Finally, transmitters with output formats other than NRZ are described. These benchtop research systems have been built to achieve long distance and high spectral efficiency in DWDM applications.

2.2 MODULATION SCHEMES AND APPLICATIONS

Depending on the purpose, optical transmitters can be configured in many ways. When low-cost, low power consumption transmitters are needed but the transmission distance is short, directly modulated laser transmitters are used. If transmitting a long distance such that intersymbol interference (ISI), chirp, and extinction ratio specifications are important, then EA, EML, or LN modulator-based transmitters are used. This section focuses on the application of the transmitters, categorized according to different needs. Among modulation formats, NRZ format has been widely used due to simplicity and low bandwidth requirements.

2.2.1 Directly Modulated Laser Transmitters

This type of the transmitter is widely used in lower-speed applications. Its size, cost, and power dissipation advantages are attractive for low cost and low power dissipation needs. For VSR applications, uncooled Fabry–Perot and DFB lasers can be used to transmit through less than 1 km of single-mode fiber at 1.3 μm and 1.55 μm. Somewhat longer short-haul applications require DFB lasers with good sidemode rejection and low chirp for operation over longer distances (Mizuhara et al., 1995; Adams et al., 1996; Aoki et al., 2000; Ebberg et al., 2000; Massara et al., 1999; Morton et al., 1997; Timofeev et al., 1999). In order to reduce cost, the transmitter is often integrated in the subsystem package.

Much attention has been paid to lowering cost and power dissipation while maintaining good transmission characteristics. Figure 2.1 shows an example of a filtered transmitter output waveform at 10 Gb/s. A clean, open eye with good extinction ratio is obtained as a result of good impedance matching

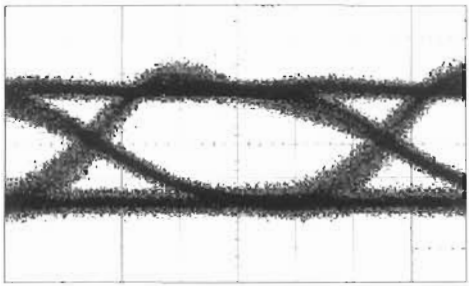


Fig. 2.1 10 Gb/s directly modulated laser transmitter output (H: 20 ps/div).

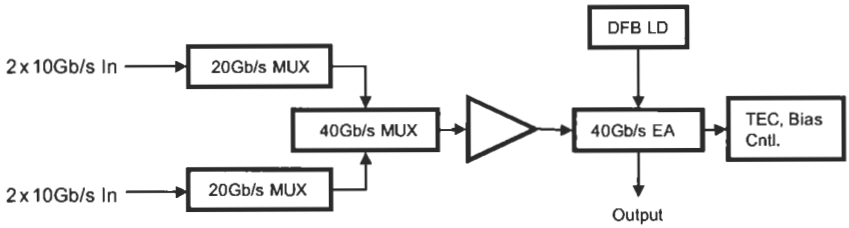


Fig. 2.2 40 Gb/s EA modulator transmitter.

between the laser input and laser-driver output in order to avoid RF reflections, which would result in ISI in the transmitted pulses.

2.2.2 Electroabsorption Modulator Transmitters

EA and EML are widely used in applications that require lower chirp than can be obtained with direct laser modulation. A 10 Gb/s EML module with a modulator-driver IC and control circuits has been developed, and 40 Gb/s EA modulators have been introduced in the market recently (Ishikawa et al., 2000; Shirai et al., 2000). Figure 2.2 shows the block diagram of a prototype 40 Gb/s transmitter employing an EA modulator. The input, 4-channel 10 Gb/s NRZ, is multiplexed using commercially available 20 Gb/s 2:1 MUX and 40 Gb/s 2:1 MUX. The output from the 40 Gb/s MUX is then amplified using a 35 GHz power amplifier and applied to the EA modulator. The output from the amplifier was about 4 Vp-p. The thermoelectric cooler and modulator bias are controlled separately. Figure 2.3 shows the electrooptic frequency response, input impedance match (S_{11}), and DC extinction ratio as a function of bias voltage (Horikawa, 1999).

Figure 2.4 shows the output eye diagram from the EA modulator transmitter at 40 Gb/s. The output power is −3 dBm, and the extinction ratio is 9 dB. It is reasonably open, considering the intrinsic jitter in the sampling scope and the bandwidth limitations in the driver amplifier and 40 Gb/s multiplexer.

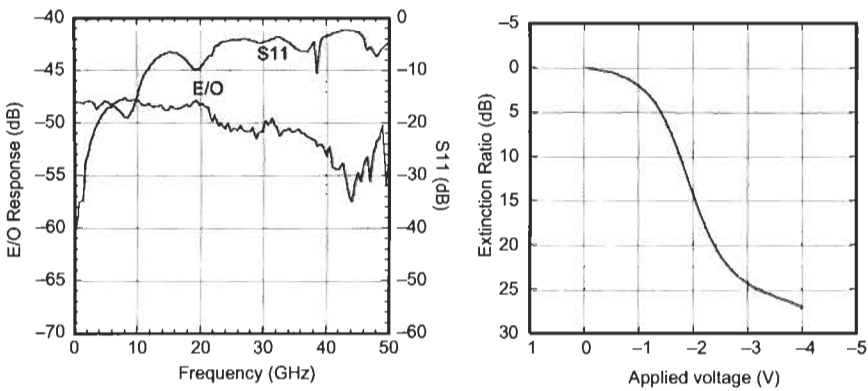


Fig. 2.3 40 Gb/s EA modulator characteristic (Horikawa, 1999).

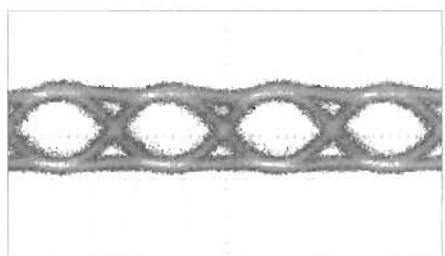


Fig. 2.4 EA transmitter output at 40 Gb/s (H: 10 ps/div).

2.2.3 Lithium Niobate Modulator Transmitters

2.2.3.1 10 Gb/s Lithium Niobate Transmitter

The Ti : $LiNbO_3$ (LN) modulator has been widely used due to availability and low chirp. Bias voltage drift has been an issue, but recent progress in fabrication processes has reduced drift to manageable levels. Figure 2.5 shows a prototype 10 Gb/s LN modulator transmitter output. The transmitter employs data drivers and a dual-electrode LN modulator for zero chirp. For stability over temperature and aging, a microcontroller monitors the LN bias condition and aligns to an optimum point. Figure 2.6 shows simulation results of eye margin at 10^{-9} BER at the receiver with different 10% to 90% rise and fall time (t_r/t_f) on the output waveform (Nuyts, 1997). In this figure, receiver bandwidth was varied, and no receiver noise was considered. A t_r/t_f of <50 ps, preferably <35 ps, is needed for maximum eye opening at 10 Gb/s. Overdriving the LN modulator helps to decrease t_r/t_f at the output, and about 30% more voltage swing than drive voltage at DC is needed for optimum operation.

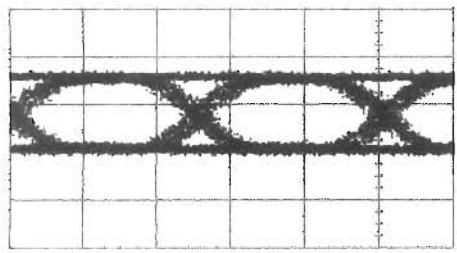


Fig. 2.5 10 Gb/s Ti : $LiNbO_3$ transmitter output (H: 40 ps/div).

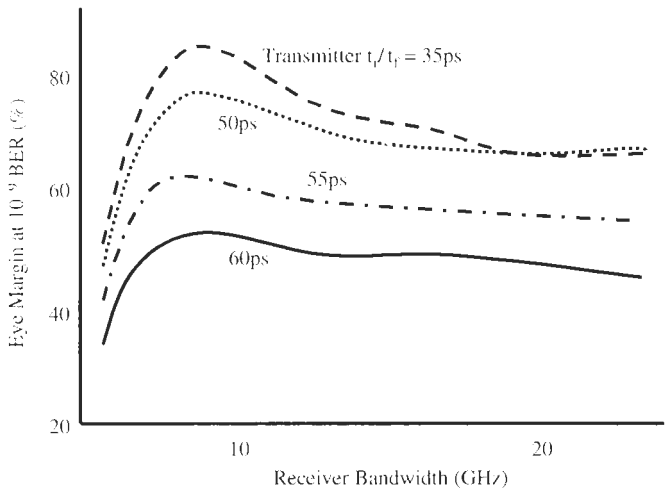


Fig. 2.6 Eye margin simulation with different t_r/t_f on transmitter output waveform (Nuyts, 1997).

2.2.3.2 40 Gb/s Lithium Niobate OTDM System

The following describes a 40 Gb/s LN-modulator-based OTDM prototype system (Chen et al., 1997). Figure 2.7 shows the system block diagram. First, a number of unmodulated CW source lasers producing n optical wavelengths are combined using an arrayed waveguide grating optical multiplexer (OMUX). This multiwavelength optical signal is connected to the first LN modulator, which has polarization maintaining fiber (PMF) at its input and output ports and is driven differentially using commercially available amplifiers with 4 Vp-p NRZ data at 20 Gb/s. In order to maintain high optical signal-to-noise ratio (OSNR), the output from the LN modulator is amplified by a PMF EDFA with +13 dBm output power. This 20 Gb/s NRZ optical signal is then fed to a second LN modulator. It is driven by a differential 20 GHz sine wave, which is generated by a 20 GHz clock limiter and GaAs power amplifier to provide a stable signal. Both LN modulators are controlled by 16-bit

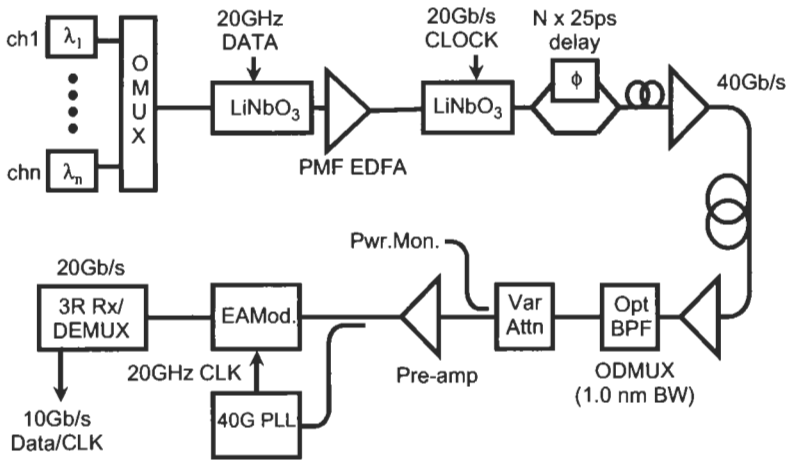


Fig. 2.7 Ti : $LiNbO_3$ modulator-based 40 Gb/s OTDM system.

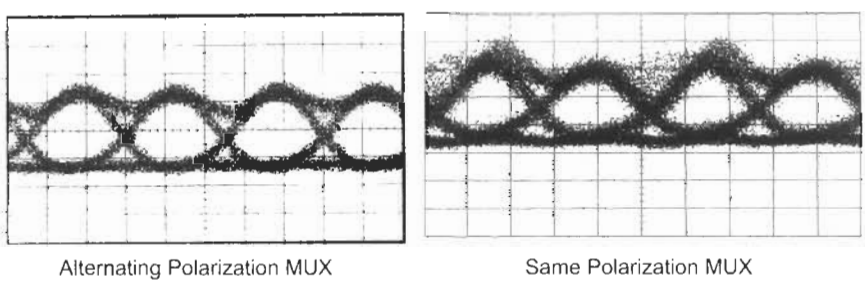


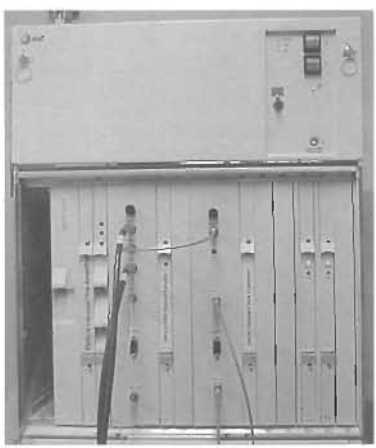
Fig. 2.8 40 Gb/s OTDM transmitter output (H: 10 ps/div).

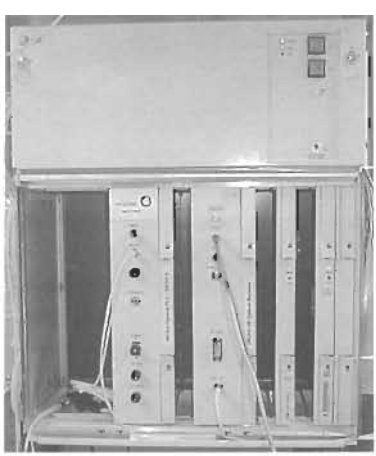
microcontrollers to search for the optimum bias points and to lock onto these points regardless of environmental and power supply voltage changes. The output from the second LN modulator, which is a zero-chirp 50% duty cycle, 20 Gb/s RZ signal, is fed to a second optical MUX via PMF. In the MUX, the signal is split into two arms using a PMF splitter and then recombined in a polarization beam splitter, with one arm incorporating a PMF delay line. The other arm is equipped with a PMF attenuator (not shown) to compensate for the delay line loss. The final output is a wavelength-division multiplexed, 40 Gb/s per channel NRZ signal. A small amount of dispersion added at the output helps decorrelate all the WDM channels.

Figure 2.8 shows the optical output from the transmitter. A very clean, low-jitter waveform is obtained when multiplexed signals are alternatively polarized. If the same polarization is used, coherent interaction is present, degrading the output eye. In order to improve system stability and to increase portability, the transmitter is assembled into a circuit pack configuration.

Figure 2.9 shows a picture of the transmitter shelf. The shelf is powered by a −48 V system power supply, which is connected to an isolated and regulated 120 V AC source.

The receiver, shown schematically in the lower portion of Fig. 2.7 and pictured in Fig. 2.10, consists of an optical bandpass filter, an EDFA preamplifier, an EA-modulator-based 2 : 1 optical demultiplexer with 40 GHz PLL timing recovery, a 20 Gb/s 3R receiver, and 2 : 1 electrical DEMUX. The EA modulator is selected to achieve low (< 1 dB) polarization dependence. The 20 Gb/s 3R receiver consists of front-end, AGC, timing, decision, and DEMUX circuits. All the transmitter and receiver circuits are assembled in microwave packages





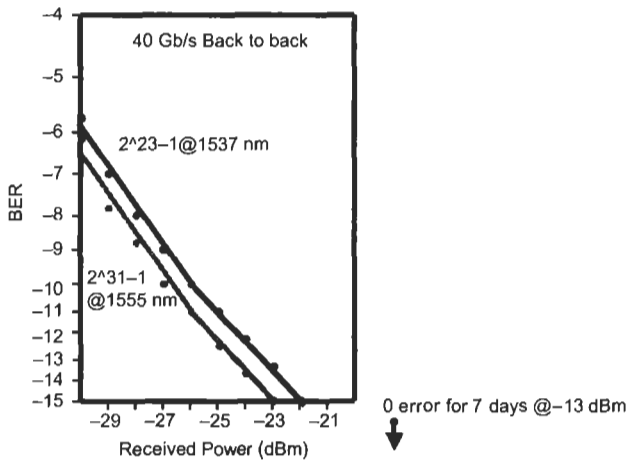


Fig. 2.11 40 Gb/s OTDM system performance.

with coplanar waveguides on dielectric substrates to ensure low-loss, mode-free microwave interconnection. All the power feeds to the ICs are filtered with EMI feedthroughs to ensure stable operation.

Since ICs from different vendors require different power supplies, compact local switching regulators are used, along with power filters to eliminate switching noise. Figure 2.11 shows BER curves with different pseudorandom pattern (PN) word lengths and optical wavelengths. Good sensitivity curves are obtained with small differences among different PN patterns. The slope change at 10^{-9} BER is due to the AGC function in the receiver. The curves show no error floors above 10^{-15} BER.

In order to investigate system stability, the received optical power was increased to −13 dBm and operated to collect errors for a long period of time. As shown in the figure, zero errors were obtained over 7 days of continuous measurement, resulting in $< 5 \times 10^{-16}$ BER.

2.2.4 Other Modulation Schemes

In this section, transmitters using coding schemes other than NRZ are described. These new schemes have not yet been commercially deployed, and most of them are in benchtop configurations in research laboratories. They are only briefly described since more complete descriptions of similar systems are found in other chapters.

2.2.4.1 Return-to-Zero

Although this scheme exhibits poorer spectral efficiency in DWDM applications due to its wider optical spectrum, it has been proven to be suitable for extending 40 Gb/s transmission length (Feiste et al., 2000). A narrow optical

pulse generator such as a mode-locked laser or an EA modulator is required to create an RZ pulse train, followed by a 40 Gb/s modulator for data encoding. Figure 2.12 shows an example eye pattern obtained using EA modulators for RZ shaping and data encoding. Due to bandwidth limitations in the sampling scope and O/E converter, the pulse shown does not necessarily represent the actual pulse shape. Narrower pulse width can be obtained by biasing the EA modulator deeper, then modulating the device with a larger RF signal. Recently, an integrated 40 Gb/s EA modulator plus semiconductor optical amplifier and RZ encoder has been reported (Oggazzaden et al., 2001) with successful transmission experiments over 100 km of Truewave fiber.

2.2.4.2 Chirped RZ

Chirped RZ (CRZ) has been shown in transmission studies to be more tolerant of fiber dispersion than regular RZ (Bergano et al., 1997) and can achieve longer transmission distance (Bakhshi et al., 2001). Figure 2.13 shows the basic block diagram. The laser source is modulated with data to create an NRZ signal, followed by an RZ modulator to create RZ pulses. Finally a phase modulator is used to synchronously modulate the RZ pulse to chirp the output with the center of the pulse having maximum chirp. The optimum phase modulation is about 1.5 radians (Bakhshi et al., 2001).

Transmission experiments conducted with chirped RZ format successfully transmitted 64 channels of 5 Gb/s signals over 7000 km fiber. Recently, an integrated 10 Gb/s CRZ transmitter has been assembled (Griffin et al., 2001). A transmission experiment resulted in an error-free transmission over a 3000 km dispersion-managed fiber loop.

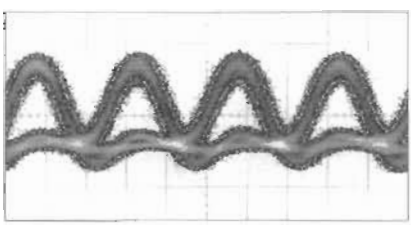


Fig. 2.12 40 Gb/s RZ transmitter output using 50 GHz p-i-n detector (H: 10 ps/div).

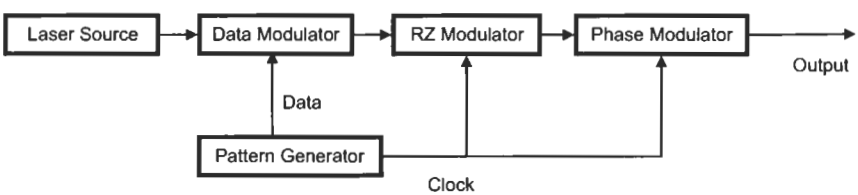


Fig. 2.13 Chirped RZ transmitter.

2.2.4.3 *Duobinary*

Duobinary coding has been used in the radio transmission, and recently was adapted to an optical transmission system (Yonenaga and Kuwano, 1997). This scheme utilizes optical phase as well as amplitude information, reducing the signal bandwidth. It is an effective method to reduce dispersion and nonlinear effects in the fiber, and spectrum efficiency is increased compared to RZ and NRZ coding. A typical block diagram of a duobinary transmitter is shown in Fig. 2.14. The duobinary signal is generated by adding one-bit delayed data to the present data. A three-level data labeled 1, 0, −1 is obtained after a narrow lowpass filter, as shown in Fig. 2.14. The optical duobinary data is generated by driving the LN modulator across the null point in its transfer curve, mapping the levels 1, 0, −1 to the optical output. The levels 1 and −1 have the same optical intensity but the opposite phase. At the receiver, a traditional direct detection receiver is used to demodulate 1, 0, and −1 into an electrical binary signal. The output spectrum is shown in Fig. 2.15. The spectrum width is reduced compared to conventional intensity modulation (IM).

2.2.4.4 *Carrier-Suppressed RZ*

Carrier-suppressed RZ has been used to compress the signal spectrum (Hirano et al., 1999; Miyamoto et al., 2000). Figure 2.16 shows the principle of operation. When the LN modulator is driven with half the frequency of the transmitting speed and biased at the null point on the transfer curve, the output

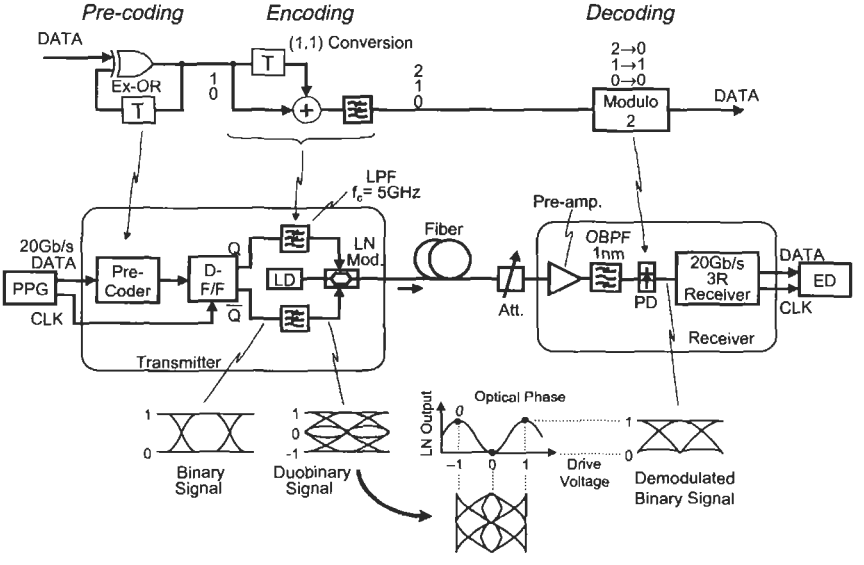


Fig. 2.14 Duobinary transmitter (T. Ono and Y. Yano, 1998. © 2001 IEEE).

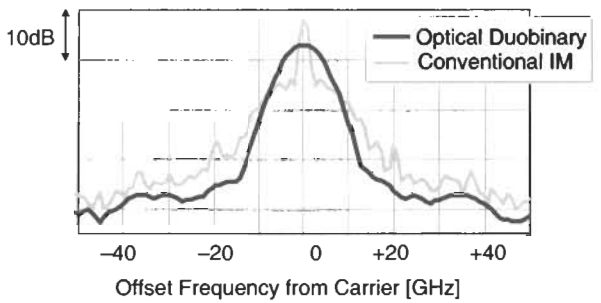


Fig. 2.15 Duobinary transmitter output spectrum (T. Ono and Y. Yano, 1998. © 2001 IEEE).

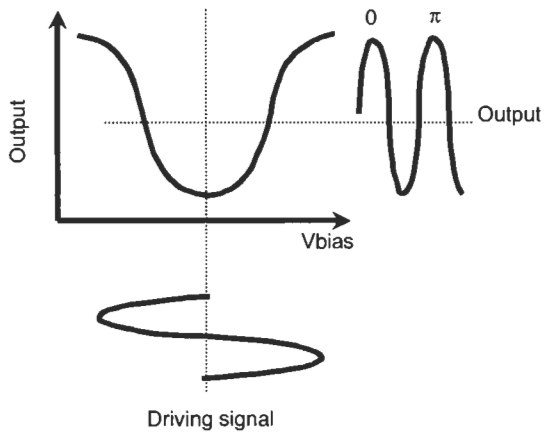


Fig. 2.16 Carrier-suppressed RZ principle of operation (Hirano et al., 1999).

pulse frequency is doubled and the relative optical phase alternates between adjacent pulses. This pulse train can modulate the transmitted data stream, resulting in alternating-optical-phase RZ, similar to duobinary modulation. This yields a narrower modulation bandwidth without a carrier in the spectrum; therefore it is more dispersion-tolerant, and higher spectral efficiency in WDM transmission can be achieved. Figure 2.17 shows a block diagram of the transmitter. The first LN modulator encodes the LD output with 40 Gb/s data, and the second modulator, driven at 20 GHz, converts the NRZ signal into a CS-RZ signal. Figure 2.18 is an output waveform at 40 Gb/s. A recent experiment succeeded in transmitting 30 WDM channels, each carrying 43 Gb/s of data, over 360 km fiber using this format. Recently, duobinary CS-RZ modulation has been demonstrated (Miyamoto et al., 2001) and has achieved the smallest relative bandwidth among RZ formats.

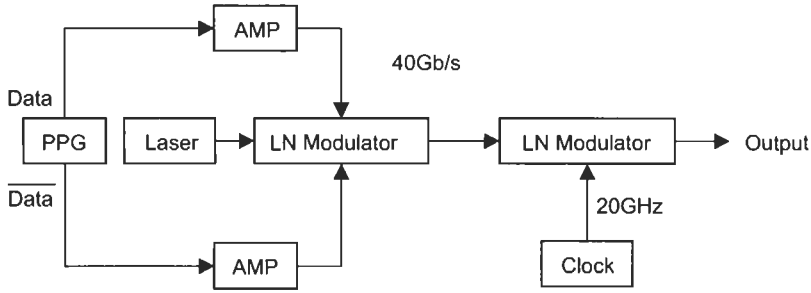


Fig. 2.17 Carrier-suppressed RZ transmitter (Hirano et al., 1999).

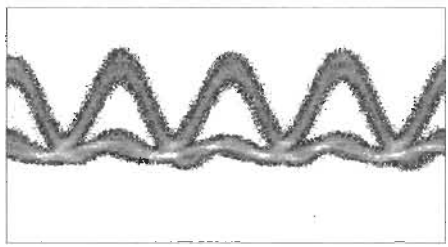


Fig. 2.18 Carrier-suppressed RZ transmitted output (H: 10 ps/div).

2.3 OTDM TRANSMITTERS FOR >100 Gb/s

There have been reports of transmission experiments with data rates as high as 1280 Gb/s (Nakazawa et al., 2000; Mikkelson et al., 2000b; Raybon et al., 2000). Most of these ultrahigh data rate transmitters have been constructed using optical time-division multiplexing (OTDM) techniques, since electronic speeds are currently limited to 80 Gb/s. OTDM requires a highly coherent high-speed laser pulse source, using techniques such as mode-locked lasers, gain-switched lasers, EA modulators, or MZ modulators. Figure 2.19 shows an example of a 160 Gb/s OTDM transmitter. In this figure, a 20 Gb/s data modulator is followed by an EA modulator to create a transform-limited 20 Gb/s short-pulse RZ signal. Optical time-division multiplexing is then employed to create 40 Gb/s, 80 Gb/s, and then 160 Gb/s data. In order to avoid coherent beat noise between adjacent bits, multiplexing with alternating polarizations is used to relax the extinction ratio requirement.

2.4 CONCLUSION

In this chapter, several different high-speed transmitter topologies have been described. The emphasis was on cost, size, and low power dissipation for directly modulated laser transmitters for short-reach applications. As the

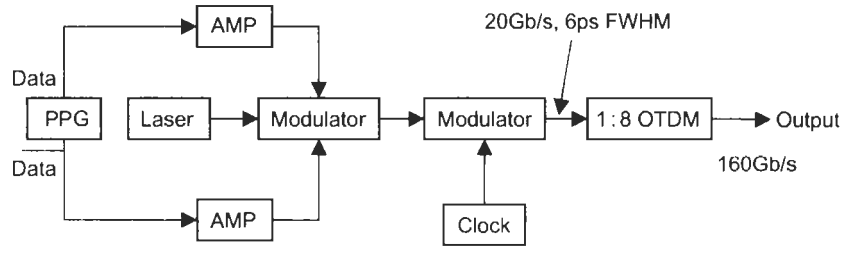


Fig. 2.19 160 Gb/s OTDM transmitter.

Table 2.1 **Transmitter Design Concepts for Various Distances**

Distance	*VSR (<2 km)*	*Short Haul (<40 km)*	*Long Haul (<100 km)*	*Very Long Haul (>100 km)*
Transmitter configuration at >10 Gb/s	1.3 μm, 1.55 μm direct modulation laser	1.3 μm, 1.55 μm external modulation (EA)	1.55 μm external modulation (LN, EA with dispersion compensation)	1.55 μm external modulation, duobinary, CRZ, CS-RZ format

transmission distance increases, reduction of chirp and ISI in the transmitting signal becomes increasingly important. External modulation schemes, such as EA and LN modulators, are good candidates for longer-distance applications. For long-reach DWDM systems, fiber nonlinearity as well as spectrum efficiency become important considerations. Several different modulation schemes were described along with their particular applications. For comparison purposes, Table 2.1 summarizes transmitter design concepts in various transmission categories.

3. High Bit-Rate Electronics for 40 Gbps and Beyond

3.1 ASIC TECHNOLOGIES FOR HIGH BIT-RATE OPTOELECTRONIC TRANSCEIVERS

3.1.1 Introduction

High-speed electronics provides an important interface between the local electronic data traffic and optoelectronic devices. The schematic diagram of

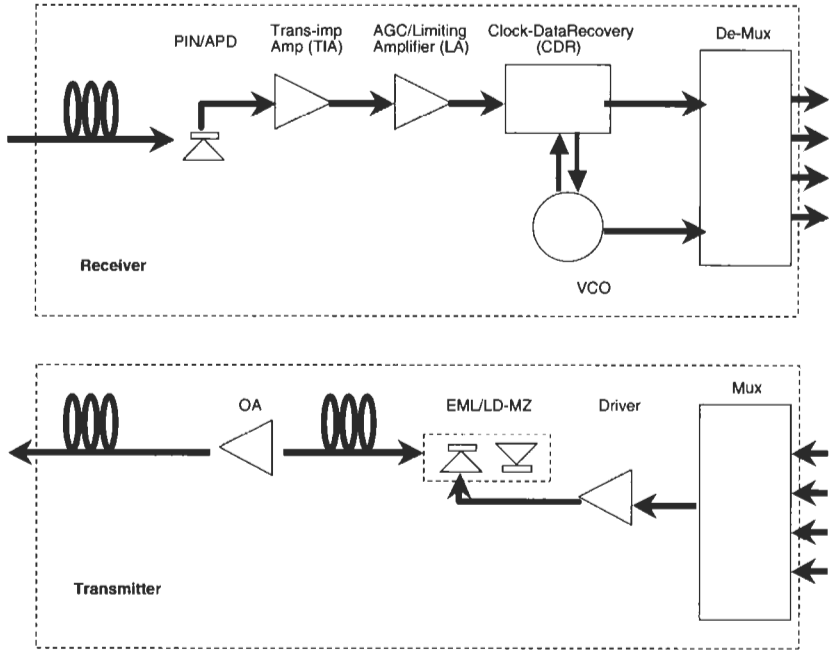


Fig. 3.1 The schematic diagram of a generic optoelectronic transceiver.

a conventional 40 Gbps optoelectronic transceiver is shown in Fig. 3.1. In the transmitter, the tributary data streams from four local 10 Gbps data channels are fed into a multiplexer and time-domain multiplexed to a single serial data stream at 40 Gbps. The modulator driver (MDD) boosts the amplitude of this combined data to modulate the electrooptical modulator such as a $LiNbO_3$ Mach-Zehnder interferometer. On the receiver side, the photocurrent from the optical detector is amplified by the transimpedance amplifier (TIA) and limiting amplifier (LA) before being fed into the clock-and-data recovery circuit (CDR). The CDR circuit recovers the incoming serial data and clock components by synchronizing the incoming data stream and local clock through phase/frequency-locked loops and a voltage-controlled oscillator (VCO). The recovered data is then broken down into lower-speed tributary channels by a demultiplexer (DeMUX).

Several kinds of high-speed ASICs are used in the transceiver—analog, digital, and mixed-mode ICs. Analog ICs with operating bandwidth as high as the bit rate are needed to handle the high-speed serial data stream. Devices such as the transimpedance amplifier (TIA), modulator driver (MDD), and voltage-controlled oscillator (VCO) must have minimum distortion and jitter generation. High-speed digital logic is needed to perform multiplexing and

demultiplexing functions with fast switching speed and minimum timing ambiguity. The phase-locked loop (PLL) in the CDR is critical to reproduce the exact timing clock and recover the received serial data stream.

In a transmission medium laden by high error rate impairments such as residual dispersion of the fiber, polarization mode dispersion (PMD), optical amplifier noise, and nonlinear effects of the optical fiber, an advanced optoelectronic receiver with a coded system is usually needed. Electronic codes are introduced in the transmitter to add redundancy in the data stream. A corresponding error correction algorithm is used in the receiver to correct random errors and to improve the error rate floor from a channel with poor optical signal-to-noise ratio (OSNR) (Chan, 1997). An electronic equalizer may also be incorporated in the receiver to compensate for transmission impairments. Figure 3.2 shows the schematic diagram of such an advanced optoelectronic transceiver. The transmitted optical data is coded with forward error correcting (FEC) codes and an equalizer IC is used in the receiver.

In the following sections, we will first survey several emerging high-speed semiconductor IC technologies and then review some of the critical ASIC functions.

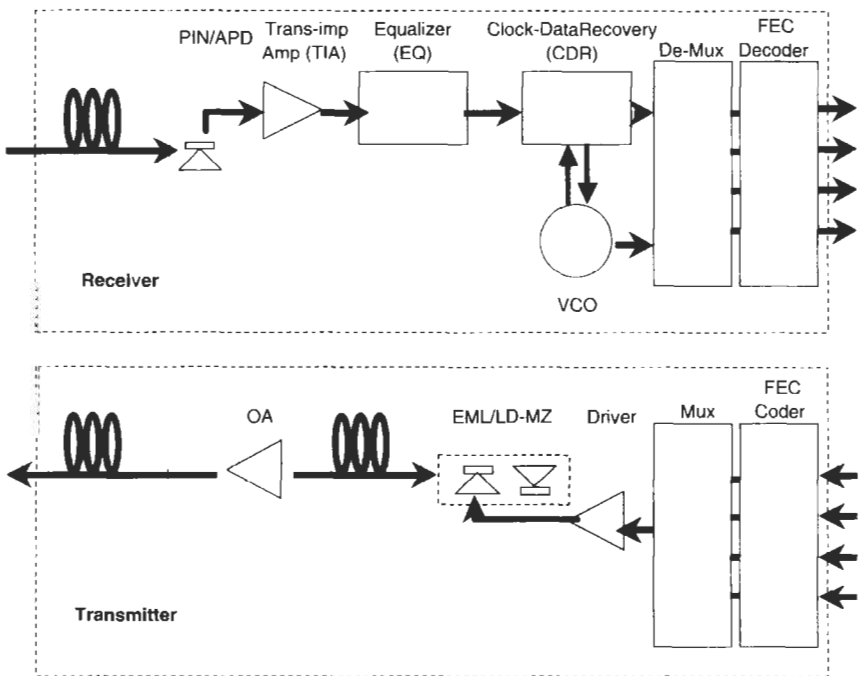


Fig. 3.2 The schematic diagram of an advanced optoelectronic transceiver with FEC line coder and equalizer.

3.1.2 High-Speed IC Technologies

Extremely high bit-rate transmission experiments up to 320 Gbps have been demonstrated with optical time-division multiplexing (OTDM) techniques with all-optical multiplexing and demultiplexing (Mikkelsen et al., 2000). However, electronic time-domain multiplexing (ETDM) techniques still have many advantages over OTDM, because electronic devices continue to enjoy high functionality, small size, low cost, and high reliability. The speed of an electronic circuit is limited mainly by the cutoff frequency of transistors. Typically, the current gain cutoff frequency f_T and the maximum oscillation frequency f_{max} are commonly used figures of merit of a transistor technology. The f_T is a good indicator of the switching speed of logic circuits, while the f_{max} is a good indicator of the bandwidth of analog circuits. In order to choose a technology to provide fast switching speed and proper system margins, a transistor technology with f_T and f_{max} better than four times the bitrate is preferred. Novel circuit technologies such as distributed amplifiers and the use of multiple-phase half-rate clocks are able to reduce the cutoff frequency requirement to twice the bitrate.

The f_T of field-effect transistors, such as MOSFET, GaAs MESFET, and GaAs/InP HEMTs, is limited by the lateral dimension of the gate electrode and the electron velocity in the host material. The f_T of these FETs is scaled up by the reduction of the gate length with advanced lithography tools. However, the cutoff frequency of bipolar transistors, such as the silicon bipolar transistor (BJT) or the heterojunction bipolar transistor (HBT), is determined mainly by the vertical layer thickness and less by the lateral dimensions.

If we track the cutoff frequencies for each IC technology and plot the best annual data from research laboratories and production lines, the resulting graph (Fig. 3.3) shows very interesting trends. The performance of transistors

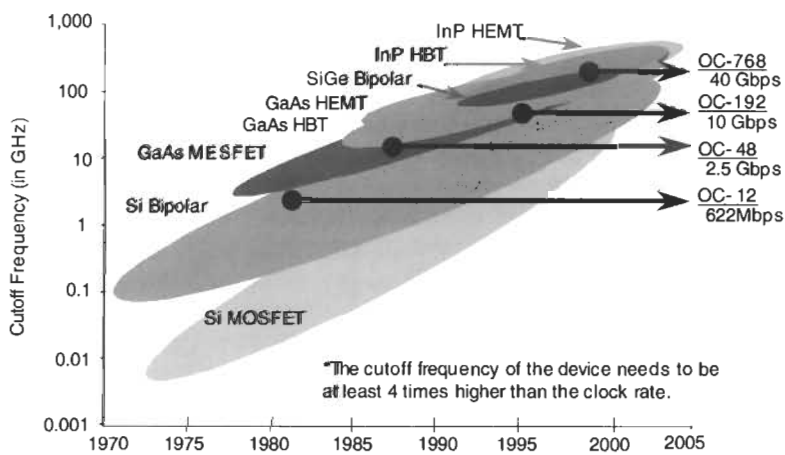


Fig. 3.3 Semiconductor technology roadmap for lightwave electronics.

improves every year mainly from the advances of fine-line lithography tools. Using the same technology, materials with better electronic properties such as GaAs, SiGe, and InP produce better transistors than the mainstream silicon CMOS devices. It is also interesting to see how physical layer ASICs have evolved for each generation of SONET standard. For example, OC-48 product development was launched in the late 1980s using GaAs MESFETs and silicon BJTs. Both were ready in early production stages with adequate f_T in the 10–15 GHz range. As the performance of CMOS production technology improved in the late 1990s, most OC-48 ASICs were ported to advanced CMOS technology for lower cost, lower power consumption, and higher integration scale. The roadmap of the evolution of ASIC technology for several SONET generations is summarized in Fig. 3.4. Today, most 10 Gbps ASICs have been implemented on GaAs and SiGe technologies with 0.5 μm line width and are being ported to advanced 0.16 μm CMOS technology. 40 Gbps ASICs are still in early stages of development as of the early 2000s. Many advanced technologies such as 0.1 μm GaAs P-HEMT, 0.1 μm InP HEMTs, GaAs HBTs, InP HBTs, and SiGe ICs are actively being evaluated. Nevertheless, a single technology will be unable to fulfill all of the requirements for a 40 Gbps transceiver. It is very common to employ a combination of several ASIC technologies in order to build a transceiver module. Figure 3.5 shows the technology tradeoffs for several critical ASIC functions. For 40 Gbps ASICs, preference lies in the use of bipolar technologies such as the InP HBT and SiGe HBT for low-power high-speed digital circuits with high transistor counts and for GaAs HEMT or InP HEMT for analog functions which require low-noise amplification and high voltage driving capability.

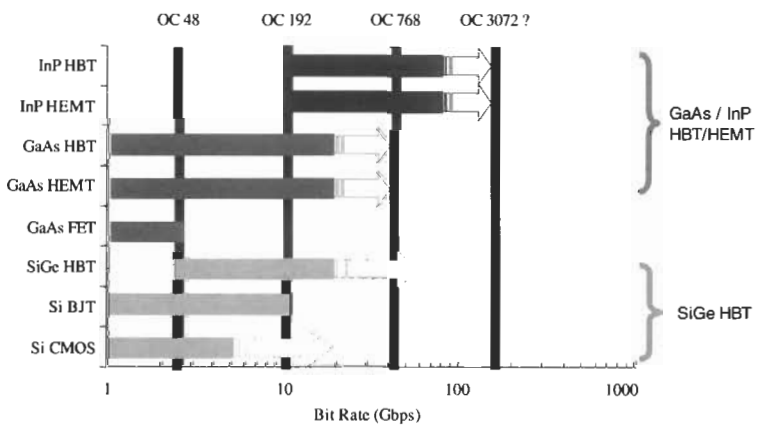


Fig. 3.4 Advanced ASIC technologies for broadband SONET lightwave transceivers.

Device Function	ASIC Technology for 40 Gbps Transceiver				
	GaAs HEMT	GaAs HBT	InP HEMT	InP HBT	SiGe HBT
MUX	--	+	-	+	++
CDR/DeMux	--	-	-	+	+
VCO	--	+	-	++	+
PreAmp	--	-	+	+	-
LA/AGC	--	+	-	++	+
Mod Driver	+	+	--	+	--

Fig. 3.5 Technology matrix for 40 Gbps optoelectronic transceiver applications.

3.2 HIGH BIT-RATE ANALOG AND MIXED-SIGNAL ELECTRONICS

Several high-frequency analog and mixed-signal functions are needed to amplify and recover high bit-rate electronic signals connected to optoelectronic devices such as photodetectors and electrooptical modulators. Broadband amplifiers such as transimpedance amplifiers and limiting amplifiers are needed to amplify the magnitude of the photocurrent for the CDR circuit to recognize the 1 and 0 state without introducing excess noise and timing jitter. The CDR will synchronize the locally generated clock signal with the incoming data stream and recover the data bits with the correct timing sequence for the digital demultiplexer circuit. On the transmitter side, a high voltage driver is also needed to drive the electrooptic (EO) modulator which requires a driving voltage that is 5 to 10 times the output from a digital multiplexer. High-speed electronic equalizers are also needed in the receiver to compensate for a portion of transmission impairments such as residual dispersion and PMD over a long fiber span. We will focus our discussions in the following sections on high-speed ASICs for the current OC-192 10 Gbps and emerging OC-768 40 Gbps applications and illustrate some of the critical design considerations of these analog function blocks.

3.2.1 Broadband Amplifiers

Frequency bandwidth, amplitude and phase response, transient behavior, and noise characteristics are key parameters of broadband amplifiers used in the optoelectronic transceiver. The bandwidth of these amplifiers should be able to handle incoming signals with an electrical spectrum ranging from as low as 8 kHz for 125 ms SONET frames to as high as the bit rate of the link. Usually the bandwidth of the NRZ signal is contained within 70% of the bit rate, e.g., a 7 GHz bandwidth is suitable for a 10 Gbps data stream. However, a higher rolloff frequency than the bit rate is preferred. This extra bandwidth

is needed to accommodate broadened signal spectra of received signals by the nonlinear effect in the fiber (i.e., self-phase modulation (SPM)) and to prevent the amplifier from introducing undesired group delay distortion near its upper rolloff frequency. The amplitude response of these analog amplifiers should be flat over the whole bandwidth within a couple of decibels to minimize intersymbol interference (ISI). The phase response should be linear up to and beyond the bit rate in order to minimize group delay dispersion. Group delay or velocity dispersion will produce unwanted timing jitter near the edges of pulses.

The frequency response of a transistor, such as current gain (h_{21}) or power gain (U), is not linear due to the parasitic RC elements as shown in Fig. 3.6. The gain at high frequency is reduced by capacitive elements at input and output junctions of the transistor and by coupling through parasitic elements. In order to have a flat gain response over a broad frequency spectrum, a feedback scheme with a passive resistive element or active transistors is used to reduce the gain at low frequencies while using a cascoded gain stage or inductive peaking scheme to boost the gain at high frequencies. For ultrabroadband amplifiers, a traveling wave configuration is used to exploit the intrinsic switching speed of the transistor. The junction capacitance of a transistor is utilized as a part of artificial transmission lines which form the input and output connections of the transistor.

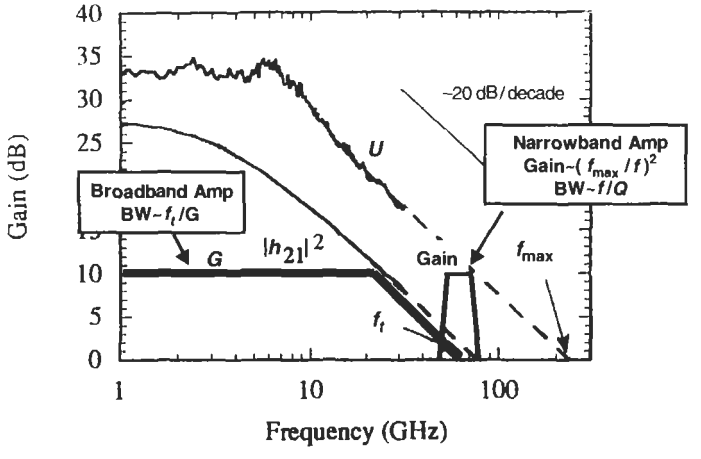


Fig. 3.6 Frequency dependence of gain from a transistor, narrowband amplifier and broadband amplifier. The current gain (h_{21}) and power gain (U) of a transistor have strong frequency dependence. Feedback elements are needed to build a broadband amplifier with a large bandwidth (BW), while resonant elements with a quality factor (Q) are needed to obtain high gain (Gain) at high frequency in narrowband amplifier.

3.2.1.1 Transimpedance Amplifier

The transimpedance amplifier (TIA) is a broadband low-noise amplifier that is used to convert and amplify the weak photocurrent I_{in} from the detector to a voltage V_{out}. Figure 3.7 shows a typical HBT circuit implementation using a lumped resistor R_f as a frequency-independent feedback element at the input gain stage followed by two output stages to achieve broadband frequency response. To the first order, the 3 dB bandwidth of this TIA is limited by the RC time constant from the equivalent Miller feedback resistor; the input capacitance of the TIA and detector can be approximated to be

$$BW_{3dB} = \frac{A_0}{2\pi R_f (C_{PD} + C_{BE,Q_{in}})} \tag{3.1}$$

where A_0 is the open-loop voltage gain of the input stage:

$$A_0 = g_{m,Q_{in}} R_C = \frac{I_{C,Q_{in}}}{k_B T} R_C \tag{3.2}$$

and the input transistor Q_{in} is biased at a collector current of I_C with a transconductance $g_{m,Q_{in}}$ at temperature T and k_B is Boltzmann's constant (e.g., $k_B T$ is 26 mV at 300°K). The frequency-dependent input-equivalent-noise power spectral density from the shot noise in the input transistor Q_{in} can be estimated as

$$\frac{d}{df}\langle i^2 \rangle_{in} = 2qI_{C,Q_{in}} \left[\frac{1}{\beta} + \frac{1}{(g_{m,Q_{in}} R_f)^2} + \left(\frac{2\pi f (C_{PD} + C_{in})}{g_{m,Q_{in}}} \right)^2 \right] \tag{3.3}$$

The overall transimpedance of this TIA is the product of the transimpedance of the first stage and the voltage gain of the output stages; it can be

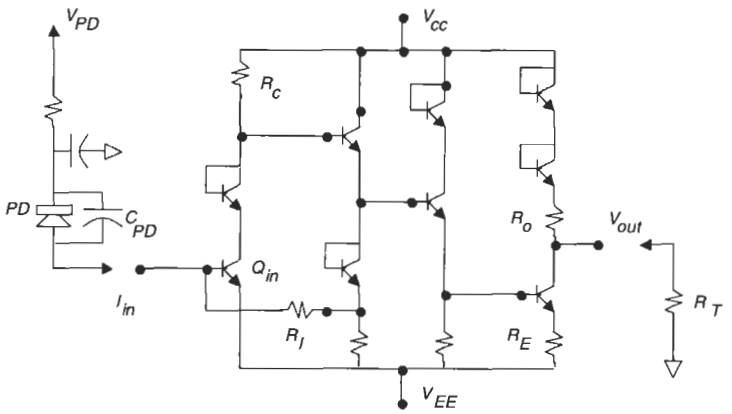
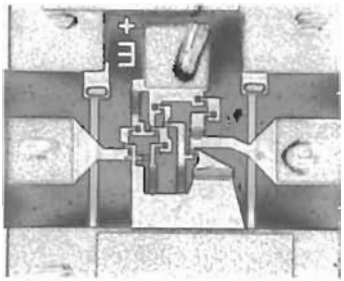
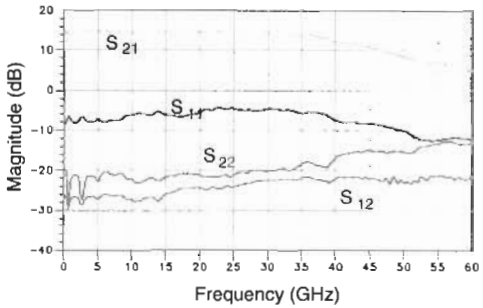


Fig. 3.7 A schematic diagram of an InP HBT transimpedance amplifier (TIA).



(a)



(b)

Fig. 3.8 (a) Microphotograph of an InP HBT TIA. The dimensions of this chip are 0.5 mm × 0.7 mm. (b) Measured small signal s-parameters over frequency from 45 MHz to 60 GHz.

approximated by

$$T_Z = \frac{V_{out}}{I_{in}} = R_f \left[\frac{R_o \parallel R_T}{R_E} \right] \tag{3.4}$$

where R_T is the external load termination resistance, R_o is the internal load resistance, and R_E is the emitter degeneration resistance of the output stage.

This TIA has been implemented using an InP HBT process at Bell Laboratories, Lucent Technologies, with a typical cutoff frequency f_T of 120 GHz. A microphotograph of this TIA is shown in Fig. 3.8a, and the measured s-parameters are depicted in Fig. 3.8b. From the measured s-parameters, a transimpedance of 220 ohms (47 dBΩ) is obtained with a 3-dB bandwidth of 56 GHz.

3.2.1.2 Traveling-Wave Amplifier (TWA)

At 40 GHz, the effective wavelength in semiconductor materials such as InP, GaAs, and silicon is approximately 2.1 mm. Even interconnecting lines of tens of microns may produce pronounced inductive effects which may cause unnecessary peaking in the frequency response and generate frequency-dependent group-delay dispersion. The passive elements and interconnect lines may no longer be seen as lumped elements, and the frequency response of TIA is not easy to predict. The traveling-wave amplifier (TWA) configuration is a very attractive alternative, making it possible to design an ultrahigh-bandwidth amplifier with a cutoff frequency close to the intrinsic transit speed in the transistor without performance issues caused by the junction capacitance (Niclas, 1984; Ayasli et al., 1982).

The idea of the TWA is to cascade multiple single-stage transistor amplifiers together, as shown in Fig. 3.9, and configure both inputs and outputs of each stage into artificial transmission lines by using the transistor junctions as parallel capacitor segments C and interstage lines as serial inductive segments

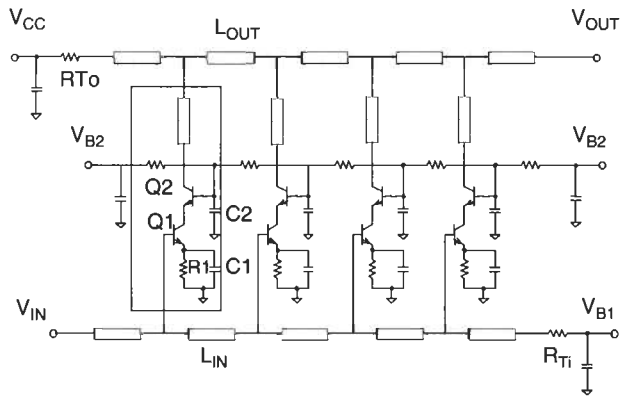


Fig. 3.9 Schematic diagram of an InP HBT traveling-wave amplifier (TWA). The dashed box shows a unitary cascode gain stage.

L to form transmission lines with extremely high cutoff frequency. The characteristic impedance Z and the cutoff frequency f_c of this artificial transmission line are:

$$Z = \left(\frac{L}{C}\right)^{1/2} \tag{3.5}$$

and

$$f_c = \frac{1}{2\pi(LC)^{1/2}} \tag{3.6}$$

The cutoff frequency f_c can be designed to be much higher than the transistor cutoff frequency. The overall amplifier bandwidth approaches the maximum oscillation frequency $f_{\max}$ of the transistor, which is mainly limited by the intrinsic carrier transit delay and dispersive elements in the transistors. The junction charging time is no longer a speed bump to the circuit response because both input and output capacitance of transistors are fully integrated into the transmission lines.

If we use the TWA to replace the TIA, then the photocurrent will drive the input transmission line of a characteristic impedance of $Z_{c,in} = (L_{IN}/C_{IN})^{1/2}$, where C_{IN} is the input impedance of the unitary gain stage formed by the cascoded transistor pair of Q1 and Q2. Usually the characteristic impedance of the input line and output line is about the same as the system impedance, $Z_0 = Z_{c,in} = Z_{c,out} = 50$ ohms. The resulting voltage gain of this TWA is

$$A_V = \frac{ng_{m,Q1}(Z_{0,in}Z_{0,out})^{1/2}}{2} \tag{3.7}$$

where g_m is the transconductance of the unitary cascoded gain stage with Q1 biased at a collector current of I_{c1}, and n is the number of stages. The maximum

number of stages is limited by the cutoff angular frequency of the amplifier w_c and serial input resistance and capacitance of Q1, $R_{in,Q1}$, and $C_{in,Q1}$:

$$n \leq \frac{2}{w_c^2 C_{in,Q1}^2 R_{in,Q1} Z_{0,in}} \tag{3.8}$$

The voltage gain and the transconductance of this HBT TWA are then simplified to

$$A_V = \frac{n g_{m,Q1} Z_0}{2} \tag{3.9}$$

and

$$T_Z = A_V Z_{0,in} = \frac{n g_{m,Q1} Z_{0,in}^2}{2} \tag{3.10}$$

The input equivalent noise current is dominated by the noise contribution from the unitary cascaded gain cell, mainly by the noise behavior of Q1. The input equivalent noise current of a HEMT TWA can be in the order of tens of picoamperes/$\sqrt{\text{Hz}}$ regime, depending on the transistor technology.

Figure 3.10(a) shows a microphotograph of the traveling-wave amplifier reported by Bell Laboratories, Lucent Technologies (Baeyens et al., 2000). The

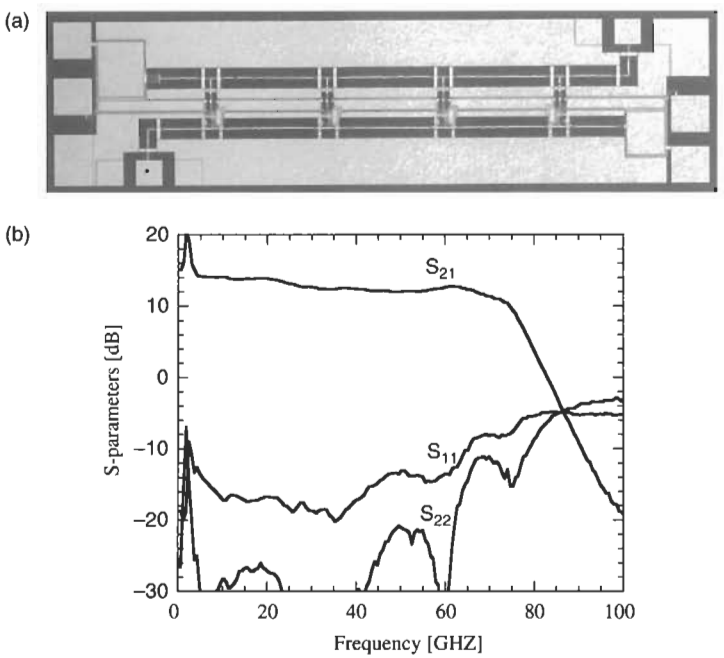


Fig. 3.10 (a) A microphotograph of a traveling-wave amplifier implemented on an InP HBT technology using coplanar transmission lines. (b) The measured frequency response of the traveling-wave amplifer indicates 12 dB gain over a bandwidth of 75 GHz. The low frequency peak is the result of poor AC grounding of this on-chip measurement from microwave wafer probes.

amplifier is implemented on InP HBT technology using coplanar transmission lines as the inductive segments in the unitary gain cell. Its measured frequency response is shown in Fig. 3.10(b) with 12 dB gain and a 3 dB rolloff frequency of 75 GHz. The data was taken on a microwave probe station on the wafer level. The peaking in the low frequency response is the result of poor AC grounding on the power supply in the microwave probes.

3.2.2 Clock and Data Recovery (CDR) Circuit

One of the most critical ASICs of the optoelectronic transceiver is the clock and data recovery (CDR) circuit. This mixed-mode IC integrates the analog clock recovery function and digital detection schemes in order to synchronize the incoming serial data stream with a locally generated clock signal. The synchronized clock is then used to recover the received data bits with the correct timing information. As shown in Fig. 3.11, the clock recovery is done by using both the phase-locked loop (PLL) and frequency-locked loop (FLL) to synchronize the voltage-controlled oscillator (VCO) to the incoming high-speed bit stream and low-frequency local system clock. The differences in the phase and frequency between the incoming data and VCO are monitored by the high-speed phase and frequency detectors and are fed back to the VCO to

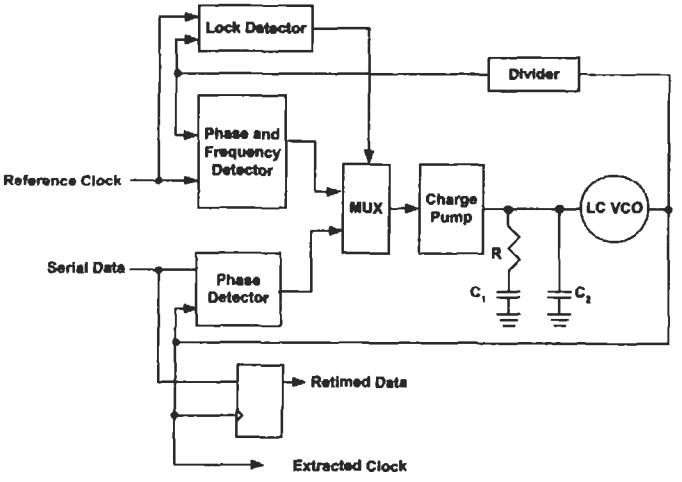


Fig. 3.11 The schematic diagram of clock and data recovery (CDR) circuit. Two loops are used to control the voltage-controlled oscillator (VCO) to recover the high-frequency clock from the incoming serial data stream and synchronize to the low-frequency local reference clock. A high-frequency phase lock loop (PLL) is used to synchronize VCO to incoming high-speed data stream, while a frequency lock loop (FLL) is engaged to lock to the low-speed local system clock. The recovered clock is then used to retrieve the incoming data bits.

correct the drift through charge pump circuitry. The recovered clock is then used to retrieve the data with a retiming latch. A detailed discussion of CDR can be found in Buchwald and Martin, 1995.

The function of the phase detector is to compare the difference between the local clock and the incoming data. A basic linear phase detector can be implemented by using either an analog mixer or a digital exclusive-OR (XOR) logic circuit as shown in Fig. 3.12. The average dc component in the output of this linear phase detector represents the phase difference between two input clocks and can be used to speed up or slow down the VCO through a low-pass filter and charge pump circuit. Figure 3.13 shows a digital frequency/phase

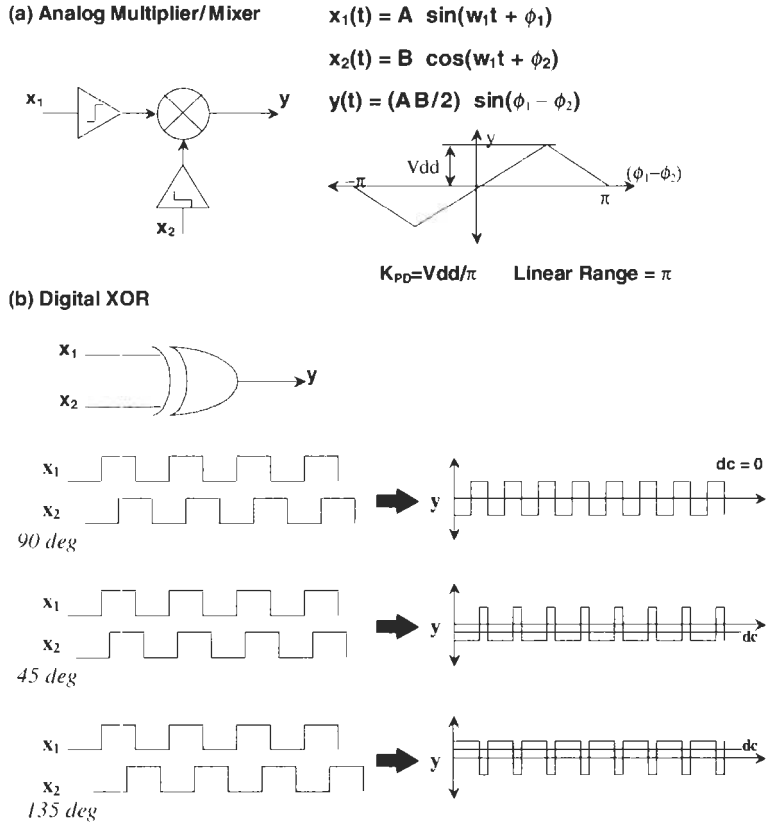


Fig. 3.12 Basic linear phase detector function can be implemented with either an analog multiplier/mixer or a digital exclusive-or (XOR) logic. (a) The low frequency component in the output of the analog multiplier/mixer indicates the phase difference between two input clocks. (b) The average output dc level of the XOR gate indicates the phase difference between two clock signals. The dc is zero when x_1 and x_2 are exactly 90 degrees apart.

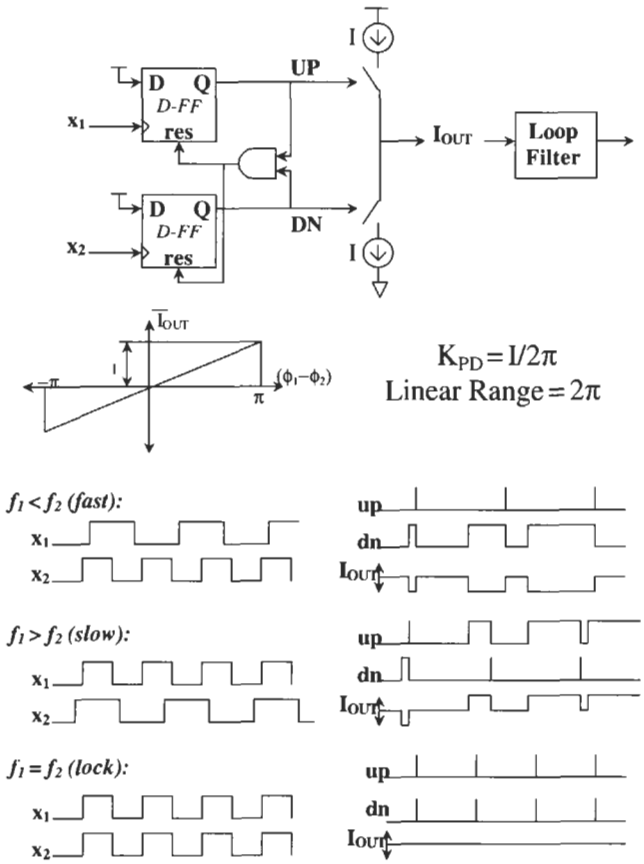


Fig. 3.13 The schematic diagram and timing chart of a digital edge-triggered digital frequency/phase detector. This circuit is insensitive to the duty cycles of the clock and can be used for both phase and frequency detection.

detector that utilizes edge-triggered data flip-flops (D-FFs) to detect the phase and frequency difference. This kind of phase detector is insensitive to the duty cycle of the incoming clocks.

To recover the clock from the detected random binary signals, the digital Alexander phase detector is widely used to implement the CDR function for lightwave circuits (Alexander, 1975). The idea is to utilize the local clock to sample the incoming random binary data stream by taking three data samples: at the middle of the bit interval before the clock transition, A; at the clock transition, T; and at the middle of the bit interval just after the clock transition, B, as shown in Fig. 3.14(a). The combination of the sampled data at A, B, and T will indicate if the local clock transition is running earlier or later than the incoming data transition through the truth table and the combinatory logic in

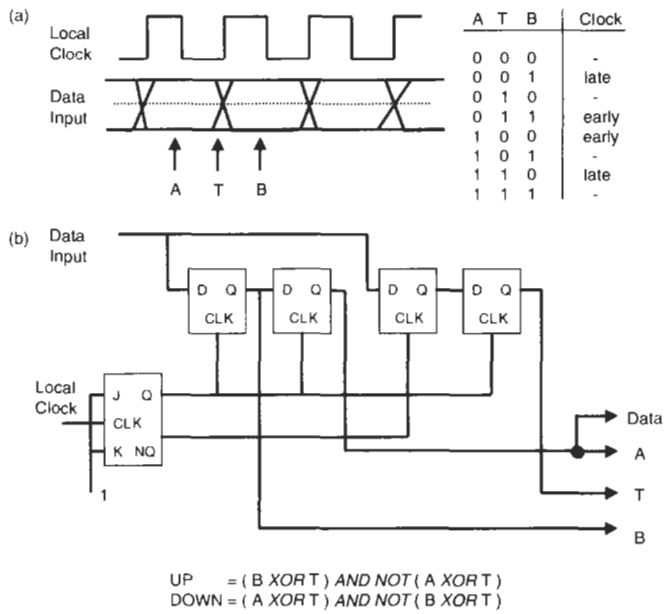


Fig. 3.14 (a) The timing diagram of an Alexander phase detector, which takes three sequential samples of the incoming random binary data A, T, and B, generated before, at, and after the data transitions timed by the local clock. (b) The up and down control signals are generated from the combinatory logic of A, B, and T to synchronize the local clock by speeding up or slowing down the VCO.

Fig. 3.14. This digital phase detector can be easily implemented with a chain of four data flip-flops as the sample-and-hold mechanism to produce A, B, and T signals from the incoming data as shown in Fig. 3.14(b). From the A, B, and T signals, UP and DOWN signals can be produced to drive the charge pump circuit and VCO to synchronize the local clock to the incoming data. As the result, both clock and data in the received serial data stream are recovered by this digital phase detector. There are many advantages of the Alexander phase detector, such as simplicity of implementation, stability over temperature and process variation, and large clock phase margin. However, the switching speed of the flip-flops and the combinatory logic has to be faster than the bit rate.

Figure 3.15 shows a complete chip of a 40 Gbps CDR designed by Bell Laboratories, Lucent Technologies, with a 120 GHz InP HBT technology. This chip integrates a high-speed digital Alexander phase detector, integrated VCO, charge pump, and a high-speed countdown divider for the external frequency-locking loop. The measured eye diagram of recovered 40 Gbps data and the spectrum of the VCO output from this CDR are depicted in Fig. 3.15.

Many CDRs utilize multiple-phase clocks at half the bit rate of the system to run the phase detection logic at lower speed. For example, the VCO

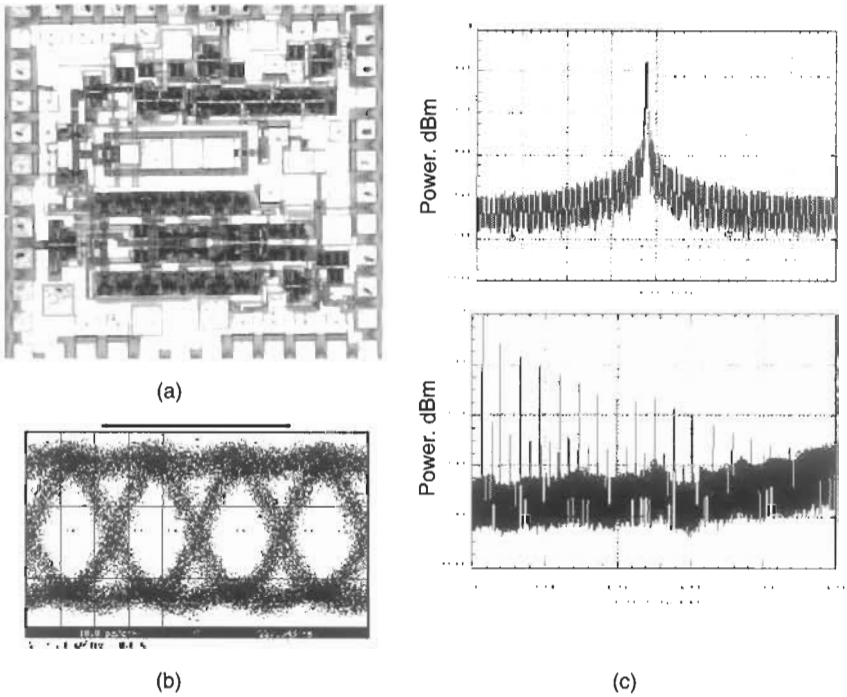


Fig. 3.15 (a) The microphotograph of single-chip CDR with integrated phase detector, VCO, charge pump, retimed data recovery, and high-speed countdown divider for external frequency locking circuit. The circuit is implemented by using InP HBT technology. The chip is measured 2.4 mm × 2.4 mm. (b) The eye diagram of the recovered data at 40 Gbps. (c) The measured output spectrum of the integrated 40 GHz VCO [Bell Laboratories, Lucent Technology].

produces four clocks at half the bitrate to drive four parallel phase detectors at four clock transitions 90° apart. This so-called "bang-bang" phase detection architecture promises wider clock phase margin with the half-rate clock and can be implemented using less demanding IC technologies with lower transistor performance (Wurzer et al., 1999). One of these implementations is illustrated by the schematic diagram of a 40 Gbps SiGe CDR/1 : 4 demux chip in Fig. 3.16 reported by Lucent Technologies (Reihold et al., 2001). The four digital phase detectors are driven by four 20 GHz clocks each with a quadrature phase difference. These quadrature phases are maintained by dividing down the 40 GHz clock with an on-chip 1 : 2 divider. The phase detection logic then produces the up and down signals for the charge pump by examining six timing samples generated by four latch chains during two data bit intervals. The retimed data is then fed into a 1 : 4 demultiplexer to regenerate four 10 Gbps data channels.

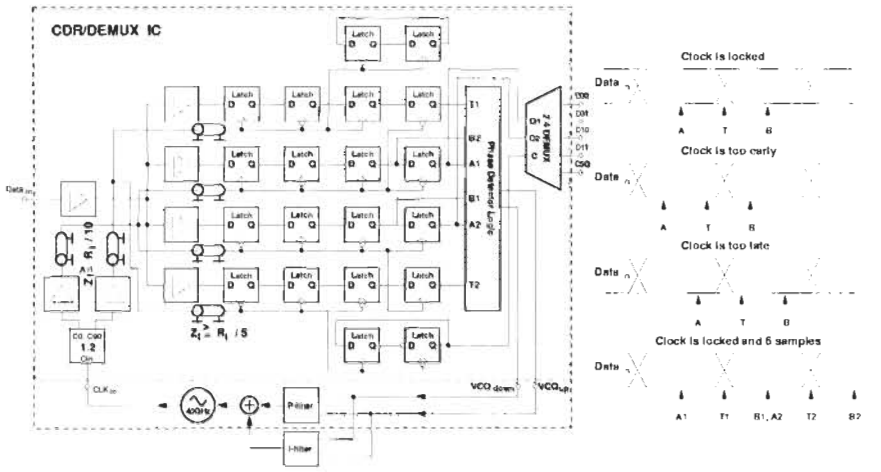


Fig. 3.16 The schematic diagram of integrated CDR/1 : 4 demux chip operating at 40 Gbps. The four digital phase detectors are driven by four quadrature phases of 20 GHz clocks produced by a divider. The phase detection logic produces the up and down signals to the charge pump and VCO circuits from six timing samples produced by four latch chains during two 40 Gbps data bits from the half-rate clock. The four clock phases are generated by a divider. The retimed data is then fed into a 1 : 4 demultiplexer to regenerate four 10 Gbps data channels.

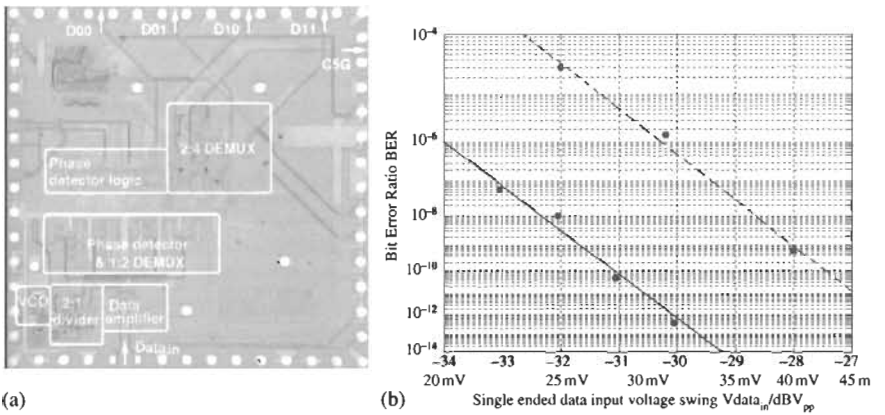


Fig. 3.17 (a) The microphotograph of a SiGe CDR/1 : 4 demux IC. (b) The measured bit-error-rate (BER) performance with respect to the input voltage swing.

This chip was realized on a 70 GHz SiGe HBT process. Figure 3.17(a) shows the microphotograph of this chip, which integrates a high-speed divider, VCO, phase detectors, 1 : 4 demultiplexer, and countdown divider for frequency locking circuitry through an external loop. The measured bit-error-rate data over the input voltage swings are shown in Fig. 3.17(b) with the integrated VCO

and with an off-chip external VCO. The BER performance of the CDR with integrated VCO is degraded by the poorer phase noise characteristics of the on-chip VCO.

3.3 HIGH-SPEED DIGITAL FUNCTIONS

3.3.1 Demultiplexer

Once the clock and data are recovered in the high-speed receiver by the CDR, the incoming data stream will be demultiplexed down into several low bit-rate data channels to perform error correction and perform add/drop functions with local traffic. The demultiplexer can readily be realized by selecting the high-speed input data stream into the lower speed channels. Figure 3.18(a) shows the schematic diagram of an InP HBT 40 Gbps 1 : 4 demultiplexer IC reported by Bell Laboratories (Mattia et al., 2000). Figure 3.18(b) shows the measured timing traces of the input 40 Gbps data stream and the demultiplexed 10 Gbps data channels before the output retiming data latches. Without the retiming output buffer, the 90° phase shift between the adjacent 10 Gbps demultiplexed channels is observed in Fig. 3.18(b).

3.3.2 Multiplexer

On the transmitter side, local tributary data channels are fed into the multiplexer and combined into a single channel at high data rate. Because the

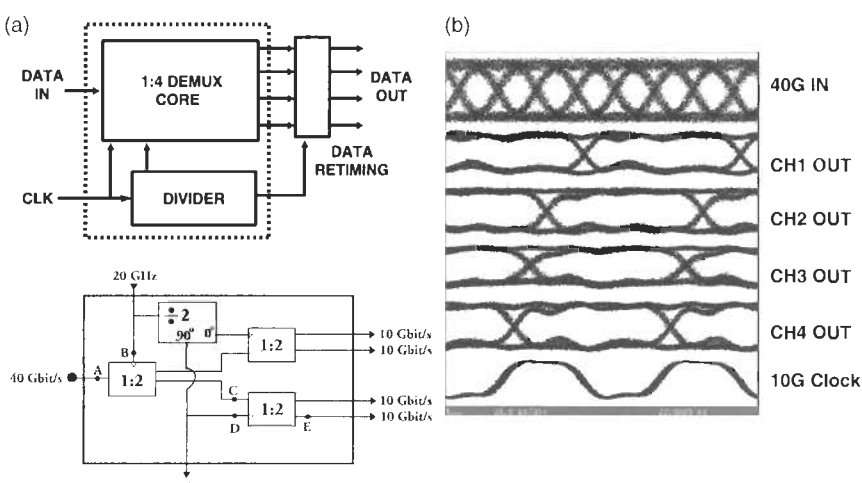


Fig. 3.18 (a) The schematic diagram of a 40 Gbps 1 : 4 demultiplexer. This ASIC was implemented on an InP HBT technology. (b) The measured output traces of the demultiplexed 10 Gbps data streams.

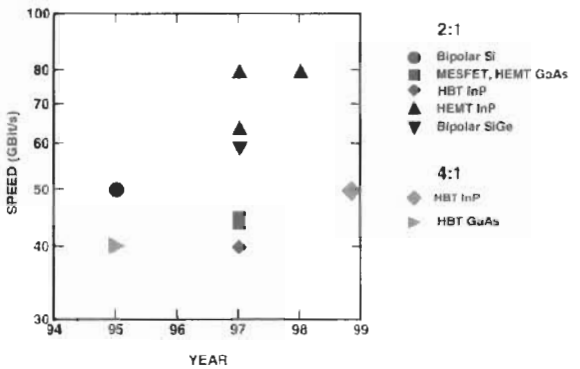


Fig. 3.19 Published speed of electronic multiplexers.

local channels come in at a fraction of the aggregated bit rate in the transmitted signal, a high-speed clock with low timing jitter needs to be generated and synchronized with the local system clock. Data from these tributary channels may arrive with various timing delays and needs to be synchronized and retimed before multiplexed into the high data rate. Figure 3.19 shows the published speed of electronic multiplexers. The highest speed of an electronic multiplexer reported to date is an 80 Gbps 2 : 1 multiplexer using an advanced InAlAs/InGaAs/InP HEMT technology reported by NTT (Otsuji et al., 1998).

The schematic diagram of an OC192 16 : 1 multiplexer, implemented on a 0.25 μm SiGe BiCMOS technology, is shown in Fig. 3.20 (Cong et al., 2000). Sixteen 622 Mbps channels are synchronized and combined into a single serial data stream at 9.953 GHz. The input buffer section of the multiplexer is shown in Fig. 3.21(a). It contains the retiming flip-flop bank to synchronize the sixteen incoming channels to the on-chip phase-locked VCO at 9.953 GHz. A FIFO register file is used to buffer the data from sixteen input channels and select the proper bit to be transmitted at the OC-192 rate of 9.953 Gbps. The multiplexer and output retiming sections are depicted in Fig. 3.21(b).

3.4 ELECTRONIC EQUALIZERS

In a high bit-rate optical transmission system, the maximum link length without electronic repeaters is limited by noise from inline optical amplifiers, fiber nonlinearity, chromatic dispersion of the fiber span, and other factors such as polarization-mode dispersion (PMD). Among proposed optical and electronic compensation schemes, advanced electronic receivers with electronic equalization techniques offer a solution that is low cost, compact, and simple. Electronic equalizers may be utilized to correct intersymbol interference (ISI) caused by transmission impairments as well as by the limited bandwidth of optical detectors and preamplifiers. There are four postdetection

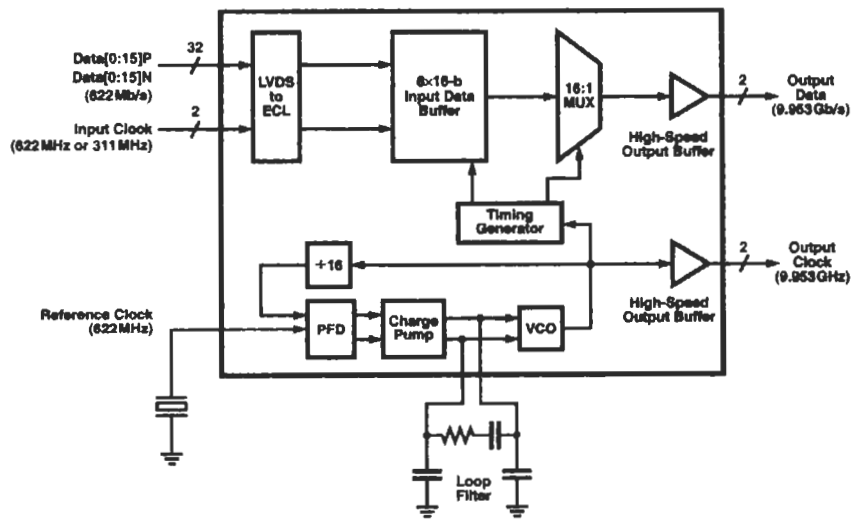


Fig. 3.20 The schematic diagram of OC-192 16 : 1 multiplexer. This chip generates the 9.953 GHz clock which is phase-locked to the 622 MHz reference clock of the transmitter. This clock is then used to synchronize and retime sixteen 622 MHz tributary channels before they are multiplexed into a single 9.953 Gbps serial data stream.

equalization techniques emerging: adaptive threshold control (ATC), feed-forward equalization (FFE), decision feedback equalizer (DFE), and the maximum-likelihood sequence estimator (MLSE). These techniques can be used individually or in combination to offer immunity to many transmission impairments and to enhance transmission distance.

In long-haul transmission, ASE from the optical amplifiers and fiber nonlinearity will contribute different amounts of noise and interference to the off and on states of the TDM optical signal at the detector (Ogawa et al., 1997). It is very important to adjust the optimal decision threshold of the limiting amplifier to maximize eye margin. Because the actual transmission impairments may drift with time, it is advantageous to have the threshold adjustment controlled by an adaptive feedback loop. The schematic diagram of an adaptive threshold control is shown in Fig. 3.22. The optimal decision threshold is determined by the error monitor circuit. The error monitoring function can be achieved by either measuring the bit-error-rate (BER) against calibrated BER-vs-Vt tables (Yonenaga et al., 2001) or measuring the eye opening margin dynamically (Ellermeyer et al., 2000). The eye monitor ASIC measures both vertical and horizontal eye opening margins and has been implemented by the Ruhr University-Alcatel group at 10 Gbps using a 50 GHz-f_T SiGe process. The eye-monitor circuit needs two level detectors to estimate the vertical eye margin and a delay locked loop (DLL) for timing and phase margins. The eye monitor function can readily be integrated into the CDR circuit.

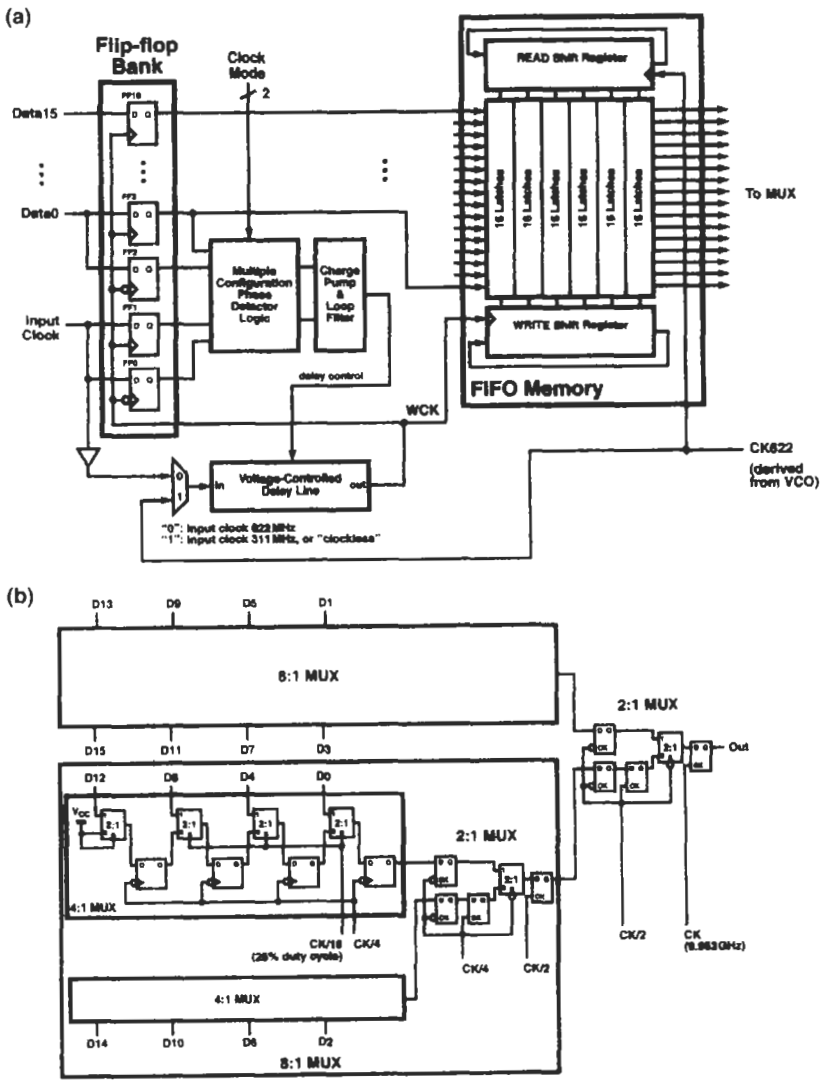


Fig. 3.21 (a) The schematic diagram shows the input buffer of the OC-192 16 : 1 multiplex. (b) The detailed configuration of the high-speed 16 : 1 multiplexer section.

To mitigate the ISI caused by transmission impairments such as PMD, residual chromatic dispersion, and nonlinear pulse propagation, several electronic adaptive equalization techniques have been proposed since the early 1990s (Winters and Gitlin, 1990). For data rates beyond 10 Gbps, only simple linear feed-forward equalizers (FFE) (Bulow et al., 1998; Cariali et al., 2000) and one-bit nonlinear decision feedback equalizers (DFE) (Moller et al., 1999;

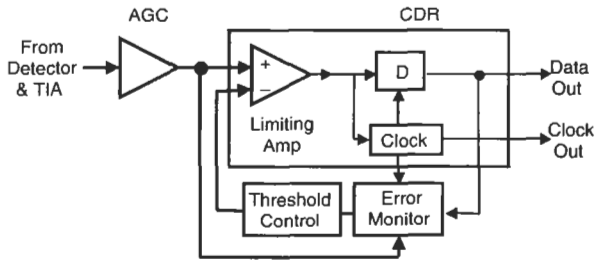


Fig. 3.22 The schematic diagram of clock-data recovery with adaptive threshold control.

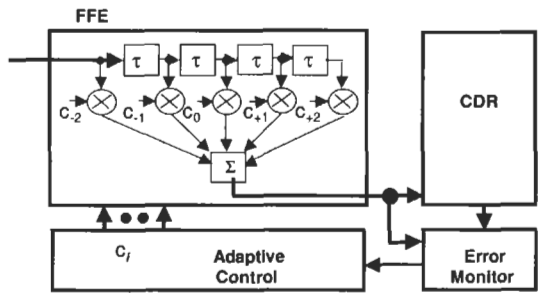


Fig. 3.23 The schematic diagram of linear feed-forward equalizer (FFE) implemented with a linear transversal filter of tapped delay lines.

Otte and Rosenkranz, 1999) have been studied and implemented because of the limitation of the processing speed of electronic circuits. At 10 Gbit/s, an electronic FFE has been experimentally demonstrated to compensate the chromatic dispersion of 1550 nm signals over 100 km of standard SMF as well as to reduce the BER penalty of 11.4 dB with PMD ($\langle\tau\rangle = 60$ ps) in a real fiber (Schlump et al., 1998). Using an electronic DFE ASIC, ISI caused by first-order PMD of up to 120 ps differential group delay was equalized at 10 Gbps (Moller et al., 1999).

The linear FFE is usually implemented as an analog transversal filter of tapped delay lines as shown in Fig. 3.23. At the input of the FFE, the received electrical signal is divided into several paths, attenuated, delayed, and recombined. The recombined signal is then amplified and sent to the CDR circuit. The delay lines are arranged in multiples of the bit period. The initial values of tap weights are determined by estimated channel characteristics (Winters and Santoro, 1990) and then can be optimized through an adaptive loop with an error monitor.

In the DFE scheme, the decision on the current bit is made based on the current input signal minus the interferences coming from previous bits and subsequent bits. The schematic diagram of a 1-bit nonlinear DFE is shown

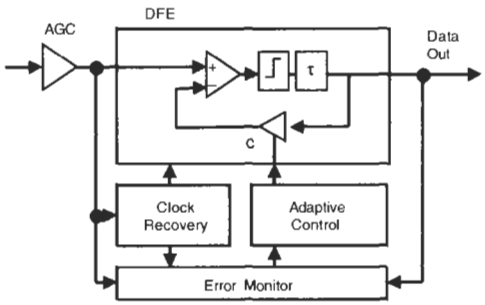


Fig. 3.24 The schematic diagram of a one-bit decision feedback equalizer (DFE).

in Fig. 3.24. The amount of feedback from the previous bit is determined adaptively by the channel characteristics using an error monitor and feedback control processor. ISI mitigation would be more effective at high data rate by using analog FFE before the nonlinear DFE.

3.5 FUTURE PROSPECTS FOR HIGH-SPEED ELECTRONIC TDM

The advance of semiconductor IC technologies has enabled basic ASIC building blocks to implement ETDM optical terminals for data rates up to 40 Gbit/s. Recently, the operating speed of a static digital frequency divider reached 70 GHz in both InP and SiGe technologies (Sokolich et al., 2000; Washio et al., 2000). Combining demonstrated broadband analog amplifiers with 100 GHz bandwidth and novel high-speed circuit design techniques, it is feasible to implement ETDM optical transceivers at 160 Gbit/s within the first decade of the 21st century. By taking advantage of the high processing speed and large-scale integration of advanced semiconductor IC technologies, many optical fiber transmission impairments at high data rates will be mitigated by low-cost high-performance electronics such as advanced electronic equalizers and spectrally efficient coders and decoders. High-speed electronics will become a critical enabler to move the capacity of ETDM/WDM optic fiber communication systems to petabits per second with individual channels running at 100 Gbit/s in the foreseeable future.

References

Adams, D., Rolland, C., Puetz, N., Moore, R., Shepherd, F., Kim, H., and Bradshaw, S. (1996). Mach-Zehnder modulator integrated with a gain-coupled DFB laser for 10 Gbit/s, 100 km NDSF transmission at 1.55 μm. *Electron. Lett.* **32**, 485–486.

Akatsu, Y., Miyagawa, Y., Miyamoto, Y., Kobayashi, K., and Akahori, Y. (1993). A 10 Gb/s high sensitivity, monolithically integrated p-i-n-HEMT optical receiver. *IEEE Photon. Technol. Lett.* **5**, 163–165.

Alexander, J. D. H. (1975). Clock recovery from random binary signals. *Electron. Lett.* **11**, 541–542.

Alexander, S. B. (1997). *Optical Communication Receiver Design.* Society of Photo-Optical Instrumentation Engineers, Bellingham, Washington.

Aoki, M., Sudo, T., Tsuchiya, D., Takemoto, D., and Tsuji, S. (2000). 85C-10 Gb/s operation of 1.3 μm InGaAlAs MQW-DFB laser. *Proc. Eur. Conf. Opt. Commun. 2000*, paper 3.3.1.

Ayasli, Y., Mozzi, R., Vorhaus, J., Reynolds, L., and Pucel, R. (1982). A monolithic GaAs 1–13 GHz traveling-wave amplifier. *IEEE Trans. Microwave Theory and Techniques* **MTT-30**, 976–981.

Baeyens, Y., Pullela, R., Mattia, J. P., Tsai, H.-S., and Chen, Y.-K. (1999). A 74-GHz bandwidth InAlAs/InGaAs-InP HBT distributed amplifier with 13-dB gain. *IEEE Microwave and Guided Wave Letters* **9**, 461.

Bakhshi, B., Vaa, M., Golovchenko, E. A., Patterson, W. W., Maybach, R. L., and Bergano, N. S. (2001). Comparison of CRZ, RZ and NRZ modulation formats in a 64 × 12.3 Gb/s WDM transmission experiment over 9000 km. *Tech. Dig. Opt. Fiber Commun. Conf. 2001*, paper WF4-1.

Barron, C. C., Mahon, C. J., Thibeault, B. J., Wang, G., Jianh, W., Coldren, L. A., and Bowers, J. E. (1994). Resonant-cavity-enhanced pin photodetector with 17 GHz bandwidth-efficiency product. *Electron. Lett.* **30**, 1796–1797.

Bergano, N., Davidson, C., Mills, M., Corbett, P., Evangelides, B., Petersen, B., Menges, R., Zyskind, J., Sulhoff, J., Srivastava, A., Wolf, C., and Judkins, J. (1997). Long-haul WDM transmission using optimum channel modulation: A 160 Gb/s (32 × 5Gb/s) 9,300 km demonstration. *Opt. Fiber Commun. Conf. 1997*, postdeadline paper PD16.

Bowers, J. E., Burrus, C. A., and McCoy, R. J. (1985). InGaAs PIN photodetectors with modulation response to millimeter wavelengths. *Electron. Lett.* **21**, 812–814.

Bowers, J. E., and Burrus, C. A. (1986). High-speed zero-bias waveguide photodetectors. *Electron. Lett.* **22**, 905–906.

Bowers, J. E., and Burrus, C. A. (1987). Ultrawide-band long-wavelength p-i-n photodetectors. *IEEE J. Lightwave Technol.* **LT-5**, 1339–1350.

Buchwald, A., and Martin, K. (1995). *Integrated Fiber-Optic Receivers.* Kluwer Academic Publishers, Norwell, Massachusetts.

Bulow, H., Schlump, D., Weber, J., Wedding, B., and Heidemann, R. (1998). Electronic equalization of fiber PMD-induced distortion at 10 Gbit/s. *Tech. Dig. Opt. Fiber Commun. Conf. 1998*, 151–152.

Campbell, J. C., Johnson, B. C., Qua, G. I., and Tsang, W. T. (1989). Frequency response of InP/InGaAsP/InGaAs avalanche photodiodes. *IEEE J. Lightwave Technol.* **LT-7**, 778–784.

Campbell, J. C. (1995). Resonant-cavity photodetectors. *Tech. Dig. Int. Electron. Devices Meeting 1995*, 575–578.

Caplan, D. O., and Atia, W. A. (2001). A quantum-limited optically-matched communication link. *Tech. Dig. Opt. Fiber Commun. Conf. 2001*, paper MM2.

Cariali, F., Martini, F., Chiappa, P., and Ballentin, R. (2000). Electronic compensation of PMD and chromatic dispersion with an IC at 10 Gbit/s transmission system. *Electron. Lett.* **36**, 889–891.

Chan, V. (1997). Coding and error correction in optical fiber communications systems. In *Optical Fiber Telecommunications III-A* (I. Kaminow and T. Koch, eds.). Academic Press, San Diego.

Chen, C. D., Kim, I., Mizuhara, O., Nguyen, T. V., Ogawa, K., Tench, R. E., Tzeng, L. D., and Yeates, P. D. (1998). 1.2-Tb/s WDM transmission experiment over 85 km fiber using 40 Gb/s line rate transmitter and 3R receiver. *Opt. Fiber Commun. Conf. 1998*, postdeadline paper PD21.

Chin, A., and Chang, T. Y. (1990). Multilayer reflectors by molecular-beam epitaxy for resonance enhanced absorption in thin high-speed detectors. *J. Vac. Sci. Technol.* **8**, 339–342.

Clark, W. R., Marittai, A., Noel, J.-P., Jatar, S., Kim, H., Jamroz, E., Knight, G., and Thomas, D. (1999). Reliable, high-gain-bandwidth product InGaAs/InP avalanche photodiodes for 10 Gb/s receivers. *Tech. Dig. Opt. Fiber Commun. Conf. 1999*, paper Tu11, 96–98.

Cohen-Johnathan, C., Giraudet, L., Bonzo, A., and Praseuth, J. P. (1997). Waveguide AlInAs avalanche photodiode with a gain-bandwidth product over 160 GHz. *Electron. Lett.* **33**, 1492–1493.

Cong, H., Logan, S., Loinaz, M., O'Brien, K., Perry, E., Polhemus, G., Scoggins, J., Snowdon, K., and Ward, M. (2001). A 10 Gbps 16 : 1 multiplexer and 10 GHz clock synthesizer in 0.25 μm SiGe BiCMOS. *Tech. Dig. IEEE Intl. Solid-State Circuits Conf. 2001*, 80–81.

Davis, G. A., Weiss, R. E., LaRue, R. A., Williams, K. J., and Esman, R. D. (1996). A 920–1650 nm high-current photodetector. *IEEE Photon. Technol. Lett.* **8**, 1373–1375.

Dentai, A. G., Kuchibhotla, R., Campbell, J. C., Tsai, C., and Lei, C. (1991). High quantum efficiency, long-wavelength InP/InGaAs microcavity photodiode. *Electron. Lett.* **27**, 2125–2127.

Ebberg, A., Auracher, F., and Borchert, B. (2000). 10 Gbit/s transmission using directly modulated uncooled MQW ridge waveguide DFB lasers in TO package. *Electron. Lett.* **36**, 1476–1477.

Ellermeyer, T., Langmann, U., Wedding, B., and Pohlmann, W. (2000). A 10-Gb/s eye-opening monitor IC for decision-guided adaptation of the frequency response of an optical receiver. *IEEE J. Solid-State Circuits* **35**, 1958–1963.

Feiste, U., Ludwig, R., Schmidt, C., Ehrke, H., and Weber, H. (2000). Unrepeated 80 Gbit/s RZ single channel transmission over 160 km of standard fiber at 1.55 μm with a large wavelength tolerance. *Tech. Dig. Opt. Fiber Commun. Conf. 2000*, paper TuD5.

Forrest, S. R. (1988). Optical detectors for lightwave communication. In *Optical Fiber Telecommunications II* (S. E. Miller and I. P. Kaminow, eds.), Academic Press, San Diego, 569–599.

Fukuchi, K., Morie, M., Tezuka, H., Suzaki, T., and Ono, T. (1995). 20-Gbit/s, 125-km standard-fiber repeaterless transmission experiment using dispersion-compensating fiber. *Tech. Dig. Opt. Fiber Commun. Conf. 1995*, 149–150.

Giboney, K., Rodwell, M., and Bowers, J. (1992). Traveling-wave photodetectors. *IEEE Photon. Technol. Lett.* **4**, 1363–1365.

Giboney, K., Nagarajan, R., Reynolds, T., Allen, S., Mirin, R., Rodwell, M., and Bowers, J. (1995). Traveling-wave photodetectors with 172-GHz bandwidth and 76-GHz bandwidth-efficiency products. *IEEE Photon. Technol. Lett.* **7**, 412–414.

Giboney, K., Rodwell, M., and Bowers, J. (1997). Traveling-wave photodetector theory. *IEEE Trans. Microwave Theory Tech.* **MTT-45**, 1313–1319.

Glave, F. E. (1972). An upper bound on the probability of error due to intersymbol interference for correlated signals. *IEEE Trans. Inform. Theory* **IT-18**, 356–363.

Goldsmith, C. L., Magel, G. A., Kanack, B. M., and Baca, R. J. (1997). Coherent combining of RF signals in a traveling-wave photodetector array. *IEEE Photon. Technol. Lett.* **9**, 988–990.

Griffin, R. A., Walker, R. G., Johnstone, R. I., Harris, R., Perney, N. M. B., Whitbread, N. D., Widdowson, T., and Harper, P. (2001). Integrated 10 Gb/s chirped return-to-zero transmitter using GaAs/AlGaAs modulators. *Opt. Fiber Commun. Conf. 2001*, postdeadline paper PD15.

Harari, J., Journet, F., Rabii, O., Jin, G., Vilcot, J. P., and Decoster, D. (1995). Modeling of waveguide PIN photodetectors under very high optical power. *IEEE Trans. Microwave Theory and Tech.* **MTT-43**, 2304–2310.

Henry, P. S. (1989). Error-rate performance of optical amplifiers. *Tech. Dig. Opt. Fiber Commun. Conf. 1989*, 170.

Hirano, A., Miyamoto, Y., Yonenaga, K., Sano, A., and Toba, H. (1999). 40 Gbit/s L-band transmission experiment using SPM-tolerant carrier-suppressed RZ format. *Electron. Lett.* **35**, 2213–2215.

Horikawa, H. (1999). Personal communication.

Humblet, P. A., and Azizoglu, M. (1991). On the bit error rate of lightwave systems with optical amplifiers. *IEEE J. Lightwave Technol.* **LT-9**, 1576–1582.

Ishibashi, T., Shimizu, N., Kodama, S., Ito, H., Nagatsuma, T., and Furuta, T. (1997). Uni-traveling-carrier photodiodes. *Tech. Dig. Ultrafast Electronics and Optoelectronics*, Opt. Soc. Amer., 1997, 166–168.

Ishikawa, K., Watanabe, K., Chiba, H., Suzuki, H., Takeyari, R., Ohhata, K., Tanoue, T., Hirata, K., Masuda, T., Hayami, R., Washio, K., and Harada, T. (2000). 40-Gbit/s optical transmitter and receiver using InP HBT and SiGe HBT ICs. *Proc. Eur. Conf. Opt. Commun. 2000*, paper 2.3.4.

Itzler, M. A., Loi, K. K., McCoy, S., and Codd, N. (2000). High-performance, manufacturable avalanche photodiodes for 10 Gb/s optical receivers. *Tech. Dig. Opt. Fiber Commun. Conf. 2000*, paper FG5, 126–128.

Jopson, R. M., Gnauck, A. H., Kasper, B. L., Tench, R. E., Olsson, N. A., Burrus, C. A., and Chraplyvy, A. R. (1989). 8 Gbit/s, 1.3 μm receiver using optical preamplifier. *Electron. Lett.* **25**, 233–234.

Kagawa, T., Kawamura, Y., and Iwamura, H. (1993). A wide-bandwidth low-noise InGaAsP-InAlAs superlattice avalanche photodiode with a flip-chip structure for wavelengths of 1.3 and 1.55 μm. *IEEE J. Quantum. Electron.* **QE-29**, 1387–1392.

Kasper, B. L. (1988). Receiver design. In *Optical Fiber Telecommunications II* (S. E. Miller and I. P. Kaminow, eds.), Academic Press, San Diego, 689–722.

Kato, K., Hata, S., Kozen, A., and Yoshida, J. (1991). High-efficiency waveguide InGaAs p-i-n photodiode with bandwidth of greater than 40 GHz. *Tech. Dig. Optical Fiber Commun. Conf. 1991*, 188.

Kato, K., Kozen, A., Muramoto, Y., Itaya, Y., Nagatsuma, T., and Yaita, M. (1994). 110-GHz, 50% efficiency mushroom-mesa waveguide p-i-n photodiode for a 1.55-μm wavelength. *IEEE Photon. Technol. Lett.* **6**, 719–721.

Kato, K. (1999). Ultrawide-band/high-frequency photodetectors. *IEEE Trans. Microwave Theory and Techniques.* **MTT-47**, 1265–1281.

Kishino, K., Ünlü, M. S., Chyi, J. I., Reed, J., Arsenault, L., and Morkoç, H. (1991). Resonant cavity-enhanced (RCE) photodetectors. *IEEE J. Quant. Electron.* **QE-27**, 2025–2034.

Kuwano, S., Takachio, N., Iwashita, K., Otsuji, T., Imai, Y., Enoki, T., Yoshino, K., and Wakita, K. (1996). 160-Gbit/s (4-ch × 40-Gbit/s electrically multiplexed data) WDM transmission over 320-km dispersion-shifted fiber. *Opt. Fiber Commun. Conf. 1996*, postdeadline paper PD25.

Lenox, C., Nie, H., Yuan, P., Kinsey, G., Homles, A. L. Jr., Streetman, B. G., and Campbell, J. C. (1999). Resonant-cavity InGaAs-InAlAs avalanche photodiodes with gain-bandwidth product of 290 GHz. *IEEE Photon. Technol. Lett.* **11**, 1162–1164.

Li, T., and Teich, M. C. (1991). Bit-error rate for a lightwave communication system incorporating an erbium-doped fibre amplifier. *Electron. Lett.* **27**, 598–600.

Lin, L. Y., Wu, M. C., Itoh, T., Vang, T. A., Miller, R. E., Sivco, D. L., and Cho, A. Y. (1996). Velocity-matched distributed photodetectors with high-saturation power and large bandwidth. *IEEE Photon. Technol. Lett.* **8**, 1376–1378.

Lin, L. Y., Wu, M. C., Itoh, T., Vang, T. A., Muller, R. E., Sivco, D. L., and Cho, A. Y. (1997a). High-power high-speed photodetectors design, analysis, and experimental demonstration. *IEEE Trans. Microwave Theory Tech.* **MTT-45**, 1320–1331.

Lin, L. Y., Wu, M. C., and Itoh, T. (1997b). Figure of merit for high-power, high-speed photodetectors. *Tech. Dig. Opt. Fiber Commun. Conf. 1997*, 40–41.

Livas, J. C. (1996). High sensitivity optically preamplified 10 Gb/s receivers. *Opt. Fiber Commun. Conf. 1996*, postdeadline paper PD4.

Lucky, R. W. (1965). Automatic equalization for digital communication. *Bell System Tech. J.* **44**, 547–588.

Ludwig, R., Pieper, W., Weber, H. G., Breuer, D., Petermann, K., Küppers, F., and Mattheus, A. (1997). Unrepeatered 40-Gbit/s RZ single-channel transmission over 150 km of standard fiber at 1.55 μm. *Tech. Dig. Opt. Fiber Commun. Conf. 1997*, 245–246.

Lugannani, R. (1969). Intersymbol interference and probability of error in digital systems. *IEEE Trans. Inform. Theory* **IT-15**, 682–688.

Makita, K., Watanabe, I., Tsuji, M., and Taguchi, K. (1995). 150 GHz GB-product and low dark current InAlGaAs quaternary well superlattice avalanche photodiodes. *Proc. Intl. Optics and Optoelectron. Conf. 1995*, 36.

Massara, A.B., Williams, K.A., White, I.H., Penty, R.V., Galbraith, A., Crump, P., and Harper, P. (1999). Ridge waveguide InGaAsP lasers with uncooled 10 Gbit/s operation at 70C. *Electron. Lett.* **35**, 1646–1647.

Matthews, J. W. (1973). Sharp error bounds for intersymbol interference. *IEEE Trans. Inform. Theory* **IT-19**, 440–447.

Mattia, J. P., Pullela, R., Baeyens, Y., Chen, Y-K., Tsai, H. S., Georgiou, G., Winkler von Mohrenfels, T., Reinhold, M., Groepper, C., Dorschky, C., Schulien, C. (2000). A 1 : 4 demultiplexer for 40 Gbps fiber-optic applications. *Tech. Dig. IEEE Intl. Solid-State Circuit Conf. 2000*, 64–65.

Mikkelsen, B., Raybon, G., Essiambre, R.-J., Stentz, A. J., Nielsen, T. N., Peckham, D. W., Hsu, L., Gruner-Nielsen, L., Dreyer, K., and Johnson, J. E. (2000a). 320 Gbps single-channel pseudolinear transmission over 200 km of nonzero-dispersion fiber. *IEEE Photonics Tech. Lett.* **12**, 1400–1402.

Mikkelson, B., Raybon, G., and Essiambre, R. (2000b). 160 Gb/s TDM transmission systems. *Proc. Eur. Conf. Opt. Commun. 2000*, paper 6.1.1.

Miyamoto, Y., Yonenaga, K., Kuwahara, S., Tomizawa, M., Hirano, A., Toba, H., Murata, K., Tada, Y., and Umeda, Y. (2000). 1.2-Tbit/s (30 × 42.7-Gbit/s ETDM optical channel) WDM transmission over 376 km with 125 km spacing using forward error correction and carrier-suppressed RZ format. *Opt. Fiber Commun. Conf. 2000*, postdeadline paper PD26.

Miyamoto, Y., Yonenaga, M., Hirano, A., and Toba, H. (2001). Duobinary carrier-suppressed return-to-zero format and its application to 100 GHz-spaced 8 × 43-Gb/s DWDM unrepeated transmission over 163 km. *Tech. Dig. Opt. Fiber Commun. Conf. 2001*, paper TuU4.

Mizuhara, O., Nguyen, T.V., Tzeng, L.D., and Yeates, P.D. (1995). 10 Gbit/s transmission experiment using direct modulation and including 8 : 1 time-division-multiplex. *Electron. Lett.* **31**, 660–662.

Moller, L., Thiede, A., Chandrasekhar, S., Benz, W., Lang, M., Jakobus, T., and Schlechtweg, M. (1999). ISI mitigation using decision feedback loop demonstrated with PMD distorted 10 Gbit/s signals. *Electron. Lett.* **35**, 2092–2093.

Morton, P.A., Shtengel, G.E., Tzeng, L.D., Yadvish, R.D., Tanbun-Ek, T., and Logan, R.A. (1997). 38.5 km error free transmission at 10 Gbit/s in standard fibre using a low chirp, spectrally filtered, directly modulated 1.55 μm DFB laser. *Electron. Lett.* **33**, 310–311.

Muramoto, Y., Kato, K., Mitsuhara, M., Nakajima, O., Matsuoka, Y., Shimizu, N., and Ishibashi, T. (1998). High-output-voltage, high speed, high efficiency uni-traveling-carrier waveguide photodiode. *Electron. Lett.* **34**, 122–123.

Murtaza, S. S., Nie, H., Campbell, J. C., Bean, J. C., and Peticolas, L. J. (1996a). Short-wavelength, high-speed, Si-based resonant-cavity photodetector. *IEEE Photon. Technol. Lett.* **8**, 927–929.

Murtaza, S. S., Tan, I.-H., Bowers, J. E., Hu, E. L., Anselm, K. A., Islam, M. R., Chelakara, R. V., Dupuis, R. D., Streetman, B. G., and Campbell, J. C. (1996b). High-finesse resonant-cavity photodetectors with an adjustable resonance frequency. *IEEE J. Lightwave Technol.* **LT-14**, 1081–1089.

Nakagawa, J., Mizuochi, T., Takano, K., Motoshima, K., Shimizu, K., and Kitayama, T. (1996). 10 Gbit/s, 285 km repeaterless optical transmission experiments with a highly sensitive optical preamplifier. *Electron. Lett.* **32**, 48–49.

Nakata, T., Takeuchi, S., Wada, S., Watanabe, I., Makita, K., Yamaguchi, M., and Torikai, T. (2000). Waveguide avalanche photodiodes with a thin InAlAs multiplication layer for over 10-Gbps high-sensitivity receivers. *Proc. Eur. Conf. Opt. Commun. 2000*, paper 5.3.5.

Nakazawa, M., Yamamoto, T., and Tamura, K. R. (2000). 1.28 Tbit/s – 70 km OTDM transmission using third- and fourth-order simultaneous dispersion compensation with a phase modulator. *Proc. Eur. Conf. Opt. Commun. 2000*, postdeadline paper 2.6.

Niclas, K. (1984). Multi-octave performance of single-ended microwave solid-state amplifiers. *IEEE Trans. Microwave Theory and Techniques* **MTT-32**, 896–908.

Nie, H., Anselm, K. A., Lenox, C., Yuan, P., Hu, C., Kinsey, G., Streetman, B. G., and Campbell, J. C. (1998). Resonant-cavity separate absorption, charge, and multiplication avalanche photodiodes with high-speed and high-gain-bandwidth product. *IEEE Photon. Technol. Lett.* **10**, 409–411.

Nielsen, T. N., Hansen, P. B., Stentz, A. J., Aquaro, V. M., Pedrazzani, J. R., Abramov, A. A., and Espindola, R. P. (1998). 10-Gbit/s repeaterless transmission at 1.3 μm with 55.1-dB power budget using Raman post- and preamplifier. *Tech. Dig. Opt. Fiber Commun. Conf. 1998*, 52–53.

Nuyts, R. (1997). Personal communication.

Ogawa, K., Tzeng, L., Park, Y.-K., and Sano, E. (1997). Advances in high bit-rate transmission systems. In *Optical Fiber Telecommunications III-A* (I. Kaminow and T. Koch, eds.), Academic Press, San Diego.

Oggazzaden, A., Lentz, C. W., Mason, T. G. B., Glogovsky, K. G., Reynolds, C. L., Przybylek, G. J., Leibenguth, J. M., Kercher, T. L., Boardman, J. W., Rader, M. T., Geary, J. M., Walters, F. S., Peticolas, L. J., Freund, J. M., Chu, S. N. G., Sirenko, A., Jurchenko, R. J., Hybersen, M. S., and Ketelsen, L. J. P. (2001). 40 Gb/s tandem electro-absorption modulator. *Opt. Fiber Commun. Conf. 2001*, postdeadline paper PD14.

Ohhira, R., Amamiya, Y., Niwa, T., Nagano, N., Takeuchi, T., Kurioka, C., Chuzenji, T., and Fukuchi, K. (1998). A high-sensitivity 40-Gbit/s optical receiver using packaged GaAs HBT-ICs. *Tech. Dig. Opt. Fiber Commun. Conf. 1998*, 155–156.

Olsson, N. A. (1989). Lightwave systems with optical amplifiers. *IEEE. J. Lightwave Technol.* **LT-7**, 1071–1082.

Onat, B. M., Gökkavas, M., Özbay, E., Ata, E. P., Towe, E., and Ünlü, M. S. (1998). 100-GHz resonant cavity enhanced Schottky photodiodes. *IEEE Photon. Technol. Lett.* **10**, 707–709.

Ono, T., and Yano, Y. (1998). Key technologies for terabit/second WDM systems with high spectral efficiency of over 1 bit/s/Hz. *IEEE J. Quantum. Electron.* **QE-34**, 2080–2088.

Otsuji, T., Murata, K., Enoki, T., and Umeda, U. (1998). 80 Gbps multiplexer IC using InAlAs/InGaAs/InP HEMTs. *Electron. Lett.* **34**, 113–114.

Otte, S., and Rosenkranz, W. (1999). A decision feedback equalizer for dispersion compensation in high speed optical transmission systems. *Tech. Digest of IEEE ICTON '99*, 19–22.

Park, Y. K., and Granlund, S. W. (1994). Optical preamplifier receivers: application to long-haul digital transmission. *Optical Fiber Technol.* **1**, 59–71.

Paslaski, J., Chen, P. C., Chen, J. S., Gee, C. M., and Bar-Chaim, N. (1996). High-power microwave photodiode for improving performance of RF fiber-optic links. *Proc. SPIE, Photon. and Radio Freq.* **2844**, 110–119.

Raybon, G., Mikkelsen, B., Stentz, A., Nielsen, T., Peckham, D., Hsu, L., Gruner-Nielsen, L., Dreyer, K., and Johnson, J. (2000). 320 Gbit/s single channel pseudo-linear transmission over 200 km of nonzero-dispersion fiber. *Opt. Fiber Commun. Conf. 2000*, postdeadline paper PD29.

Reihold, M., Dorschky, C., Pullela, R., Rose, E., Mayer, P., Pachke, P., Baeyens, Y., Mattia, J.-P., and Kunz, F. (2001). A fully-integrated 40 Gbps clock and data recovery/1 : 4 DEMUX IC in SiGe technology. *Tech. Dig. IEEE Intl. Solid-State Circuits Conf. 2001*, 84–85.

Saltzberg, B. R. (1968). Intersymbol interference error bounds with application to ideal bandlimited signaling. *IEEE Trans. Inform. Theory* **IT-14**, 563–568.

Schlump, D., Wedding, B., and Bulow, H. (1998). Electronic equalisation of PMD and chromatic dispersion induced distortion after 100 km standard fibre at 10 Gbit/s. *Proc. 24th Eur. Conf. Opt. Commun. 1998*, **1**, 535–536.

Shimbo, O., and Celebiler, M. I. (1971). The probability of error due to intersymbol interference and Gaussian noise in digital communication systems. *IEEE Trans. Commun. Tech.* **COM-19**, 113–119.

Shimizu, N., Watanabe, N., Furata, T., and Ishibashi, T. (1997). High-speed InP/InGaAs uni-traveling-carrier photodiodes with 3-dB bandwidth over 150 GHz. *Device Res. Conf. '97*, Fort Collins, CO.

Shirai, M., Watanabe, K., Ido, T., Ishikawa, K., Suzuki, H., and Takeyari, R. (2000). Low driving voltage optical modulators for 40 Gbit/s applications. *Proc. Eur. Conf. Opt. Commun. 2000*, paper 1.3.4.

Sieniawski, D. (1998). Low cost 10 Gb/s receiver module. *48th Electronic Components and Technol. Conf.*, IEEE Components, Packaging, and Manufacturing Technology Society, 1998, paper S05P2.

Smets, R. C. J., Jennen, J. G. L., de Waardt, H., Teichmann, B., Dorschky, C., Seitz, R., Reid, J. J. E., Tiemeijer, L. F., Kuindersma, P. I., and Boot, A. J. (1997). 114 km, repeaterless, 10 Gbit/s transmission at 1310 nm using an RZ data format. *Tech. Dig. Opt. Fiber Commun. Conf. 1997*, 269–270.

Smith, R. G., and Kasper, B. L. (1997). Optical receivers. In *The Communications Handbook* (Jerry D. Gibson, ed.), CRC Press, Boca Raton, 789–802.

Smith, R. G., and Personick, S. D. (1980). Receiver design for optical fiber communication systems. In *Semiconductor Devices for Optical Communications, Topics in Applied Physics*, Vol. 39 (H. Kressel, ed.), Springer-Verlag, Berlin, 89–160.

Sokolich, M., Fields, C., Shi, B., Brown, Y. K., Montes, M., Martinez, R., Kramer, A. R., Thomas, S., and Madhav, M. (2000). A low power 72.8 GHz static frequency divider implemented in AlInAs/InGaAs HBT IC technology. *IEEE 22nd Annual GaAs IC Symposium 2000*, 81–84.

Takahata, K., Muramoto, Y., Akatsu, Y., Akahori, Y., Kozen, A., and Itaya, Y. (1996). 10-Gb/s two-channel monolithic photoreceiver array using waveguide p-i-n PD's and HEMTs. *IEEE Photon. Technol. Lett.* **8**, 563–565.

Takara, H., Shake, I., Uchiyama, K., Kamatani, O., Kawanishi, S., and Sato, K. (1998). Ultrahigh-speed optical TDM signal generator utilizing all-optical modulation and optical clock multiplication. *Opt. Fiber Commun. Conf. 1998*, postdeadline paper PD16.

Takeuchi, T., Nakata, T., Makita, K., and Torikai, T. (2001). A high-power and high-efficiency photodiode with an evanescently coupled graded-index waveguide for 40 Gb/s applications. *Tech. Dig. Opt. Fiber Commun. Conf. 2001*, paper WQ2.

Tan, I., Dudley, J., Babic, D., Cohen, C., Young, B., Hu, E., Bowers, J., Miller, B., Koren, U., and Young, M. (1994). High quantum efficiency and narrow absorption bandwidth of the wafer-fused resonant $In_{0.53}Ga_{0.47}As$ photodetectors. *IEEE Photon. Technol. Lett.* **6**, 811–813.

Tan, I., Sun, C., Giboney, K. S., Bowers, J. E., Hu, E. L., Miller, B. I., and Capik, R. J. (1995). 120-GHz long-wavelength low-capacitance photodetector with an air-bridged coplanar metal waveguide. *IEEE Photon. Technol. Lett.* **7**, 1477–1479.

Taylor, H. F., Eknoyan, O., Park, C. S., Choi, K. N., and Chang, K. (1990). Traveling wave photodetectors. *Optoelectron. Signal Processing Phased-Array Antennas II*, 59–63.

Timofeev, F.N., Kostko, I.A., Bayvel, P., Berger, O., Wyatt, R., Kashyap, R., Lealman, I.F., and Maxwell, G.D. (1999). 10 Gbit/s directly modulated, high temperature-stability external fibre grating laser for dense WDM networks. *Electron. Lett.* **35**, 1737–1739.

Tucker, R. S., Taylor, A. J., Burrus, C. A., Eisenstein, G., and Wiesenfeld, J. M. (1986). Coaxially mounted 67 GHz bandwidth InGaAs pin photodiode. *Electron. Lett.* **22**, 917–918.

Tzeng, L. D., Mizuhara, O., Nguyen, T. V., Ogawa, K., Watanabe, I., Makita, K., Tsuji, M., and Taguchi, K. (1996). A high-sensitivity APD receiver for 10-Gb/s system applications. *IEEE Photon. Technol. Lett.* **8**, 1229–1231.

Ünlü, M. S., Kishino, K., Chyi, J. I., Arsenault, L., Reed, J., Mohammad, S. N., and Morkoç, H. (1990). Resonant cavity enhanced AlGaAs/GaAs heterojunction phototransistors with an intermediate InGaAs layer in the collector. *Appl. Phys. Lett.* **57**, 750–752.

Wake, D., Spooner, T. P., Perrin, S. D., and Henning, I. D. (1991). 50 GHz InGaAs edge-coupled pin photodetector. *Electron. Lett.* **27**, 1073–1075.

Washio, K., Hayami, R., Ohue, E., Oda, K., Tanabe, M., Shimamoto, H., and Kondo, M. (2001). 67-GHz static frequency divider using 0.2-μm self-aligned SiGe HBTs. *IEEE Trans. Microwave Theory and Techniques* **49**, 3–8.

Welstrand, R. B., Jiang, H., Zhu, J. T., Liu, Y. Z., Pappert, S. A., and Yu, P. K. L. (1997). High-speed and high-saturation power semiconductor waveguide photodetector structures. *Tech. Dig. Opt. Fiber Commun. Conf. 1997*, 39–40.

Wey, Y., Giboney, K., Bowers, J., Rodwell, M., Silvestre, P., Thiagarajan, P., and Robinson, G. (1993). 108-GHz GaInAs/InP p-i-n photodiodes with integrated bias tees and matched resistors. *IEEE Photon. Technol. Lett.* **5**, 1310–1312.

Wey, Y., Giboney, K., Bowers, J., Rodwell, M., Silvestre, P., Thiagarajan, P., and Robinson, G. (1995). 110-GHz GaInAs/InP double heterostructure p-i-n photodetectors. *IEEE J. Lightwave Technol.* **LT-13**, 1490–1499.

Williams, K. J., Esman, R. D., and Dagenais, M. (1994). Effects of high space-charge fields on the response of microwave photodetectors. *IEEE Photon. Technol. Lett.* **6**, 639–641.

Williams, K. J., Esman, R. D., Wilson, R. B., and Kulick, J. D. (1998). Differences in p-side and n-side illuminated p-i-n photodiode nonlinearities. *IEEE Photon. Technol. Lett.* **10**, 132–135.

Williams, K. J., and Esman, R. D. (1999). Design considerations for high-current photodetectors. *IEEE J. Lightwave Technol.* **LT-17**, 1443–1454.

Winters, J. H., and Gitlin, R.D. (1990). Electrical signal processing techniques in long-haul fiber-optic systems. *IEEE Trans. Commun.* **38**, 1439–1453.

Winters, J., and Santoro, M. (1990). Experimental equalization of polarization dispersion. *IEEE Photon. Technol. Lett.* **2**, 591–593.

Wu, M., and Itoh, T. (1993). Ultrafast photonic-to-microwave transformer (PMT). *IEEE LEOS Summer Topical Meeting Dig. Optical Microwave Interactions.*

Wurzer, M., Böck, J., Knapp, H., Zirwas, W., Schumann F., and Felder, A. (1999). A 40 Gbps integrated clock and data recovery circuit in a 50 GHz f_T silicon bipolar technology. *IEEE J. Solid-State Circuits* **34**, 1320–1324.

Yamashita, K., Oikawa, Y., Yamamota, T., Ohnishi, H., Mikawa, T., and Hamano, H. (1997). High-performance 10-Gbit/s optical receiver using avalanche photodiode and heterojunction bipolar transistor ICs. *Tech. Dig. Opt. Fiber Commun. Conf. 1997*, 16–17.

Yonenaga, K., and Kuwano, S. (1997). Dispersion-tolerant optical transmission system using duobinary transmitter and binary receiver. *IEEE J. Lightwave Technol.* **LT-15**, 1530–1537.

Yonenaga, K., Matsuura, A., Kuwahara, S., Yoneyama, M., Miyamoto, Y., Hagimoto, K., and Noguchi, K. (1998). Dispersion-compensation-free 40-Gbit/s × 4-channel WDM transmission experiment using zero-dispersion-flattened transmission line. *Opt. Fiber Commun. Conf. 1998*, postdeadline paper PD20.

Yonenaga, K., Sano, A., Yoneyama, M., Kuwahara, S., Miyamoto, Y., and Toba H. (2001). Automatic dispersion equalisation using bit error rate monitoring in 40 Gbps optical transmission system. *Electron. Lett.* **37**, 187–188.

Yun, T. Y., Han, J. H., and Park, M. S. (1995). 10 Gbit/s optical receiver of lossless tuned pin-HEMT with high performance. *Electron. Lett.* **31**, 1688–1689.

Yun, T. Y., Park, M. S., Han, J. H., Watanabe, I., and Makita, K. (1996). 10-gigabit-per-second high-sensitivity and wide-dynamic-range APD-HEMT optical receiver. *IEEE Photon. Technol. Lett.* **8**, 1232–1234.

Index to Volumes IVA and IVB

OPTICAL FIBER TELECOMMUNICATIONS IVA

COMPONENTS

Edited by
Ivan Kaminow / Tingye Li

Optical Fiber Telecommunications IV (A & B) is the fourth in a respected series that has chronicled the progress in the research and development (R&D) of lightwave communications from its infancy in 1979 to its vibrant maturity today. Active participants and authorities in the field have written the chapters, and the coverage of the field is complete. The level of presentation is aimed at engineers working in the field but at the same time tutorial introductions make the chapters accessible to students and professionals. Even managers, system purchasers and operators, lawyers, financial analysts, and venture capitalists may find material of interest.

Volume A is devoted to progress in optical component R&D. Topics include design of optical fiber for a wide range of applications, new materials for fiber amplifiers, erbium-doped fiber amplifiers, Raman amplifiers, electrooptic modulators, optical switching, optical switch fabrics, planar lightwave devices, fiber gratings, semiconductor lasers for a wide range of lightwave applications, semiconductor amplifiers, all-optical regeneration, and high bit-rate electronics.

Volume B is devoted to lightwave systems and system impairments and compensation. System topics include growth of the Internet, network architecture evolution, undersea systems, ultra long haul, pseudo-linear high speed TDM transmission, cable TV evolution, access networks, gigabit Ethernet, and photonic simulation tools. Impairment and compensation topics include nonlinear effects, polarization mode dispersion, bandwidth efficient modulation formats, error control coding, and electronic equalization techniques.

Ivan Kaminow retired from Bell Labs in 1996 after a 42-year career, mostly in lightwave research. At first, he conducted seminal studies on electrooptic modulators and materials, Raman scattering in ferroelectrics, integrated optics (including titanium-diffused lithium niobate modulators), semiconductor lasers (including the DBR laser, ridge waveguide InGaAsP laser and multi-frequency laser), birefringent optical fibers, and WDM lightwave networks. Later, as head of the Photonic Networks and Components Department, he led research on WDM components (including the erbium-doped fiber amplifier, waveguide grating router and the fiber Fabry-Perot), and on WDM local and wide area networks. He is a member of the National Academy of Engineering and a recipient of the John Tyndall award. He now consults on lightwave technology.

Tingye Li retired from AT&T in December 1998 after a 41-year career at Bell Labs and AT&T Labs in microwave, antenna, laser, and lightwave research. His early work on laser resonator modes established the basis for the understanding of laser operation. Since the late 1960s, Li and his groups have conducted pioneering studies on lightwave technologies and systems. As head of the Lightwave Systems Research Department he led the work on amplified WDM transmission systems, and advocated their deployment for upgrading network capacity. He was deeply involved in the concurrent R&D work at AT&T on WDM systems from 1988 until his retirement. He is a member of the National Academy of Engineering and is a recipient of the OSA/IEEE John Tyndall award. He now consults on lightwave technologies and systems.

ACADEMIC PRESS
An Elsevier Science Imprint

UPC
6 08628 51720 9

EAN
90339
9 780123 951724
ISBN 0-12-395172-0