

# SILICON PHOTONICS

*Silicon Photonics: The State of the Art* Edited by Graham T. Reed  
© 2008 John Wiley & Sons, Ltd. ISBN: 978-0-470-02579-6

# SILICON PHOTONICS

## THE STATE OF THE ART

Edited by

**Graham T. Reed**

*Advanced Technology Institute,  
University of Surrey, UK*



John Wiley & Sons, Ltd

Copyright © 2008 John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,  
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk  
Visit our Home Page on [www.wiley.com](http://www.wiley.com)

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to [permreq@wiley.co.uk](mailto:permreq@wiley.co.uk), or faxed to (+44) 1243 770620.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The Publisher is not associated with any product or vendor mentioned in this book.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

#### ***Other Wiley Editorial Offices***

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDongall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, ONT, L5R 4J3, Canada

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

#### ***British Library Cataloguing in Publication Data***

A catalogue record for this book is available from the British Library

ISBN: 978-0-470-02579-6

Typeset in 10/12pt Times by Aptara Inc., New Delhi, India  
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire

*To Alison, Hannah and Matthew for making it all worthwhile.*

# Contents

<b>Foreword</b>	<b>xi</b>
<b>About the Editor</b>	<b>xvii</b>
<b>List of Contributors</b>	<b>xix</b>
<b>Acknowledgements</b>	<b>xxiii</b>
<b>1 Introduction: The Opto-Electronic Integrated Circuit</b>	<b>1</b>
<i>Richard Soref</i>	
1.1 A Few Words About History	2
1.2 The Possibilities for OEICs	4
1.3 The Present Status of OEICs	6
1.4 Silicon-based GeSn and SiGeSn Technology	9
1.5 OEICs for the Near, Mid and Far Infrared	9
1.6 Opto-Eletronic Integration with Ultimate CMOS and Post CMOS	10
1.7 Nanophotonics Integrated with Nanoelectronics	10
1.8 Conclusion	12
References	13
<b>2 Silicon Photonic Waveguides</b>	<b>15</b>
<i>G. Z. Mashanovich, G. T. Reed, B. D. Timotijevic and S. P. Chan</i>	
2.1 Introduction	15
2.2 Planar Waveguides	15
2.3 Rib Waveguides	23
2.4 Strip Waveguides	28
2.5 Coupling to Small Silicon Waveguides	31
2.6 Novel Silicon Waveguide Structures	36
2.6.1 Free-standing Waveguides	36
2.6.2 Hollow Waveguides	40
2.7 Summary and Conclusions	43
References	44
<b>3 Silicon-based Photonic Crystal Structures: From Design to Realization</b>	<b>47</b>
<i>Dennis W. Prather, Shouyuan Shi, Janusz Murakowski, Garrett Schneider, Ahmed Sharkawy, Caihua Chen and BingLin Miao</i>	
3.1 Theory of Photonic Crystals and Photonic Bandgap Structures	47
3.1.1 Analogy Between Photonic and Semiconductor Crystals	50
3.1.2 Analyzing Photonic Bandgap Structures	51

3.1.3 Doping of Photonic Crystals	57
3.1.4 Waveguides and Nano-cavities in Photonic Crystals	58
3.2 Fabrication and Characterization of Photonic Crystals	60
3.2.1 Planar/2D Photonic Crystal Fabrication	61
3.2.2 3D Photonic Crystal Fabrication	67
3.3 Overview of Photonic Bandgap Applications in Silicon Photonic Integrated Circuits	72
3.4 Silicon Photonic Crystal Dispersion-based Applications	75
3.4.1 Non-channel Waveguide	75
3.4.2 A Unidirectional Emitter	78
3.4.3 Sub-surface Silicon Optical Bus Based on Three-dimensional Photonic Crystals	80
3.4.4 Dispersion-based Variable Beam Splitter	82
3.4.5 Photonic Crystal Analog-to-digital Converter	84
3.5 Conclusions	86
References	86
<b>4 Optical Modulators in Silicon Photonic Circuits</b>	<b>95</b>
<i>F.Y. Gardes, G.T. Reed, G.Z. Mashanovich and C.E. Png</i>	
4.1 Introduction	95
4.2 Mechanisms for Optical Modulation in Silicon	96
4.2.1 Absorption	97
4.2.2 Optical Modulation Mechanisms in Silicon	97
4.3 A Brief History of Silicon Modulator Designs	106
4.4 Conclusion	141
References	142
<b>5 Silicon Lasers</b>	<b>147</b>
<i>Bahram Jalali, Dimitris Dimitropoulos, Varun Raghunathan and Sasan Fathpour</i>	
5.1 Introduction	147
5.2 Why Bulk Silicon Cannot Amplify Light	147
5.2.1 Physics of Optical Gain in Bulk Silicon	148
5.3 Approach 1: Quantum Confinement	153
5.4 Approach 2: Erbium Doping and Other Approaches	156
5.4.1 Hybrid Integration of III-V and Silicon	159
5.5 Approach 3: Raman Effect	160
5.5.1 Physics of Raman Scattering in Silicon	161
5.5.2 Limitation of Active Carrier Removal Using a p-n Junction	166
5.5.3 Electrical Power Dissipation Caused by Active Carrier Sweep-out	168
5.5.4 Raman Wavelength Conversion	170
5.5.5 Comparison of Raman with Other Optical Nonlinearities in Silicon	172
5.6 Experimental Realization of Raman-based Silicon Photonic Devices	172
5.7 GeSi Raman Devices	178
5.8 Mid-wave IR (MWIR) Silicon Photonics	180
5.8.1 Prospects of Mid-infrared Silicon Raman Devices	181
5.9 Energy Harvesting in Silicon Raman and Other Nonlinear Optical Devices	182
5.10 Summary	186
References	186

---

<b>6 Optical Detection Technologies for Silicon Photonics</b>	<b>191</b>
<i>A.P. Knights and J.D.B. Bradley</i>	
6.1 Introduction	191
6.2 Photodetector Theory	192
6.2.1 Band-edge Photo-excitation of Charge Carriers	192
6.2.2 Recombination and Generation via Deep Levels	193
6.2.3 p–n Junction Theory of Semiconductors	195
6.2.4 Photodiodes	200
6.3 Photodetector Performance Characteristics	202
6.3.1 Quantum Efficiency	202
6.3.2 Responsivity	203
6.3.3 Response Time	203
6.3.4 Detector Noise	204
6.4 Sub-bandgap Detection	205
6.4.1 Hybirdization	206
6.4.2 SiGe-based Devices	207
6.4.3 Integration of Ge Detectors with SOI Waveguide Geometries	213
6.4.4 Sub-bandgap Detection via Chemical Doping	217
6.4.5 Defect-enhanced Infrared Response	220
6.5 Conclusions	225
References	225
<b>7 Passive Silicon Photonic Devices</b>	<b>229</b>
<i>Ansheng Liu, Nahum Izhaky and Ling Liao</i>	
7.1 The Waveguide Bend	229
7.2 The Directional Coupler	233
7.3 The Multimode Interference Coupler	237
7.4 The Y-junction	242
7.5 The Mach–Zehnder Interferometer	245
7.6 Bragg Gratings	250
7.7 Fabry–Perot Resonators	255
7.8 Ring Resonators	259
7.9 Conclusion	263
References	264
<b>8 Integration</b>	<b>269</b>
<i>Cary Gunn</i>	
8.1 Motivation for Integration	269
8.2 CMOS Integration Approaches	269
8.2.1 Hybrid Integration	269
8.2.2 Monolithic Integration into CMOS	270
8.3 Detailed Description of the Front-end Integration Approach	272
8.3.1 Transistor and CMOS Process Selection	272
8.3.2 CMOS Process Selection	274
8.3.3 Passive CMOS Waveguide	275
8.3.4 Active CMOS Waveguide	275

8.3.5 <i>Implants and Activation</i>	277
8.3.6 <i>Edge-coupling vs Surface-coupling and Reliability</i>	277
8.3.7 <i>Germanium Integration</i>	279
8.4 State of the Art	285
8.4.1 <i>An Integrated 40 Gbps Wavelength Division Multiplexed Transceiver</i>	285
8.4.2 <i>An Integrated 10 GHz Opto-electronic RF Oscillator</i>	289
8.4.3 <i>Aurora</i>	293
8.5 Conclusions	294
References	295
<b>9 Silicon Photonic Applications</b>	<b>297</b>
<i>Richard Jones, Haisheng Rong, Hai-Feng Liu and Mario Paniccia</i>	
9.1 Introduction	297
9.2 Communications and Interconnects	298
9.2.1 <i>Radio-over-fiber (RoF) RF Applications</i>	304
9.3 Nonlinear Optical Effects in Silicon and Applications	307
9.3.1 <i>Silicon Amplifiers and Lasers</i>	308
9.3.2 <i>Wavelength Conversion</i>	312
9.4 Sensing	314
9.4.1 <i>Physical Sensors</i>	314
9.4.2 <i>Chemical Sensors</i>	316
9.4.3 <i>Biochemical Sensors</i>	317
9.4.4 <i>Integrated Lab-on-a-chip</i>	319
9.5 Summary and Conclusions	320
Acknowledgements	320
References	321
<b>Index</b>	<b>327</b>

# Foreword

Civilized society has traditionally measured progress by the materials used to fashion structures and implements. However, the Ages of Stone, Bronze and Iron have now been followed by the Information Age where bits are the implements and knowledge is the product. During the last decade, silicon photonics has emerged as a viable alternative to compound semiconductor materials platforms and as the only solution to mega-unit volume manufacture of photonic integrated circuits. A twenty-year gestation period for novel research concepts to become pervasive commercial products is typical. With silicon photonics scheduled for high-volume market applications in the 2010–2015 window, products and advanced prototypes have already entered the market. Now, more than ever before, is the time to document the science and technology for a knowledge base to launch the new era.

The Information Age is progressing in four consecutive phases of development: telecommunication; computation; imaging; and learning. In the telecommunication phase, smoke signals became Morse code telegraphy, then copper-based, analog voice transmission, and finally digital voice/video/data content. The emergence of digital electronics in the computation phase drove early electronic switching for voice, and later massive bandwidth demand for data communication. As the imaging phase began, the advent of image communication with the World Wide Web created further bandwidth demand. The learning phase is incubating with interactive computation and imaging activities that drive even higher bandwidth requirements, and that will ultimately eliminate software with neural architectures of distributed information and adaptive intelligence.

The hallmark of the Information Age has been unparalleled growth in both content and components. About thirty years ago, I attended a strategy session of my employer, AT&T Bell Laboratories. The question under consideration was whether optical fiber transmission media would ever be commercially viable. After considerable comparison of various figures-of-merit, the decision was that the current transition from twisted-pair copper to coaxial cable would more than account for the anticipated growth in the voice market (now known as POTS, ‘plain old telephone service’). The analysis was correct, but we did not ask the right question. In 2000, twenty-five years later, a joint MIT-industry team founded the Communication Technology Roadmap, <http://mph-roadmap.mit.edu>, to consider the technology supply chain issues that were created by the enormous growth of telecommunications and the limits of component supply and performance. The erbium-doped fiber amplifier had enabled an information capacity scaling of optical networks with wavelength division multiplexing (WDM) to provide for ten years of growth before the year had ended. The inconsistencies of the network-build cycle time, the technology-development costs and the consequent need for consolidation among the >700 companies with product offerings that once were served by one company led, by the end

of the next year, to the ‘bursting of the telecom bubble’. Demand for components disappeared with bandwidth in excess; hardware became a commodity; and tens of thousands of optical communications specialists had no place to practice. Would we ask the right question this time?

History suggested that the proper question should be (to paraphrase Tolstoy), ‘How much bandwidth does a person need?’ The best subject for analysis was a typical teenaged student who was reported to be watching television, talking on a cellphone, working on a computer, playing a digital game and when asked ‘What are you doing?’, replied ‘My homework!’. A quick calculation showed that one megabit-per-second could exceed all of the data needs of this individual. We quickly realized that we had asked the wrong question again. A new transformation had occurred. People were not the only consumers of information; machine-to-machine communication for commerce and data storage was dominant. This change in network use drove demand for ‘shorter-reach’ bandwidth at the network edge. Premium ‘long-haul’ services had become a commodity, and the new communication value point was with servers and storage.

The first release of the Communication Technology Roadmap in May of 2005 posted as its primary conclusion,

Photonics technology will be driven by *electronic–photonic synergy and short (<1 km) reach interconnection*. This direction will ignite a major shift in leadership of the optical component industry from information transmission (telecom) to information processing (computing, imaging).

The lesson of the Information Age has been that the *best measure of progress is the rate of change*. The telecom, optical component paradigm is customized, discrete components with fiber ‘pigtails’ and a total addressable market that peaks at a few hundred thousand units per year. The new need is tens of millions of standardized components per year with low cost and high functionality that is enabled by large-scale integration. As an example, one server machine in 2010 will deliver 6 PFLOP ( $6 \times 10^{15}$  floating-point operations) performance that is enabled by the parallelism of interconnected computational units. One such server cluster will require  $\sim 2.5$  million optical interconnects at 10 Gb/s with a budget of \$20–40M for optical transceivers. The optical transmission needs of a single data center dwarf the earlier vision for telecom. *This emerging demand for high-volume production at low cost of standard components with integrated functionality is the demand pull for silicon photonics*.

The mere existence of a ‘state of the art’ for silicon photonics is both logical and illogical. The commercial success of fiber optic communication systems is based on the transparency of silica ( $\text{SiO}_2$ ). Planar lightguide circuits (PLCs) based on doped-silica waveguides were the first stage of transformation of fiber optics to a planar platform. It would seem natural to exploit the materials compatibility of the Si/ $\text{SiO}_2$  materials systems for photonics, since many of the planar processing issues have already been solved for electronic integrated circuit chips. In addition, silicon presents a high-thermal-conductivity substrate; it is transparent at  $\lambda = 1550$  nm; and its refractive index can be actively modified by free-carrier injection or temperature change; and it has a high refractive index ( $n = 3.5$ ) that enables scaling to dimensions of  $\lambda/n$  ( $\sim 500$  nm for silicon waveguides).

However, silicon photonics is a misnomer to those specializing in active photonic devices: light emitters, photodetectors, modulators; since silicon has an indirect gap band structure that offers weak interaction between electrons and photons. In addition, the design principles of the highly successful fiber platform are based on optimized single devices and materials

diversity... the diametric opposite of the materials and process integration mindset of CMOS circuit design and fabrication. Now the perennial question of electronics has been transposed to photonics: ‘Will silicon, once again, render III–V compounds to being the *materials of the future?*’

Has photonic technology advanced on the silicon platform? This question is a good one, and it can be answered affirmatively. The chapters in this book’s presentation of the state of the art document that progress. If the driver for silicon photonics is high performance/cost enabled by monolithic device integration, then the key metrics to follow are speed, power and footprint. The high index contrast of the Si/SiO<sub>2</sub> materials system allows record reduction in photonic device footprint with new performance achievement. Micrometer-dimensioned silicon ring resonator devices provide optical filter response with wide (>20 nm) free spectral range (FSR). Monolithic, waveguide-integrated germanium photodetectors convert 1550 nm light to electrical current with >95% quantum efficiency and >10 GHz bandwidth, simultaneously. Optical modulators have been reduced to small sizes whose low capacitance dissipates record low power. Integrated data links have been created in low-cost, CMOS technology process flows. These demonstrations are creating a standard set of materials, processes, and design/fabrication tools. With photonic capability throughout the interconnection hierarchy, from chip to board to box, the effortless parallelism of optics has become an architectural necessity. As the electronic chip industry encounters scaling roadblocks in power density, bandwidth and latency, a realization is emerging that integrated silicon photonics is the only option for continued exponential increase in chip performance with time. For server cluster interconnection the expectations are low cost, low power, less space/Gb/s, lower shielding costs, and improved cable management.

Where does silicon photonics face its most difficult challenge? This is also a good question, and its answer could determine the ultimate viability of a commercial silicon photonic technology. Practitioners of the seminal III–V materials platform would answer that an efficient light emitter is missing. Those from the silicon electronics side would ask for more standardization in design and processing. One can envision a discrete optical power supply chip that provides photons to photonic integrated circuits in the same fashion that electrical current is supplied to electronic chips. It is clear that wavelength division multiplexing is critical to the performance advantage of optics over electronics for short-reach applications. The optical power supply of the future will deliver greater than fifty wavelengths of light for WDM and pulses of light for global synchronization. The materials platform for this optical power supply will be determined and implemented within the next decade.

How important is the standardization barrier? Standardization is critical for component vendors to grow the market, and it is critical for the silicon platform to achieve economies of scale. The first release of the Communications Technology Roadmap described a photonics component industry dynamic that has since been dubbed ‘the death spiral’. The scenario goes as follows. Faced with a decreasing sales volume, Company A attempts to secure its near-term viability by customizing products for a current customer. The rationale is that at least that customer will be secure, because Company A is the only source for the customized component. A byproduct of this strategy is that the rest of the market has not designed for that particular customization, and the total addressable market for that product from Company A decreases. Revenues decrease; investment in research and development decrease; and competitiveness to participate in the next generation of technology is decreased. The only way to win through customization is to increase the variety of product offerings and lose the economies of scale.

Standardization leads to consolidation of the industry to a smaller number of vendors in the short term, but to a healthier industry in the long term. Customization (proliferation of variety in product offerings) leads to the ‘death spiral’ for which each company’s total addressable market decreases in the short term with consequent decreasing revenues, research and new products. If the silicon photonics platform becomes ubiquitous, as it has for electronics, it will offer a common set of components that will serve not only one application, but multiple applications in multiple markets.

Standardization of the silicon photonics platform is essential. An integrated circuit chip represents >\$100M in development costs that must be recovered by rapid deployment to the market with high volume sales. The low-cost manufacturing infrastructure for silicon chips requires that a standard set of tools and processes be employed. Mere substitution of silicon for the diversity of materials used in the optoelectronics industry is not a viable option. The huge investment over the past two decades in silicon fabrication knowledge and infrastructure must be leveraged to give silicon photonics a scalable advantage in both factors of the performance/cost metric.

What are the performance metrics for success? Electronic–photonic convergence is the technology train on which the emergence of silicon photonics is riding. Power efficiency, bandwidth, latency, footprint and functionality create the systems performance boundaries. The critical devices for an integrated microphotonic chip are the waveguide, the modulator and the photodetector. High index contrast is essential for electronic–photonic convergence to enable dimensional scaling to  $>10^6$  devices/chip. The Si/SiO<sub>2</sub> system gives a refractive index ratio of 3.5/1, the highest index contrast ever employed in a photonic integrated circuit. An operating wavelength of  $\lambda = 1550$  nm is ideal for this system, because it is transparent, because germanium photodetectors and modulators are capable and compatible and because it conforms to the silica fiber transparency minimum. For electronic–photonic circuits of high complexity, silicon waveguides should operate with *propagation loss*  $\sim 0.1 \text{ dB/cm}$ . Waveguide-integrated germanium photodetectors exhibit the best performance of any comparable optoelectronic device at 1550 nm. The small size, capacitance and transit distance of the device enable high speed at low applied voltage. The capability of controlled evanescent coupling of light from the waveguide to the detector (absorption along the detector length) yields an extended spectral width for photodetection at theoretical responsivity ( $\sim 1.2 \text{ A/W}$ ). For the most demanding, next-generation applications, the waveguide-integrated photodetector should deliver a *bandwidth  $\times$  external quantum efficiency*  $\sim 75 \text{ GHz}$ . The optical modulator is the most power demanding of the integrated optoelectronic devices. Mach–Zehnder interferometric devices are large with high capacitance and high power dissipation. Ring resonator devices are small, but they must be tuned for wavelength stability. Waveguide integrated electro-absorption devices offer the best combination of small size and spectral stability, but insertion loss must be carefully managed. For advanced computational and imaging applications, the waveguide integrated modulator should be capable of *3dB bandwidth*  $>25 \text{ GHz}$ , with a *switching power*  $<150 \text{ pJ}$ , with an *extinction ratio*  $>5$ .

Silicon photonics enables electronic–photonic convergence. This new functionality is not simply breaking the bandwidth barrier. It is design freedom for the information appliance. It enables a distributed footprint with minimal latency penalty. It is freedom from electromagnetic interference. It is design simplicity, aesthetics and new form factors. It is power efficiency and enhanced connectivity through WDM. Two essential topics are not covered in this state-of-the-art compendium, because no known solutions exist. They are chip packaging and integrated

design. Optoelectronic packaging is still on the fiber optics platform: fiber-pigtailed, discrete devices or small-scale combinations of integrated devices. There is no pluggable chip carrier for electronic–photonic chips without a permanent fiber attach. Design is still based on Maxwell’s equations. There is no higher level of abstraction for a photonic circuit theory to support electronic–photonic circuit simulation tools. These missing pieces will arrive, and we will find that others are required, as well. This book certifies that the basic device building blocks for silicon photonics are in place and well founded.

Will silicon photonics have a societal impact as great as silicon electronics? The early indicators answer ‘Yes’. Interactive data applications such as games and virtual life are becoming dominant forms of discretionary activity. The Information Age version of *boutique cache* is ownership of an ideal virtual domicile with select virtual accoutrements, friends and pets. This transformation has even taken the form of buying virtual clothing for virtual pets with real currency. This emerging infrastructure drives markets for vast hardware to support on-demand communications from server and storage machines and to imaging appliances. The mobile appliance market reached 1B units in 2006 to serve the new interactive citizen. Real-time, virtualization computation is beyond the capability of current hardware, and the future will feature parallel arrays of special purpose nodes to efficiently create second worlds. Civilization has progressed using natural materials to build shelter, to build agricultural implements, to build weapons, to build factories and now to build virtual worlds. One can only wonder what direction civilization will take with this new ability to satisfy basic human needs within the virtual world of information.

**Professor Lionel C. Kimerling**

# About the Editor

## **Graham T Reed BSc (Hons), PhD, FIEE, CEng University of Surrey, UK**

Graham Reed is Professor of Optoelectronics and Head of Department, Electronic Engineering, at the University of Surrey in the UK. He graduated in 1983 with a First Class Honours degree in Electronic and Electrical Engineering. Subsequently he obtained a PhD in Integrated Optics in 1987. After a brief period as leader of the Electro-Optics Systems Group at ERA Technology Ltd, he joined the University of Surrey in 1989, where he established the Silicon Photonics Group. As such this was one of the pioneering groups in silicon photonics, and has made a significant impact upon the state of the art. The group is currently the leading group in the UK in this field, and Professor Reed is acknowledged as the individual who initiated research on silicon photonic circuits and devices in the UK. The work has been carried out with collaborators from all around the world, both from academic and industrial institutions. Professor Reed has published extensively in the international scientific literature, has contributed presentations to numerous international conferences both as a submitting and an invited speaker, and has served on a variety of international committees.

# List of Contributors

## *Chapter 1*

### **Richard Soref**

Air Force Research Laboratory, Sensors Directorate, Electromagnetics Technology Division, Hanscom AFB, MA 01731, USA. Email: Richard.Soref@hanscom.af.mil

## *Chapter 2*

### **G. Z. Mashanovich**

Silicon Photonics Group Manager, School of Electronics and Physical Sciences, Advanced Technology Institute, University of Surrey, Guildford, GU2 7XH, UK. Email: g.mashanovich@surrey.ac.uk

### **G.T. Reed**

Professor of Optoelectronics, Head of Department of Electronic Engineering, School of Electronics and Physical Sciences, Advanced Technology Institute, University of Surrey, Guildford, GU2 7XH, UK. Email: g.reed@surrey.ac.uk

### **B. D. Timotijevic**

Research Student, School of Electronics and Physical Sciences, Advanced Technology Institute, University of Surrey, Guildford, GU2 7XH, UK. Email: b.timotijevic@surrey.ac.uk

### **S. P. Chan**

Intel Technology (M) Sdn. Bhd., Kulim 09000, Malaysia.

## *Chapter 3*

### **Dennis. W. Prather**

Professor of Electrical and Computer Engineering, Department of Electrical and Computer Engineering, University Of Delaware, 140 Evans Hall, Newark, DE 19716, USA. Email: dprather@ee.udel.edu

### **Shouyuan Shi**

Electrical and Computer Engineering Department, University Of Delaware, 140 Evans Hall, Newark, DE 19716, USA.

### **Janusz Murakowski**

Electrical and Computer Engineering Department, University of Delaware, 140 Evans Hall, Newark, DE 19716, USA.

**Garrett Schneider**

Electrical and Computer Engineering Department, University Of Delaware, 140 Evans Hall, Newark, DE 19716, USA.

**Ahmed Sharkawy**

Assistant Research Professor of Electrical Engineering, Electrical and Computer Engineering Department, University Of Delaware, 140 Evans Hall, Newark, DE 19716, USA.

**Caihua Chen**

Electrical and Computer Engineering Department, University Of Delaware, 140 Evans Hall, Newark, DE 19716, USA.

**BingLin Miao**

Electrical and Computer Engineering Department, University Of Delaware, 140 Evans Hall, Newark, DE 19716, USA.

***Chapter 4*****G. T. Reed**

Professor of Optoelectronics, Head of Department of Electronic Engineering, School of Electronics and Physical Sciences, Advanced Technology Institute, University of Surrey, Guildford, GU2 7XH, UK. Email: g.reed@surrey.ac.uk

**F. Y. Gardes**

Research Student, School of Electronics and Physical Sciences, Advanced Technology Institute, University of Surrey, Guildford, GU2 7XH, UK. f.gardes@surrey.ac.uk

**G. Z. Mashanovich**

Silicon Photonics Group Manager, School of Electronics and Physical Sciences, Advanced Technology Institute, University of Surrey, Guildford, GU2 7XH, UK. Email: g.mashanovich@surrey.ac.uk

**C. E. Png**

Senior Research Engineer, The Institute of High Performance Computing, A\*Star 1 Science Park Road, #01-01 The Capricorn Singapore Science Park 2, Singapore 117528. Email: pngce@ihpc.a-star.edu.sg

***Chapter 5*****Bahram Jalali**

Professor of Electrical Engineering, Electrical Engineering Department, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095-1594, USA. Email: jalali@ucla.edu

**Dimitris Dimitropoulos**

Electrical Engineering Department, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095-1594, USA.

**Varun Raghunathan**

Electrical Engineering Department, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095-1594, USA.

**Sasan Fathpour**

Electrical Engineering Department, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095-1594, USA. Email: sasan@ee.ucla.edu

***Chapter 6*****A. P. Knights**

Assistant Professor, Department of Engineering Physics, McMaster University, Hamilton, Ontario L8S 4L7, Canada. Email: aknight@univmail.cis.mcmaster.ca

**J. D. B. Bradley**

Department of Engineering Physics, McMaster University, Hamilton, Ontario L8S 4L7, Canada.

***Chapter 7*****Ansheng Liu**

Intel Corporation, 2200 Mission College Blvd., SC12-326, Santa Clara, CA 95054, USA. Email: ansheng.liu@intel.com

**Nahum Izhaky**

Intel Corporation, S. B. I. Park Har Hotzvim, Jerusalem, 91031, Israel.

**Ling Liao**

Intel Corporation, 2200 Mission College Blvd., SC12-326, Santa Clara, CA 95054, USA. Email: ling.liao@intel.com

***Chapter 8*****Cary Gunn**

Co-Founder, Chief Technology Officer, Luxtera, Inc., 1819 Aston Avenue, Suite 102, Carlsbad, CA 92008, USA. Email: cary@luxtera.com

***Chapter 9*****Richard Jones**

Intel Corporation, Photonics Technology Lab, 2200 Mission College Boulevard, SC12-326, Santa Clara, CA 95054, USA. Email: richard.jones@intel.com

**Haisheng Rong**

Intel Corporation, Photonics Technology Lab, 2200 Mission College Boulevard, SC12-326, Santa Clara, CA 95054, USA.

**Hai-Feng Liu**

Intel Corporation, Photonics Technology Lab, 2200 Mission College Boulevard, SC12-326, Santa Clara, CA 95054, USA.

**Mario Paniccia**

Intel Fellow, Intel Corporation, Director, Photonics Technology Lab, 2200 Mission College Boulevard, SC12-326, Santa Clara, CA 95054, USA. Email: mario.paniccia@intel.com

# Acknowledgements

The editor is grateful to all of the contributing authors, who have combined their vast collective knowledge and intellect to produce the chapters that comprise this text. I am also particularly grateful to my colleagues, staff and students, both past and present, for making day to day life both interesting and challenging, and thus, either directly or indirectly, shaping my reasons for the production of this book. For their very specific help in putting together the chapters of this book that bear my name, I must particularly thank Dr Goran Mashanovich, and Fred Gardes, without whom these chapters may never have seen completion.

# 1

## Introduction: The Opto-Electronic Integrated Circuit

Richard Soref

The history, present status and future prospects of silicon-based opto-electronic integrated circuits (OEICs) are reviewed here in order to provide a framework for the state-of-the-art discussions in this book. Before beginning the survey, let's consider some of the terminology used in opto-electronics. The Optical Society of America decided, after much debate, to replace many usages of the historic term 'optical' by the adjective 'photonic,' although 'optical' still has strong currency. The development of silicon 'photonics' is motivated largely by the desire to develop silicon-based photonic integrated circuits (PICs) since the functionality of a circuit is more diverse and significant than that of an individual photonic component. Components can be categorized as passive or active. Passive means that the optical function is fixed and constant. Active means that the function is variable and responsive to an external actuation-or-control signal. Lightwave emitters, amplifiers, detectors and modulators are active. The term 'active' is broad because it denotes a host of physical interactions such as electro-optical, thermo-optical, acousto-optical, magneto-optical, electro-mechanical-optical, chemical-optical, bio-optical and optical-optical (which includes linear and nonlinear opto-optics).

If we consider active PICs generally, we find that the actuation-and-control signals are usually applied to the silicon chip (with bonded wires, for example) from off-chip devices—at least that is how it has been done in the past. The Si-based opto-electronic integrated circuit (OEIC) is a very important special case of an active PIC in which the electronic controllers and drivers are integrated 'seamlessly' with the optical components in the same chip. This is an electronic-and-photonic integrated circuit (EPIC) in silicon: a chip-scale marriage of electronics ICs and photonic ICs. Because the 'EPIC chip' is generally associated with the EPIC program of the US Defense Advanced Research Projects Agency (DARPA), I shall favor the OEIC term in order to avoid linking the acronym to a sponsor.

The thesis of this chapter is that OEICs will, after further R&D, become the dominant form of active PICs, the one with the greatest global impact. That's why I have chosen to focus my

attention here upon OEICs. The expectation is that OEIC chips will be very cost-effective, compact, reliable, efficient, and highly integrated—that they will solve communication problems not resolvable by optics or electronics alone. There is also the synergy speculation that new functionality will be obtained by combining optics and electronics as described here. I believe that readers of this chapter can help make this OEIC dream a reality.

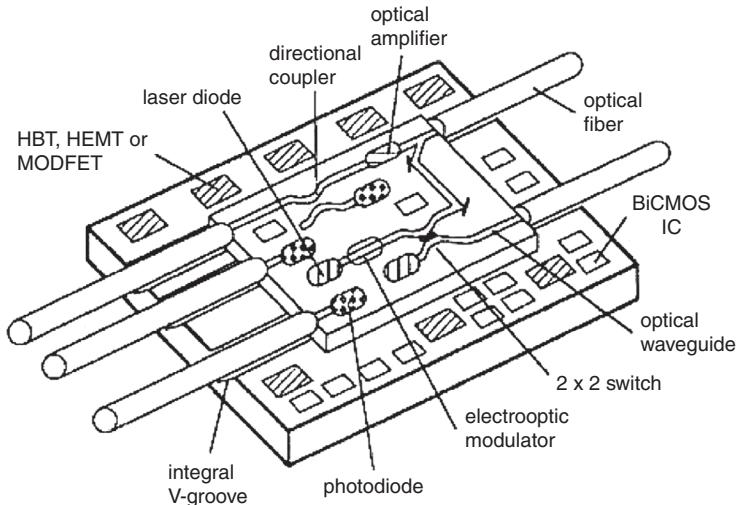
## 1.1 A Few Words About History

Silicon photonic components were conceived in the mid 1980s [1, 2], and those early experiments dealt with waveguides and  $2 \times 2$  electrooptical switches. As more people entered the field, numerous practical component structures were invented and tested. Many of the important possibilities for silicon-based photonics were recognized at the outset: that waveguides made of undoped crystal silicon could, in principle, have low propagation loss at the 1330 and 1550 nm fiber-optic telecomm bands because of silicon’s transparency at wavelengths longer than its 1200 nm indirect band-gap wavelength, and that waveguided active-and-passive photonic components could be interconnected on a chip to create a PIC. To construct an active PIC, metal wires from one (or more) electrical IC chips can be joined to the PIC chip; but that arrangement is neither cost effective nor compact, and electrical parasitics are present. The OEIC provides the best photonic control technique because electrical drivers are connected internally to photonics (intra-chip) over short paths. Parasitics are low and the scale of integration is high compared with the multi-chip.

For active control of light, the 1986 paper [2] proposed the free-carrier plasma effect and acousto-optic diffraction in Si. It was suggested that plasma-related electro-refraction (ER) and electro-absorption (EA) would arise from generating electrons and holes by above-gap light (optical pumping) or from carrier injection (electrical pumping). Relatively weak EA and ER from the Si Kerr and Franz-Keldysh electric-field effects were also mentioned, and tailoring of the waveguide-core bandgap energy through the use of SiGe alloys was proposed. In the 1980s, modulation via carrier accumulation and depletion were not known and it was only in 2004 that those methods were reported.

The 1986 paper cited two motivations for silicon photonics R & D that are still valid today: (1) leveraging the infrastructure of the huge, global silicon microelectronics industry in order to make highly sophisticated silicon photonic devices, and (2) integrating silicon photonics monolithically on a chip containing fast VLSI silicon electronics so as to create an OEIC. Regarding leverage, the unstated assumption was that the knowledge, experience, capital investment, manufacturing tools, and design automation developed for silicon electronics would make possible high-volume manufacture of PICs and OEICs at low cost. This high-volume focus is very strong today.

By 1993, the vision of silicon OEICs came into sharper focus as exemplified in the superchip ‘thought experiments’ of Abstreiter [3] and the present author [4], who viewed OEICs as extending the reach of electronic ICs by offering new functionality with high performance. As shown in Figure 1.1, integration of photonics with CMOS was proposed, as was integration with Si bipolar, BiCMOS, and SiGe/Si heterobipolar transistors. Despite the promise, little was done on actual OE integration during 1993–2003. The turning point came in 2004 when integration experiments commenced in SOI CMOS foundries due to government investment. Progress has also been slow on photonic integration with bipolar electronics. This is both a challenge and opportunity.



**Figure 1.1** The silicon OEIC superchip as proposed in 1993 [4]

In the early years, OE integration with CMOS was felt to be important and is now viewed as supremely important. Today, CMOS captures about 90% of integrated circuit revenues, and the world-wide investment in CMOS fabrication facilities amounted to 500 billion dollars in 2005. With ingenuity, opto-electronic engineers could capture this tremendous CMOS momentum. Their hard work could spawn a huge ‘CMOS photonics’ industry (Luxtera’s term).

For the past thirty years, Group III–V semiconductor devices have had a strong presence in photonics, and today, III–Vs are the mainstays for efficient LEDs, laser diodes, semiconductor optical amplifiers, photodetectors, electroabsorption modulators and electrooptical switches, either resonant or nonresonant. This dominance holds for the near and mid infrared. It will be difficult for silicon devices or IV–IV devices to supplant III–V devices in most of those photonic areas. In the 1980s, it was felt that devices made solely from group IV materials could not challenge the III–V dominance; however, the situation evolved during the next two decades and in 2006 it appears that IV–IVs can make real inroads into the active device scene and will eventually compete with III–Vs in some areas.

Cleverness is required to overcome the limitations imposed by the indirect bandgap of silicon. Such adroitness has been the hallmark of Group IV researchers who have worked diligently towards practical LEDs, amplifiers, detectors, modulators, switches and (yes) lasers, discovering new physics in the process and/or applying known physics in novel ways. They are on the ‘verge of success’. In 2006, the US Air Force Office of Scientific Research (AFOSR) launched a multi-university research initiative (MURI) to create silicon-based lasers—a project that will emphasize intrinsic approaches such as nanostructured silicon, light-from-germanium, quantum-cascade lasing and band-to-band lasing in GeSn (discussed below) as well as extrinsic approaches such as erbium-oxygen complexes in silicon.

In 1985, a ‘network’ made of silicon did not exist, but the component field blossomed in the ensuing 20 years. People built and tested a dozen different waveguide types (of which silicon-on-insulator is the leader) and they constructed the active devices mentioned above. They made directional couplers, optical power splitters, power combiners, TE-TM converters, TE-TM

splitters, two- and three-dimensional tapered-waveguide couplers, surface-grating couplers, echelle gratings, vertical couplers, arrayed-waveguide filters, add-drop multiplexers, resonant filters, transversal filters, variable optical attenuators, multimode to single-mode transitions, active-cladding devices, magneto-optical isolators, quantum-layered devices, self-assembled superlattices, microring resonators, slotted resonators, Si/organic composites, Si/ferroelectric composites, wavelength converters, Raman amplifiers, Raman lasers, and four-wave mixers. The goal now is to introduce standardization, to create design libraries and design tools for manufacturing the ‘important set’ of components—the so-called complete suite. As research goes forward, existing components become refined and new ones are invented—a virtuous cycle.

I shall close this section with a few words about OEICs that can operate in the visible and near-infrared at wavelengths shorter than the 1.2  $\mu\text{m}$  indirect-bandgap wavelength of silicon. Silicon photodiodes are ideal for detecting these shorter wavelengths, and these PDs have been deployed in optical receivers which can be part of an OEIC. Although silicon waveguides are opaque over the visible and very-near IR, transmission of 0.4–1.2  $\mu\text{m}$  light in ‘silica-on-silica waveguides’ (germania-doped- $a$ -SiO<sub>2</sub>/SiO<sub>2</sub>/Si) and silicon-oxynitride-on SiO<sub>2</sub>-on-Si waveguides was demonstrated years ago. As to short-wavelength sources, III-V emitters are available, and even a silicon PN junction diode when reverse-biased will emit yellow light, while forward biasing of a silicon PIN diode produces 1.1  $\mu\text{m}$  bandedge light. It is intriguing that a 550 nm Si P<sup>+</sup>NN<sup>+</sup> avalanche emitter can be internally modulated at 10 Gb/s [5], perhaps obviating the need for lasers in short interconnects. In summary, the components needed for a short-wave OEIC are here today and there are good prospects for such a chip. In fact, we can say that the CCD and CMOS imaging chips used ‘everywhere’ in video cameras are one type of short-wave OEIC. It would be stretching a definition to say that an LCD display panel containing silicon thin-film transistors is an ‘OEIC on glass’.

## 1.2 The Possibilities for OEICs

The development of OIECs is motivated by potential applications in communications (mainly optical interconnects), RF/microwave/mmwave signal processing, digital signal processing, smart sensing, environmental monitoring, imaging, biomedicine, spectroscopy, lab-on-a-chip, optical logic, and transistor IC testing, among others. New ‘OEIC-centric’ computer architectures may include optical backplanes, optical data buses and optical clock distribution. Mario Paniccia proposed silicon micromachining (Si V-grooves, etc.) for smart OIEC packaging [6], and Cary Gunn capsulized OEIC motivation as: reduced system cost, enhanced product yields and novel system opportunities [7].

The applications are diverse, and happily there is a diversity of approaches that can be taken to implement OEICs for specific needs. The possibilities include: (1) monolithic or heterogeneous integration; (2) lasers and LEDs located on the chip or off the chip, or both on and off; (3) optical interfacing with free-space or fiber-guided lightbeams; (4) single-wavelength operation or wavelength-multiplexed operation with or without code-division multiple access; (5) operation with digital signals and/or analog signals; (6) photonic integration with CMOS or BiCMOS or SiGe CMOS or bipolar or heterobipolar; (7) use of multiple waveguiding layers or a single guide layer; (8) use of a substrate consisting of Si or SOI or strained SOI or SGOI or GOI or even SiCOI; (9) use of photonic components constructed from silicon and/or germanium and/or group IV alloys, both binary and ternary; (10) use of bulk photonic structures and/or quantum-confined structures and/or photonic-crystal structures; (11) fabrication of an

electronics layer below or above photonic layers; (12) operation at a wavelength in the visible, near IR, mid IR, long wave IR, or far IR. Let's explore these briefly.

What is monolithic and what is heterogeneous? The distinction is not entirely clear. Monolithic refers to all-silicon construction or to 'all-in-group-IV' heterostructures. III-V and II-VI devices bonded to Si exemplify heterogeneous (also known as hybrid) integration. Epitaxial III-Vs on Si 'might be' monolithic. An organic or ferroelectric cladding on a silicon waveguide is probably monolithic. Both mono- and hetero-integration are useful. Whichever is most cost-effective in a particular application is going to win. Monolithic will probably be best in the high-volume future.

Electrically pumped silicon-based lasers are the holy grail in this photonics field. But there are two issues here: demonstrating that the laser works, and showing that it can be manufactured with high yield at low cost. If not manufacturable, the glamorous device may not be widely adopted.

For the OEIC, electrically pumped and/or optically pumped lasers can be located on-chip and/or off-chip. Electrical pumping on-chip would seem best, although there is a caveat here: the lasers must not consume 'too much' power, otherwise the modern VLSI chip, which is probably limited today by heat dissipation rather than transistor scaling, will tilt into a 'thermal meltdown' situation. To summarize: the useful options for future OEICs are: electrical silicon lasers on-chip, hybrid integration of electrical III-V lasers on chip, an off-chip laser that optically pumps several silicon Raman lasers on-chip, and off-chip lasers of various kinds that communicate via fiber optics with an OEIC containing no lasers.

Optical interfacing refers to the coupling of free-space or fiber-optic guided lightbeams into and out of the OEIC. Free and/or guided coupling are feasible by means of on-chip surface gratings and the tapered waveguides developed in recent years.

Digital signaling with time-division multiplexing can give high OEIC speeds; however, ultra-high data rates of, say, 50–200 Gb/s, are going to require wavelength-division multiplexing and demultiplexing in the OEIC transceiver. Code division multiple access (CDMA) is currently proposed as a useful adjunct to WDM to multiply the wavelength channel capacity. For analog situations, the broadband OEIC can perform a myriad of RF functions such as agile RF filtering. The frequency response of group IV transistors extends now into the millimeter-wave region, and Michal Shur announced emission and detection of terahertz radiation in silicon 'plasma transistors'.

The panoply of Si-based electronics now includes SiGe and SiGeC hetero bipolar transistors, SiGe BiCMOS and MOSFETs in silicon-germanium-on-insulator (SGOI). 100 Gb/s bipolar logic exists today. Although the scope of bipolar markets does not rival that of CMOS, I believe that there are application-specific niches for purely bipolar OEICs that can be exploited in the coming years. Another valuable class of OEICs will combine bipolar and CMOS electronics on-chip for mixed-signal analog-and-digital use (or all-digital application).

Presently OEICs use SOI, and there are strained-layer photonic components on the horizon to enhance OEIC capabilities. To accommodate those components, buffer layers (virtual substrates) on SOI may be required, or the OEIC could be based upon SSOI, SGOI, GOI or SiCOI.

As spelled out below, the recent advent of tin-containing alloys in group IV devices opens up possibilities for efficient monolithic lasers, LEDs, amplifiers, modulators and detectors that employ direct-bandgap valence-to-conduction transitions or intersubband transitions. It seems reasonable to speculate that Sn-alloy photonics will eventually be a major presence in OEICs.

Nanophotonics, discussed below, relies upon photonic crystals, plasmon optics, nanocrystals, carbon nanotubes, as well as alloy-based quantum wells, wires, and dots. For ULSI, there

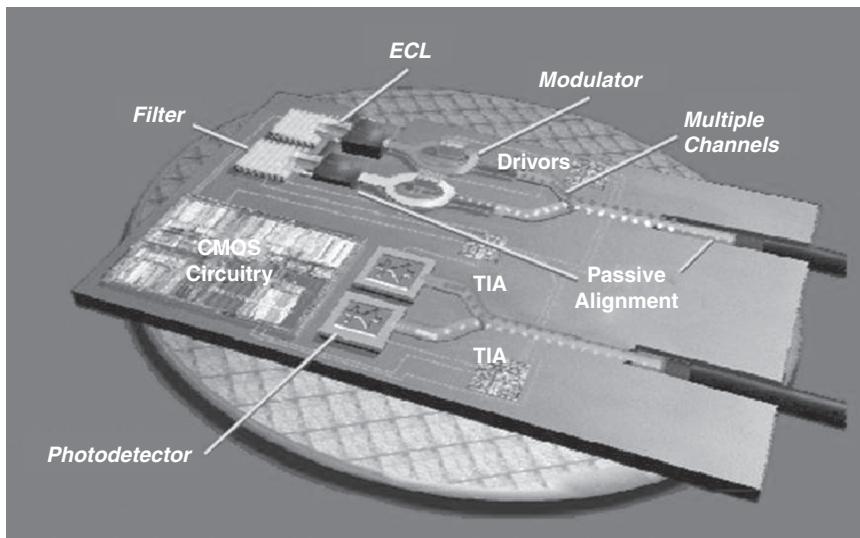
is motivation to use photonic nanostructures in the 2010 generation of OEICs. Whether the nanostructures are truly CMOS compatible remains to be seen. That probably hinges upon the complexity of the required processing.

Described below, the multi-layer vertical three-dimensional integration of silicon waveguides and photonic devices in an OEIC is an excellent possibility currently under study. Other options are a single photonic layer above the transistors, or a subterranean photonics layer below the electronics.

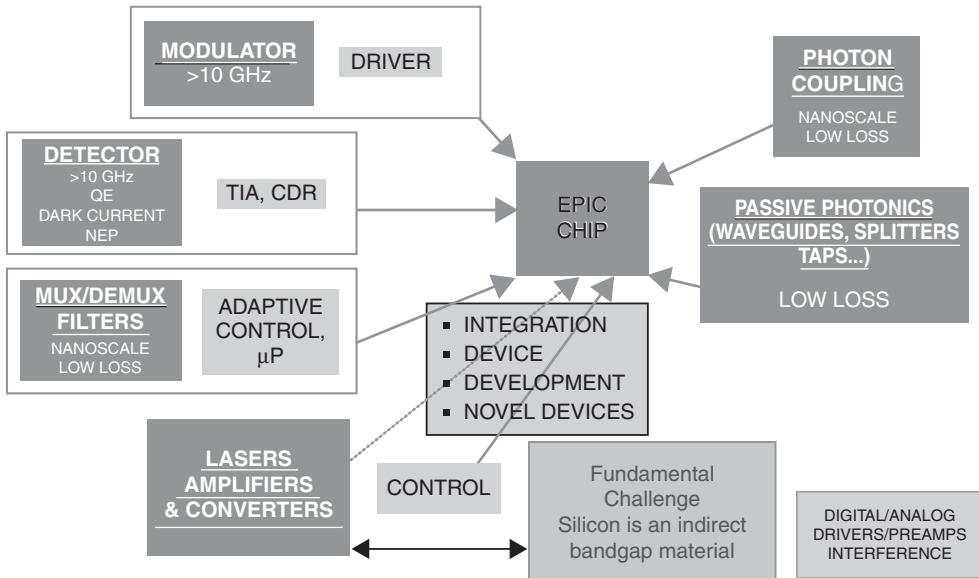
The wide infrared spectrum, as mentioned below, is available for exploitation in OEICs. If and when long-wave infrared (LWIR) OEICs are implemented, they will bring new performance and economics to infrared technology.

### 1.3 The Present Status of OEICs

University research has consistently played a key role in silicon photonics; however, industry and governments are now spurring this field forward at an accelerated rate. An industry leader, Intel Corporation, foresees a cost-driven transition to optical interconnects in computers that could ‘revolutionize future servers and enterprise networks’ [6, 8, 9, 10]. Intel chose to invest in silicon photonics because they felt OEICs could alleviate electronic bottlenecks, enabling ultrafast processing in a new generation of chips-and-computers—a generation more capable and cost effective than ever before. This will be a convergence of computing and communications. ‘Today, optics is a niche technology. Tomorrow, it’s the mainstream of every chip we build’—(Patrick Gelsinger, Intel Sr. Vice President). Intel has already achieved excellent results in silicon photonics. Much of the Intel work is aimed toward the monolithic chip illustrated in Figure 1.2. They have the capability to fulfill that vision.



**Figure 1.2** Perspective view of the future-generation monolithic silicon OEIC proposed by Intel Corporation (courtesy of Dr M. Paniccia)

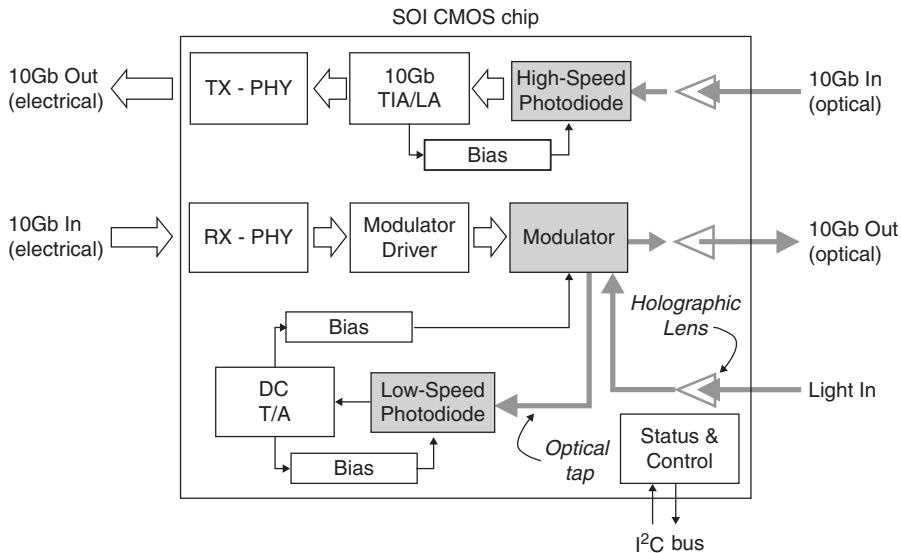


**Figure 1.3** The challenge of electronic and photonic integrated circuits in silicon as envisioned on DARPA EPIC (courtesy of Dr Jag Shah)

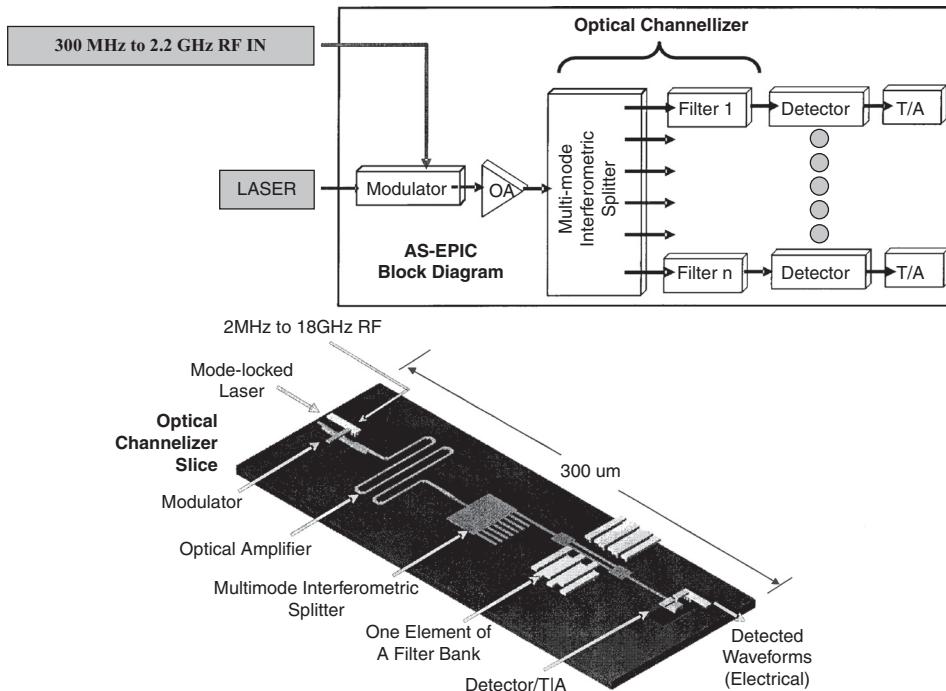
Government agencies, especially AFOSR and DARPA, believe that highly developed Si OEICs will meet both military and commercial needs. Both organizations have invested resources to advance silicon photonics technology by sponsoring a series of research contracts. AFOSR (Dr Gernot Pomrenke) has funded small-business technology-transition research on silicon-based photonic components and is currently launching a MURI to develop silicon-based lasers. DARPA (Dr Jagdeep Shah) has selected two teams to develop OEICs in commercial state-of-the-art SOI CMOS production facilities: one team led by BAE systems, the other by Luxtera Incorporated (website: [www.darpa.mil/mto/epic](http://www.darpa.mil/mto/epic)). The preliminary results on this four year EPIC project are quite encouraging. The reader is referred to [11–17] for more details.

The challenge of EPIC is illustrated in Figure 1.3 which shows the OE pairing envisioned for the new chip [11, 12]. A more specific EPIC goal is shown in the schematic diagram of Figure 1.4 which presents the Luxtera team's digital-signal Phase 1 OEIC fiber-optic transceiver [18–21]. Using the 130 nm Freescale production plant, the Luxtera team has created in a single layer of silicon a group of excellent components including a holographic surface-grating coupler for fibers, a carrier-depletion 10 Gb/s modulator, a trimmable three-channel demultiplexing filter and a  $1 \times 2$  electrooptical switch.

For electronic warfare application, the BAE systems team is working on an analog-signal OEIC to be constructed in their rad-hard 90 nm Manassas foundry. As shown in Figure 1.5, this is a miniaturized RF-channelizer optical receiver actuated by an off-chip laser [22–23]. An incoming broadband RF signal modulates the lightbeam on-chip, and the optical filter banks pick off spectral ‘slices’ of the RF waveform. The inset in Figure 1.5 shows the OE layout for one slice of three in Phase I. The RF energy in each ‘bin’ is determined by a dedicated on-chip photodetector. BAE has already demonstrated an optical filter that has a



**Figure 1.4** Schematic of the 10Gb/s 1.55μm silicon OEIC transceiver built by the Luxtera Inc. team for EPIC (courtesy of Dr C. Gunn)



**Figure 1.5** Schematic of the RF-channelizing 1.55 μm silicon OEIC receiver being built by the BAE Systems team for EPIC (courtesy of Dr M. Grove)

1 GHz electrical bandwidth (a world record), 0.35 dB/cm strip waveguides, two independent levels of Si wayguiding in a double-SOI structure, a *pin*-diode carrier-injection microring-resonator analog modulator, and a Ge-on-Si Franz-Keldysh intensity modulator. As these two projects go into Phases II and III, more sophisticated OEIC products will emerge.

## 1.4 Silicon-based GeSn and SiGeSn Technology

Because of the 4% lattice mismatch and thermal expansion mismatch, Ge and Ge-rich SiGe do not grow well over ‘large’ areas of silicon, although local-area epitaxy works well, even in Si trenches. In the past, people have grown a strain-relieved buffer of SiGe on silicon for use as a virtual substrate. This strategy can be extended to tin alloys. Since the chip of the future may contain SiGe, Ge, GeSn, SiGeSn and other column IV materials, I would rather use the term ‘group IV photonics’ to describe that network, rather than silicon photonics. The increased use of germanium and the advent of low-defect SiGeSn alloys are two important recent developments in group IV photonics. Currently, Arizona State University is the leader in GeSn techniques (see the references cited in [14–17]). Via low-temperature chemical-vapor deposition, a strain-relaxed buffer layer of GeSn or of SiGeSn has been grown directly upon silicon, and the layer has a relatively low defect density at its top surface, making that surface a good template for subsequent growth of coherent, strained-layer heterostructures such as Ge/GeSn multi-quantum wells (MQWs). To make thick MQWs, strain balance is employed—a new paradigm in which the wells and barriers have alternating tensile and compressive inplane strain with respect to the buffer.

The ‘revolutionary’ aspects are; that the QW layer in Ge/GeSn or GeSn/SiGeSn can have a direct bandgap for tin concentrations above 10%, for example, and that the QW band alignment can be Type I in the GeSn/SiGeSn system. The conduction-subband to valence-subband wavelength can be as short as 1550 nm in MQWs. More typically, in bulk heterostructures (layer thickness >10 nm), the band-to-band wavelength is in the 1.8–10  $\mu\text{m}$  region. Tin-containing heterostructures can be built that mimic those in GaAs/AlGaAs, InP/InGaAsP (and related) III–V devices. The implication of the above discussion is that the group IV Ge/GeSn/SiGeSn heterodevices are (or could be) equivalent in their lasing, modulation, and detection to those of well-known III–V compound semiconductor devices. Thus, in effect, a compound semiconductor technology becomes available in group IV.

The open questions for OEICs are: whether the Sn-containing heterostructures are ‘acceptable’ (noncontaminating, etc) in a CMOS foundry and whether a CVD GeSn growth step can be included in the CMOS production process (as is UHV CVD SiGe). If the answer to both questions is yes, then new horizons appear in Si-based OEICs for monolithic integration of efficient sources and detectors.

## 1.5 OEICs for the Near, Mid and Far Infrared

In a recent paper, I identified several kinds of silicon-based waveguide structures that can, in theory, provide low-loss propagation over large portions of the wide infrared spectrum stretching beyond 1.2  $\mu\text{m}$  [15]. For the near infrared, these types are silicon-on-insulator, silicon-on-sapphire and silicon-on-Si<sub>3</sub>N<sub>4</sub>. The near- and mid-range IR are handled by suspended silicon-membrane rib waveguides (air-clad above and below), while for mid- and

long-wave operation, the heterostructured rib of Ge-on-Si or GeSn-on-Si will do the job. To obtain coverage of the entire 1.2–100  $\mu\text{m}$  wavelength spectrum, a hollow, rectangular, air-filled waveguide should work well, provided that the inner claddings are alternating layers of SiGe and Si. Proof is still needed that these structures have low loss as predicted. Assuming a favorable outcome, then I foresee that waveguided Si-based infrared networks are feasible across 1.2–100  $\mu\text{m}$ .

For long-wave operation, most of the telecomm passive waveguided components can be scaled up in size from their 1.55  $\mu\text{m}$  dimensions and can be fabricated with lower-resolution lithography. Narrow-gap semiconductors will serve on-chip as mid-wave/long-wave detectors and sources. Those could be monolithic GeSn/Ge structures or hybrid-integrated III–V devices. Transistor ICs can be integrated with the long-wave photonics in the same manner as for telecomm OEICs. However, the OEIC may have to be cooled when the wavelength exceeds 10  $\mu\text{m}$  and the sources are on-chip. For those situations, special low-temperature transistors would be required. The inference of this discussion is that silicon-based LWIR OEICs, when properly designed, can operate anywhere within 1.2–100  $\mu\text{m}$  (this is a new paradigm, awaiting demonstration). As to applications of these OEICs, some will be new and others much like known applications: free-space infrared wireless, sensor fusion, imaging and medical diagnostics, for example.

## 1.6 Opto-Electronic Integration with Ultimate CMOS and Post CMOS

As transistors become more nano-sized, photonic structures must shrink to ‘keep pace’ for OE integration. Transistor scaling has followed an electronics Moore’s law with packing density doubling every 18 months. Work on nano-photonics could, over time, produce a similar decrease in device dimensions, fulfilling a ‘Moore’s law for photonics’, although the scaling of PICs will ‘always’ be behind IC scaling since the ICs ‘started first’.

The semiconductor industry, currently at 130 and 90 nm Fabs, has set ambitious goals for CMOS improvements. Watkins and Bishop [24] discuss 45 nm and ‘below’ CMOS structures. They say that the international roadmap for silicon ULSI predicts that devices at the 45 nm node will be in production by 2010, with 32 nm devices to follow by 2013. Ultimately, CMOS is limited by quantum-confinement effects and electron tunneling through a few atomic layers of oxide. (Similarly, the minimum useable perimeter size for photonic devices is around 10 nm). CMOS represents charge-based computational electronics, and alternatives to such electronics are sought for the 2010–2020 era because CMOS will ‘hit the wall’ then and new types of nanoelectronic devices will be needed to fulfill the Moore’s-law vision of ever-denser integration. ‘Post CMOS’ means that CMOS will still be ubiquitous—very much in use—but that new computing paradigms will have entered the scene. Horst Stormer points out that we shall, at the end of scaling (at the end of Moore’s Law) reach the smallest silicon transistor, which will then be the standard ‘brick’ for all future (charge based) silicon buildings [25]. ‘With the brick a commodity, it all resides with the architect (to make further progress)’, and a lot of progress remains. Beyond lithography lies biology, which may mean molecular self-assembly.

## 1.7 Nanophotonics Integrated with Nanoelectronics

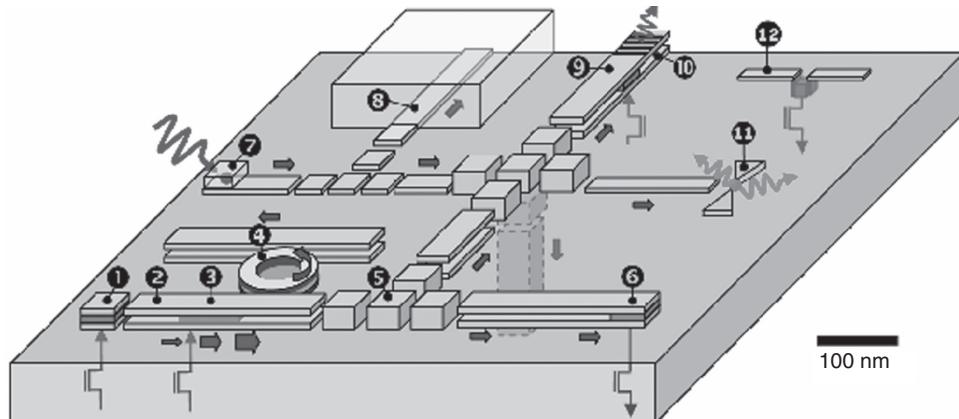
Today, we see researchers using combined microphotonic and nanophotonic techniques, with micro dominant and nano emerging strongly. Regarding micro, the passive and active versions

tend to have an elongated footprint whose length is greater than width. This differs from the ‘squarish’ transistor footprint. Further, the width and height of microphotonic waveguides tend to be of order  $\lambda_0/n$ , the wavelength in silicon—dimensions larger than transistor dimensions. However, that discrepancy does not, of itself, limit the chip-scale OE integration because, for example, the first-generation OEICs reported in 2006 [7] will contain around 50 photonic components and 100 000 transistors—implying that it is natural and acceptable to have the scale of optical integration smaller than that of electronic integration; but in the 2010 OEIC decade, nanophotonics will be the ‘natural partner’ of nanoelectronics.

Optical diffraction limits the mode size and device dimensions in conventional group IV waveguides; however, plasmon optics offers the chance to make devices a factor-of-ten smaller than the diffraction limit—subwavelength in size. Photonic-crystal (PhC) structures including line-defect and self-collimated waveguides also confine the infrared modal fields to deep subwavelength dimensions, and these photonic-lattice devices, together with slow-light dispersion engineering, provide electro-optical interaction lengths that are  $\sim 50 \times$  shorter than those in conventional waveguides. I imagine that group IV slow-light PhCs and plasmonics will become natural companions to nanoelectronics, allowing dense OE integration to proceed. Nanophotonics, of which photonic-crystal and plasmon-optic devices are leading contenders, requires sophisticated electromagnetic (EM) design tools to reach certain design or performance goals for OE application. Yablonovich points out that the desired design is the numerical solution of an ‘inverse problem’ in which one works backward from the goal [26]. A new generation of rational inverse-design algorithms and software will enable such nanophotonics.

Silicon spintronics along with plasmonics have been proposed as nano-approaches to computing for ‘post-CMOS’ silicon chips.. The newly formed Nanoelectronics Research Initiative (NRI) will push forward the computational power of chips. Their areas of research include phonon engineering for heat removal/diversion, directed self-assembly of nanoscale components, spintronics, plasmonics and nano modeling. Silicon spintronics would include spin transistors for logic and nonvolatile storage. Researchers from chemistry and biology will help tap into new ideas. The details of how OE integration will play out in the post-CMOS era are not clear. Conventional photonics could be used—or, one might, for example, attempt to use the same physical principles to actuate the photonics and the computing elements. That would imply a new paradigm such as opto-spintronics, a most difficult task, since silicon, without magnetic impurities, is not highly magneto-optic because its Verdet constant is relatively low. Perhaps direct-gap GeSn alloy heterostructures would be useful for spin transistors since direct-gap GaAs alloys have worked well for spin devices.

A convergence of microphotonic, photonic crystal, nanocrystal, and carbon nanotube techniques is driving the nanophotonics thrust. Plasmonics recently gained momentum due to an AFOSR MURI that was launched in 2004 ([website: www.plasmonmuri.caltech.edu](http://www.plasmonmuri.caltech.edu)) Some of this work is silicon based and aimed towards CMOS compatibility. The optical mode in a silicon dielectric waveguide and the surface plasmon polariton in a plasmonic waveguide are both EM waves. That’s why it is feasible to make an infrared-wave transition from the dielectric waveguide to the plasmonic waveguide, and vice versa, although phase-matching structures are needed to aid the transition. This implies that plasmonics and photonics can be monolithically integrated on silicon. Alternatively, all of the EM signal processing can be done entirely in plasmonics by utilizing a variety of plasmonic components as proposed by Brongersma [27]—an on-chip plasmonic network of the kind illustrated in Figure 1.6. The idea here is that a complete suite of active and passive plasmonic components can be interconnected on-chip.



**Figure 1.6** A ‘futuristic’ Si-based plasmonic IC chip that is proposed by the Stanford University team (courtesy of Dr M. Brongersma); transistors can be integrated here to make a PEIC

For this purely plasmonic case, the integration of transistors on the same chip would create a plasmo-electronic integrated circuit (PEIC). More likely is the combination of photonic and plasmonic components on-chip. I would call this a plasmonic-OEIC or POEIC. I suspect that plasmonics will eventually cover 1.2–100  $\mu\text{m}$ .

The latest trend in plasmonic waveguides is to make strip-like composite structures consisting of a metal–insulator–metal sandwich (MIM), which could be M/Si/M. To implement such guiding, I would propose a suspended silicon membrane (such as SOI locally undercut) that could be a conventional rib guide, clad below and above by air, or the membrane could become a plasmonic guide when coated below and above by thin metal. A plasmon metal such as copper would be better than silver or gold for CMOS compatibility.

## 1.8 Conclusion

The era of silicon-based optoelectronic integrated circuits has just begun and will probably be long-lived. Significant OEIC progress has already been reported by teams at Intel, Luxtera and BAE Systems. Group IV nanophotonics integrated with ULSI SOI CMOS will likely dominate the OEIC field. Like computer chips, silicon OEICs do have the potential to become pervasive in everyday life. Engineers and scientists have a chance to make this potential a reality. If and when the OEIC promise is realized, it will in its course have generated a new and vigorous ‘CMOS photonics industry’ with considerable economic impact.

The techniques for making low-cost high-performance OEICs are varied. They include heterogeneous (hybrid) integration, and most importantly, monolithic integration. Group IV bipolar and heterobipolar electronics can contribute along with CMOS. The sources of light can be on or off the OEIC chip (or off-plus-on). Direct-gap group IV materials are becoming available. The several light beams that are processed by the OEIC chip can travel through free space, to and from the chip, or they can communicate via fiber optics linked to the chip, or they can be generated on the chip. Generally, the OEIC contains a waveguided photonic integrated circuit whose waveguides would be made of group IV semiconductor materials for wavelengths

beyond 1.2  $\mu\text{m}$ . Alternatively, for transmitting visible and near IR light, the OEIC waveguides would consist of silicon oxynitride or silica. An important benefit of group IV is that the silicon OEIC can operate at a wavelength that is anywhere within the vast visible-to-terahertz spectrum that stretches from 0.4 to 150  $\mu\text{m}$ .

## References

- [1] R. A. Soref and J. P. Lorenzo, ‘Single-crystal silicon—a new material for 1.3 and 1.6  $\mu\text{m}$  integrated-optical components’, *Electronics Letters*, **21**, 1985, 953–954.
- [2] R. A. Soref and J. P. Lorenzo, ‘All-silicon active and passive guided-wave components for  $\lambda = 1.3$  and 1.6  $\mu\text{m}$ ’, *IEEE Journal of Quantum Electronics*, **QE-22**, 1986, 873–879.
- [3] G. Abstreiter, ‘Engineering the future of electronics’, *Physics World, UK*, **5**, 1992, 36–39.
- [4] R. A. Soref, ‘Silicon-based optoelectronics’, *Proceedings of the IEEE*, **81**, 1993 1687–1706.
- [5] L. W. Snyman, H. Aharoni and M. du Plessis, ‘A dependency of quantum efficiency of silicon CMOS  $n^+pp^+$  LEDs on current density’, *IEEE Photonics Technology Letters*, **17**, 2005, 2041–2043.
- [6] M. Paniccia, ‘Silicon optoelectronics: applications and recent results’, *SPIE Proceedings* **6115**, San Jose, CA, 2006.
- [7] A. Huang, C. Gunn, G.L. Li, S. Mirsiadi, A. Narasimha and T. Pinguet, ‘A 10 Gb/s photonic modulator and WDM MUX/DEMUX integrated with electronics in 0.13  $\mu\text{m}$  SOI CMOS’, *Proceedings of the IEEE International Solid State Circuits Conference*, San Francisco, CA, 2006.
- [8] N. Izhaky, M. Morse, S. Koehl, O. Cohen, D. Rubin, A. Barkat, G. Sarid, R. Cohen and M. Paniccia, ‘Development of CMOS compatible integrated silicon photonic devices’, *IEEE J. of Selected Topics in Quantum Electronics*, **12** (special issue on silicon photonics), 2006, 1688–1698.
- [9] M. Salib, M. Morse, and M. Paniccia, ‘Opportunities and integration challenges for CMOS-compatible silicon photonic and optoelectronic devices’, *Digest of 1st IEEE International Conference on Group IV Photonics*, Hong Kong, China, 2004, pp 1–6.
- [10] M. Paniccia, A. Liu, N. Izhaky and A. Barkai, ‘Integration challenge of silicon photonics with microelectronics’, *Digest of 2nd IEEE International Conference on Group IV Photonics*, Antwerp, Belgium, 2005, pp 20–22.
- [11] J. Shah, ‘DARPA’s programs on photonic integration’, *Digest of OSA Topical Conference on Integrated Photonics Research and Applications*, San Diego, 2005.
- [12] J. Shah, ‘DARPA’s EPIC program; electronic and photonic integrated circuits on silicon’, *Digest of IEEE 2nd International Conference on Group IV Photonics*, Antwerp, Belgium, 2005.
- [13] R. A. Soref, ‘Silicon photonic components’, invited paper, *Digest of 9th International Symposium on Contemporary Photonics Technology*, Tokyo, Japan, 2006, pp 35–38.
- [14] R. A. Soref, ‘The past, present and future of silicon photonics’, *IEEE Journal of Selected Topics in Quantum Electronics*, **12** (special issue on silicon photonics), 2006, 1678–1687.
- [15] R. A. Soref, S. J. Emelett and W. Buchwald, ‘Silicon waveguided components for the long-wave infrared region’, *Digest of 1st International Optical Microsystems Conference*, Capri, Italy, published in *Journal of Optics A (IOP)*, paper JWA1 UK, **8**, 2006, 840–842.
- [16] R. A. Soref, ‘Recent advances in silicon photonic components’, *Digest of OSA Topical Conference on Nanophotonics for Information Systems*, San Diego, CA, 2005.
- [17] R. A. Soref, ‘Silicon photonics technology: past, present future’, *SPIE Proceedings* **5730**, San Jose, CA, 2005.
- [18] C. Gunn, ‘CMOS photonics technology platform’, *SPIE Proceedings* **6125**, San Jose, CA, 2006.
- [19] C. Gunn, ‘DARPA EPIC program update: Luxtera’s development of a 100 Gb transceiver using CMOS photonics technology’, *Government Microcircuit Applications and Critical Technologies Conference*, San Diego, CA, 2006.
- [20] C. Gunn, ‘CMOS Photonics™—SOI learns a new trick’, *Proceedings of the 2005 IEEE International SOI Conference*, Honolulu, HI, 2005.
- [21] Cary Gunn, ‘CMOS Photonics technology: Enabling optical interconnects’, *Digest, 2nd IEEE International Conference on Group IV Photonics*, Antwerp, Belgium, 2005.
- [22] L. C. Kimerling, A. B. Apsel, M. Beals, D. Carothers, Y. Chen, T. Conway, D. Gill, M. Grove, C. Hong, M. F. Lipson, J. Liu, J. Michel, S. S. Patel, A. T. Pomerene, M. Rasras, D. K. Sparacinc, A. E. White and C. Wong, ‘Recent advances in CMOS compatible integrated photonics’, invited paper 6125-02, *SPIE Proceedings* **6125**, San Jose, CA, 2006.

- [23] M. J. Grove, ‘Recent advances in CMOS-compatible integrated photonics’, *Government Microcircuit Applications and Critical Technologies Conference*, San Diego, CA, 2006.
- [24] J. J. Watkins and D. J. Bishop, ‘Fabrication of sub-45-nm structures for the next generation of devices: a lot of effort for a little device,’ *MRS Bulletin Magazine*, **30**, 2006, 937.
- [25] H. L. Stormer, ‘Silicon forever! Really?’ *Proceedings of the IEEE ESSDERC*, Grenoble, France, 2005, pp 3–6.
- [26] E. Yablonovich, ‘Silicon nano-photonics: where the photons meet the electrons’, *Proceedings of the IEEE ESSDERC*, Grenoble, France, 2005, pp 23–25.
- [27] M. Brongersma, D. A. B. Miller and S. Fan, ‘Chip-scale plasmon-enabled nanophotonics’, *1st Annual Progress Review on AFOSR MURI* (H. Atwater Principal Investigator), Stanford, CA, 2006, ([www.plasmonmuri.caltech.edu](http://plasmonmuri.caltech.edu)).

# 2

# Silicon Photonic Waveguides

G. Z. Mashanovich, G. T. Reed, B. D. Timotijevic and S. P. Chan

## 2.1 Introduction

The essential building block of every photonic circuit is a waveguide. Waveguides can be classified by the number of dimensions in which the light is confined. A planar waveguide confines the light in one dimension and therefore is a 1D waveguide. Channel waveguides (or 2D waveguides) confine the light in two dimensions. There also exist structures (photonic crystals) that confine light in the three dimensions.

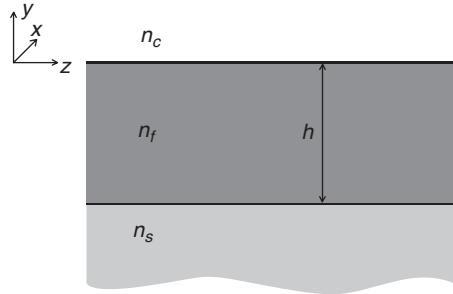
The first silicon waveguides were reported in the mid 1980s, in silicon on doped silicon [1, 2], silicon on sapphire [3], silicon germanium [4], and Silicon on Insulator (SOI) [5, 6]. The SOI platform, first reported in 1989, has by far, become the most popular of the four waveguide systems. As SOI is extensively used in the microelectronics industry, it is also one of the most promising materials for optical/electrical integration.

The research in waveguiding in silicon started with planar waveguides [7], and continued with large rib waveguides [8]. Recently, there has been a trend to reduce waveguide dimensions and consequently both small rib [9] and strip waveguides [10] have been investigated.

In this chapter, more detailed analysis of planar waveguides is given than presented in [11]. Guidelines for single mode and birefringence free conditions for relatively small rib and strip waveguides are then discussed. As the current trend is to reduce the cross-sectional dimensions of the waveguides, coupling to small silicon waveguides is also discussed. Finally, novel waveguide designs, that have been recently proposed, are discussed.

## 2.2 Planar Waveguides

The simplest optical waveguide structure is the step-index planar waveguide. It has already been considered in [11] by the ray approach. In this text, we will implement a more rigorous analysis, based on Maxwell's equations. A planar waveguide is shown in Figure 2.1. It consists of a high-index dielectric layer with thickness  $h$  surrounded on either side by materials with



**Figure 2.1** Planar waveguide

lower refractive index. In other words, the refractive index of the film  $n_f$  is higher than the refractive index corresponding to the substrate  $n_s$  and the upper cladding  $n_c$ . In addition, we assume the that  $n_s > n_c$ , that is the *asymmetric* waveguide. The film is infinite in extent in the  $xz$ -plane, and finite in the  $y$ -direction. We choose the direction of propagation to be along the  $z$ -axis, and we place the  $y = 0$  coordinate at the top interface between the film and the upper cover (Figure 2.1).

We assume that the light is propagating through a dielectric (conductivity  $\sigma = 0$ ), nonmagnetic (magnetic permeability  $\mu = \mu_0$ ), isotropic and linear ( $\mathbf{D} = \epsilon \mathbf{E}$ ) medium. In that case we have the following Maxwell's equations:

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \quad (2.1)$$

$$\nabla \times \mathbf{H} = \epsilon \frac{\partial \mathbf{E}}{\partial t} = \epsilon_0 n^2 \frac{\partial \mathbf{E}}{\partial t} \quad (2.2)$$

where  $\mathbf{E}$ ,  $\mathbf{H}$ ,  $\mu_0$ ,  $\epsilon_0$ , and  $n$  are the electric field, magnetic field, free space permeability, free space permittivity and refractive index of the medium, respectively.

For an optically inhomogeneous medium refractive index is position dependent,  $n = n(\mathbf{r})$ . Following the similar procedure as in [11], from Equations (2.1) and (2.2) we can derive wave equations for  $\mathbf{E}$  and  $\mathbf{H}$ . Derivations of the wave equation for  $\mathbf{H}$  is given below, and the process is identical for the electrical field  $\mathbf{E}$ .

We start with taking the curl operator on Equation (2.2):

$$\nabla \times (\nabla \times \mathbf{H}) = \nabla \times \left( \epsilon_0 n^2 \frac{\partial \mathbf{E}}{\partial t} \right) \quad (2.3)$$

We apply the vector identity on the left-hand side of Equation (2.3)

$$\nabla \times (\nabla \times \mathbf{H}) = \nabla (\nabla \cdot \mathbf{H}) - \nabla^2 \mathbf{H} \quad (2.4)$$

From Maxwell's equation

$$\nabla \mathbf{B} = 0 \Rightarrow \frac{1}{\mu_0} \nabla \mathbf{H} = 0 \quad (2.5)$$

Therefore, Equation (2.4) can be reduced to

$$\nabla \times (\nabla \times \mathbf{H}) = -\nabla^2 \mathbf{H} \quad (2.6)$$

and consequently Equation (2.3) can be written in the following form

$$-\nabla^2 \mathbf{H} = \nabla (\varepsilon_0 n^2) \times \frac{\partial \mathbf{E}}{\partial t} + \varepsilon_0 n^2 \frac{\partial}{\partial t} (\nabla \times \mathbf{E}) \quad (2.7)$$

From Equations (2.2) and (2.1) we have

$$\frac{\partial \mathbf{E}}{\partial t} = \frac{1}{\varepsilon_0 n^2} \nabla \times \mathbf{H} \quad (2.8)$$

and

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \quad (2.9)$$

If we substitute Equations (2.8) and (2.9) into (2.7) we obtain the wave equation for  $\mathbf{H}$  for inhomogeneous media:

$$\nabla^2 \mathbf{H} + \frac{1}{n^2} \nabla n^2 \times (\nabla \times \mathbf{H}) = \mu_0 \varepsilon_0 n^2 \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (2.10)$$

Similarly, we can derive the following wave equation for  $\mathbf{E}$ :

$$\nabla^2 \mathbf{E} + \nabla \left( \frac{1}{n^2} \nabla n^2 \mathbf{E} \right) = \mu_0 \varepsilon_0 n^2 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (2.11)$$

To derive this equation, the last Maxwell's equation

$$\nabla \mathbf{D} = \rho \quad (2.12)$$

is used, where  $\rho$  is the charge density.

For homogeneous media, Equations (2.10) and (2.11) reduce to the well-known wave equations

$$\nabla^2 \mathbf{H} = \mu_0 \varepsilon_0 n^2 \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (2.13)$$

$$\nabla^2 \mathbf{E} = \mu_0 \varepsilon_0 n^2 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (2.14)$$

Equations (2.10) and (2.11) indicate that the components of the electric and magnetic field vectors ( $\mathbf{E}_x$ ,  $\mathbf{E}_y$ ,  $\mathbf{E}_z$ ,  $\mathbf{H}_x$ ,  $\mathbf{H}_y$ , and  $\mathbf{H}_z$ ) are coupled. Therefore, a scalar equation for each component cannot be established.

If the refractive index depends only on a single Cartesian coordinate  $n = n(y)$ , as in planar waveguides, the electric and magnetic fields take the following form:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(y) e^{j(\omega t - \beta z)} \quad (2.15)$$

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H}(y) e^{j(\omega t - \beta z)} \quad (2.16)$$

where  $\omega$  is the angular frequency and  $\beta$  propagation constant. We again assume the propagation along the  $z$ -axis.

For a particular propagation constant  $\beta$ , the field distributions are completely determined. Therefore, if the polarisation of light is known, a mode is defined by its propagation constant.

We will examine two different polarisations of light. The transverse electric (TE) polarisation is given by:  $\mathbf{E} = \mathbf{E}_x$ ,  $\mathbf{E}_y = \mathbf{E}_z = 0$ ,  $\mathbf{H}_x = 0$ , while the transverse magnetic (TM) is given by  $\mathbf{H} = \mathbf{H}_x$ ,  $\mathbf{H}_y = \mathbf{H}_z = 0$ ,  $\mathbf{E}_x = 0$ . We will now derive the eigenvalue equation for TM polarisation. The process is similar for TE polarisation.

From Equation (2.16) we have

$$\mathbf{H} = \mathbf{H}_x = H_x(y)e^{j(\omega t - \beta z)}\mathbf{u}_x \quad (2.17)$$

where  $\mathbf{u}_x$  is unity vector parallel with  $x$ -axis. From Equation (2.10) we can get:

$$\nabla^2 [H_x(y)e^{j(\omega t - \beta z)}\mathbf{u}_x] + \frac{1}{n^2}\nabla n^2 \times [\nabla \times H_x(y)e^{j(\omega t - \beta z)}\mathbf{u}_x] = \mu_0 \epsilon_0 n^2 \frac{\partial^2}{\partial t^2} [H_x(y)e^{j(\omega t - \beta z)}\mathbf{u}_x]$$

or

$$\begin{aligned} \frac{d^2 H_x(y)}{dy^2} e^{j(\omega t - \beta z)} \mathbf{u}_x + \frac{1}{n^2} \frac{dn^2}{dy} \mathbf{u}_y \times \left[ -\frac{dH_x(y)}{dy} e^{j(\omega t - \beta z)} \mathbf{u}_z - j\beta H_x(y)e^{j(\omega t - \beta z)} \mathbf{u}_y \right] \\ - H_x(y)\beta^2 e^{j(\omega t - \beta z)} \mathbf{u}_x = -\mu_0 \epsilon_0 n^2 \omega^2 \frac{\partial^2}{\partial t^2} H_x(y)e^{j(\omega t - \beta z)} \mathbf{u}_x \end{aligned}$$

Therefore:

$$\begin{aligned} \frac{d^2 H_x(y)}{dy^2} e^{j(\omega t - \beta z)} \mathbf{u}_x - \frac{1}{n^2} \frac{dn^2}{dy} \frac{dH_x(y)}{dy} e^{j(\omega t - \beta z)} \mathbf{u}_x - \beta^2 H_x(y)e^{j(\omega t - \beta z)} \mathbf{u}_x \\ + \mu_0 \epsilon_0 n^2 \omega^2 \frac{\partial^2}{\partial t^2} H_x(y) e^{j(\omega t - \beta z)} \mathbf{u}_x = 0 \end{aligned}$$

and finally

$$\frac{d^2 H_x(y)}{dy^2} - \frac{1}{n^2} \frac{dn^2}{dy} \frac{dH_x(y)}{dy} + [k_0^2 n^2(y) - \beta^2] H_x(y) = 0 \quad (2.18)$$

where  $k_0$  is given by

$$k_0^2 = \left( \frac{2\pi}{\lambda_0} \right)^2 = \mu_0 \epsilon_0 \omega^2 \quad (2.19)$$

In a region of constant refractive index, the TM wave equation (2.18) becomes

$$\frac{d^2 H_x(y)}{dy^2} + [k_0^2 n^2(y) - \beta^2] H_x(y) = 0 \quad (2.20)$$

Similarly, we can derive the TE wave equation

$$\frac{d^2 E_x(y)}{dy^2} + [k_0^2 n^2(y) - \beta^2] E_x(y) = 0 \quad (2.21)$$

Equation (2.20) is a second-order differential equation, and to solve it, additional conditions must be imposed. Therefore, we apply the boundary conditions at the interfaces, which leads to the continuity of  $H_x$  and  $(1/n^2)dH_x/dy$  at the boundary.

If we designate propagation constant  $\beta$  as

$$\beta = k_0 N \quad (2.22)$$

where  $N$  is the effective refractive index of the mode, and if we assume that

$$k_0 n_c < k_0 n_s < \beta < k_0 n_f$$

we have

$$n_c < n_s < N < n_f$$

Let us now solve Equation (2.20):

$$\frac{d^2 H_x(y)}{dy^2} - (\beta^2 - k_0^2 n_c^2) H_x(y) = \frac{d^2 H_x(y)}{dy^2} - \gamma_c^2 H_x(y) = 0, \quad y \geq 0 \quad (2.23)$$

$$\frac{d^2 H_x(y)}{dy^2} + (k_0^2 n_f^2 - \beta^2) H_x(y) = \frac{d^2 H_x(y)}{dy^2} + k_f^2 H_x(y) = 0, \quad 0 > y > -h \quad (2.24)$$

$$\frac{d^2 H_x(y)}{dy^2} - (\beta^2 - k_0^2 n_s^2) H_x(y) = \frac{d^2 H_x(y)}{dy^2} - \gamma_s^2 H_x(y) = 0, \quad y \leq -h \quad (2.25)$$

where  $\gamma_c$ ,  $\gamma_s$  and  $k_f$  are positive real numbers. The solution of the three different Equations (2.23–2.25) can be expressed as

$$H_x(y) = A e^{-\gamma_c y}, \quad y \geq 0 \quad (2.26)$$

$$H_x(y) = B \cos(k_f y) + C \sin(k_f y), \quad 0 > y > -h \quad (2.27)$$

$$H_x(y) = D e^{\gamma_s (y+h)}, \quad y \leq -h \quad (2.28)$$

where  $A$ ,  $B$ ,  $C$  and  $D$  are coefficients which can be determined from the boundary conditions. From Equations (2.26) and (2.27) and continuity of the magnetic field at the  $y = 0$  boundary, we can obtain the relation between coefficients  $A$  and  $B$ :

$$\begin{aligned} A e^{-\gamma_c \cdot 0} &= B \cos(k_f \cdot 0) + C \sin(k_f \cdot 0) \\ A &= B \end{aligned} \quad (2.29)$$

Similarly, from Equations (2.26, 2.27, 2.29) and continuity of  $1/n^2 \cdot [dH_x(y)/dy]$  at  $y = 0$  we can find the relation between coefficients  $A$  and  $C$ :

$$\begin{aligned} -\frac{1}{n_c^2} A \gamma_c e^{-\gamma_c \cdot 0} &= \frac{1}{n_f^2} [-B k_f \sin(k_f \cdot 0) + C k_f \cos(k_f \cdot 0)] \\ -A \frac{\gamma_c}{n_c^2} &= C \frac{k_f}{n_f^2} \Rightarrow C = -A \frac{\gamma_c}{n_c^2} \cdot \frac{n_f^2}{k_f} \end{aligned} \quad (2.30)$$

From Equations (2.27–2.30) and continuity of the magnetic field at  $y = -h$  boundary we can find coefficient  $D$ :

$$\begin{aligned} B \cos(k_f \cdot h) - C \sin(k_f \cdot h) &= D e^{\gamma_s \cdot 0} \\ D &= A \left[ \cos(k_f \cdot h) + \frac{\gamma_c n_f^2}{k_f n_c^2} \sin(k_f \cdot h) \right] \end{aligned} \quad (2.31)$$

The last boundary condition relation between the first derivatives of  $H_x(y)$  at  $y = -h$  gives the transcendental equation, from which we can find the propagation constant  $\beta$ .

$$\frac{1}{n_f^2} \frac{dH_x}{dy} \Big|_{(x=-h)} = \frac{1}{n_s^2} \frac{dH_x}{dy} \Big|_{(x=-h)} \quad (2.32)$$

$$\frac{1}{n_f^2} [B k_f \sin(k_f h) + C k_f \cos(k_f h)] = \frac{1}{n_s^2} \gamma_s D e^{\gamma_s \cdot 0} \quad (2.33)$$

$$\frac{1}{n_f^2} \left[ A k_f \sin(k_f h) - A \frac{\gamma_c n_f^2}{k_f n_c^2} k_f \cos(k_f h) \right] = \frac{\gamma_s A}{n_s^2} \left[ \cos(k_f h) + \frac{\gamma_c n_f^2}{k_f n_c^2} \sin(k_f h) \right] \quad (2.34)$$

If we divide this equation by  $\cos(k_f \cdot h)$  we obtain:

$$\begin{aligned} \tan(k_f h) - \frac{\gamma_c n_f^2}{k_f n_c^2} &= \frac{\gamma_s n_f^2}{k_f n_s^2} + \frac{\gamma_s n_f^2 \gamma_c n_f^2}{k_f n_s^2 k_f n_c^2} \cdot \tan(k_f h) \\ \tan(k_f h) \left[ 1 - \frac{\gamma_c \gamma_s n_f^4}{k_f^2 n_c^2 n_s^2} \right] &= \frac{\gamma_c n_f^2}{k_f n_c^2} + \frac{\gamma_s n_f^2}{k_f n_s^2} \\ \tan(k_f h) &= \left[ \frac{\gamma_c}{k_f} \left( \frac{n_f}{n_c} \right)^2 + \frac{\gamma_s}{k_f} \left( \frac{n_f}{n_s} \right)^2 \right] / \left[ 1 - \frac{\gamma_c \gamma_s}{k_f k_f} \left( \frac{n_f^2}{n_c n_s} \right)^2 \right] \end{aligned} \quad (2.35)$$

Similarly, we can derive a transcendental equation for TE polarisation:

$$\tan(k_f h) = \left( \frac{\gamma_c}{k_f} + \frac{\gamma_s}{k_f} \right) / \left( 1 - \frac{\gamma_c \gamma_s}{k_f k_f} \left( \frac{n_f^2}{n_c n_s} \right)^2 \right) \quad (2.36)$$

Once the propagation constant  $\beta$  is determined from Equation (2.35), the field amplitudes are known for all regions of the waveguide using Equations (2.26)–(2.28). We designate each unique field distribution as a mode of propagation. Coefficient  $A$  (the other coefficients are

expressed in terms of  $A$  via Equations (2.29–2.31)) is related to the power carried in the waveguide and can be determined from the Poynting vector:

$$P_z = \frac{1}{2} \int_{-\infty}^{\infty} E_x H_y dy = \left( \frac{\beta}{2\omega\mu_0} \right) \int_{-\infty}^{\infty} |E_x|^2 dy \quad (2.37)$$

Only a finite number of modes will be guided, because there are a finite number of solutions to the eigenvalue equation for  $\beta$ . Therefore, the spectrum of  $\beta$  for guided modes is discrete, unlike the spectrum for unguided (radiation) modes which is continuous.

If Equation (2.35) has several solutions the waveguide is multi-mode. If the waveguide supports only one mode we refer to it as single mode (or monomode) waveguide. It is also possible that there is no solution for the transcendental equation for a particular wavelength, and in this case the waveguide cannot support a guided mode.

Once the propagation constant  $\beta$  has been calculated, the coefficients  $\gamma_c$ ,  $k_f$  and  $\gamma_s$  (Equations 2.23–2.25) can be determined, and Equations (2.26–2.28) written in the following form, using Equations (2.29–2.31):

$$H_x(y) = Ae^{-\gamma_c y}, \quad y \geq 0 \quad (2.38)$$

$$H_x(y) = A \left[ \cos(k_f y) - \frac{\gamma_c}{k_f} \frac{n_f^2}{n_c^2} \sin(k_f y) \right], \quad 0 > y > -h \quad (2.39)$$

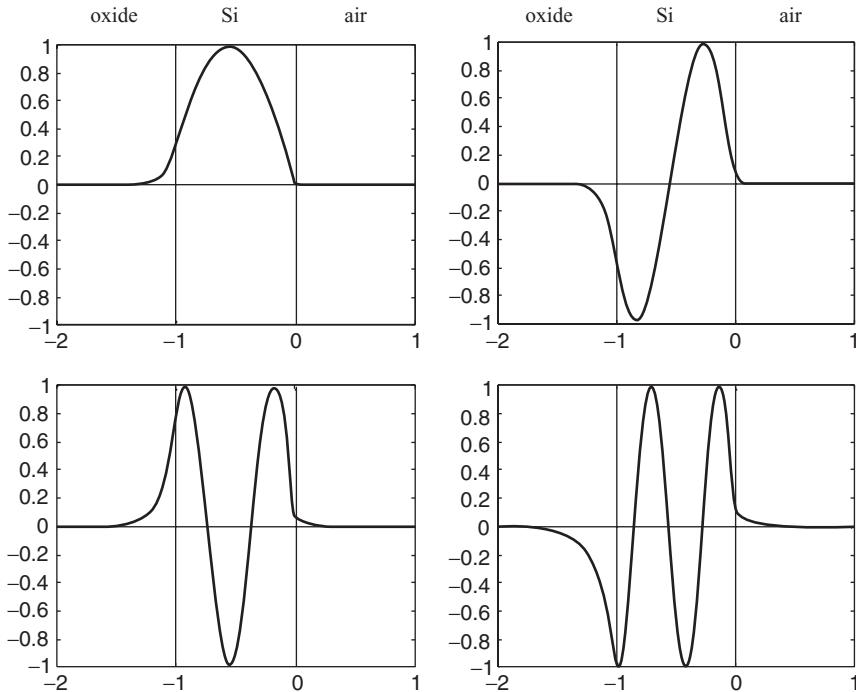
$$H_x(y) = A \left[ \cos(k_f \cdot h) + \frac{\gamma_c}{k_f} \frac{n_f^2}{n_c^2} \sin(k_f \cdot h) \right] e^{\gamma_s(y+h)}, \quad y \leq -h \quad (2.40)$$

As can be seen from these equations, the magnetic field decreases exponentially in the cover and substrate, whilst its dependence is sinusoidal in the film. The field penetration in the cover is lower than in the substrate, which is a consequence of  $n_c < n_s$  (Figure 2.2). It is also interesting to note that as the mode number  $m$  increases, the field penetration into the cover and substrate is deeper as can be seen in Figure 2.2. This is expected as the propagation constant  $\beta$  of the mode decreases with the increase of  $m$ , thus reducing the values of  $\gamma_c$  and  $\gamma_s$ . In other words, higher-order modes are less confined.

If the waveguide thickness decreases, there is a certain value of  $h$  for which the coefficient is  $\gamma_s = 0$ . We say that the wave reaches its cut-off point or simply that the waveguide is cut-off when this condition is reached. When  $\gamma_s$  becomes imaginary, the evanescent field in the substrate region becomes a radiation field, and the wave is no longer guided by the waveguide. In this case, we have radiation modes in the waveguide and the light is no longer confined in the film, but can leak to the substrate (Figure 2.3), and the cover. As the wave loses the power whilst propagating along the waveguide, these modes are called leaky modes.

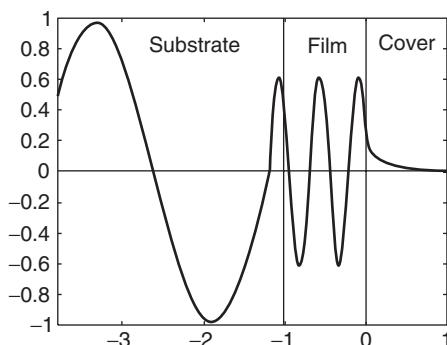
The solutions for leaky modes are not discrete. There exists a continuous and infinite number of values for the propagation constant, and therefore an infinite number of solutions for the electromagnetic field distribution in the structure.

For more complex waveguide structures, such as strip and rib waveguides, there is no analytical solution of the waveguide equation, and therefore numerical solutions must be sought. In order to increase accuracy of the modelling and to investigate different issues of

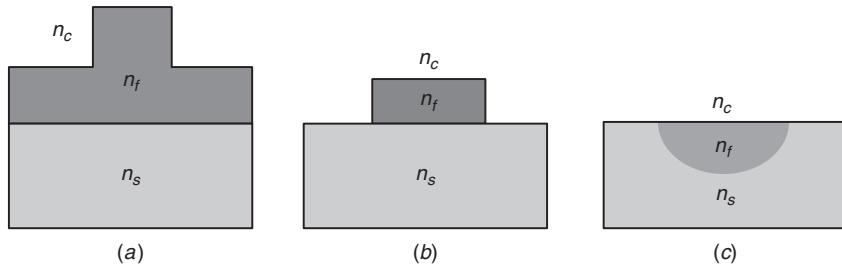


**Figure 2.2** Magnetic field modal patterns for the first four TM modes in a planar SOI waveguide. Reproduced from [13]

the light propagation in the waveguides, several numerical methods have been developed: the finite element method (FEM), beam propagation method (BPM), finite difference time domain (FDTD), just to name a few. It is beyond the scope of this chapter to go into details of these methods, but it is worth noting that they have been used in the analysis presented below.



**Figure 2.3** Substrate radiation mode in an asymmetric planar waveguide



**Figure 2.4** Three main types of channel waveguides: (a) rib; (b) strip; (c) buried waveguide

## 2.3 Rib Waveguides

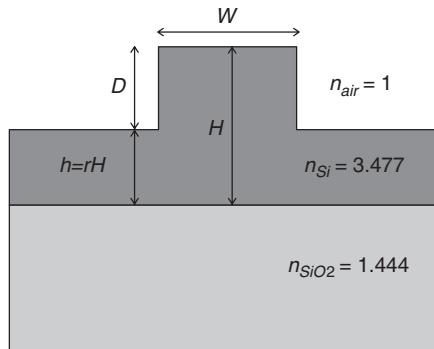
The three most common basic channel waveguide structures are: rib (Figure 2.4a), strip (Figure 2.4b) and buried (Figure 2.4c) waveguides. In this section, design guidelines for single mode and birefringence-free small rib waveguides are given, whilst in the next section, such rules are given for strip waveguides.

Unlike silica waveguides which can be relatively easily designed to be single mode, SOI waveguides with dimensions larger than a few hundred nanometres in cross-section will support multiple modes. Such multimode waveguides are usually undesirable in photonic circuits as their operation can be seriously compromised by the presence of multiple modes.

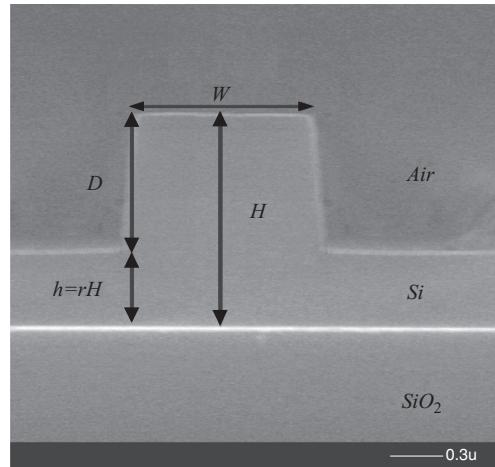
It has been shown that large rib waveguides in SOI (Figure 2.5) could be designed so as to be monomodal [13]. These waveguides have been studied extensively by a number of researchers e.g. [14–17] to find single mode behaviour and low loss propagation. Large rib waveguides are interesting because they are multi-micrometre in cross-sectional dimensions (of the order of 5  $\mu\text{m}$ ) facilitating low-loss coupling to and from optical fibres. Soref *et al.* [13] first proposed a simple expression for the single-mode condition (SMC) of such waveguides:

$$\frac{W}{H} \leq 0.3 + \frac{r}{\sqrt{1 - r^2}} \quad (\text{for } 0.5 \leq r < 1) \quad (2.41)$$

where  $r$  is the ratio of slab height to overall rib height, and  $W/H$  is the ratio of waveguide width to overall rib height (Figure 2.5). Their analysis of the waveguides was limited to shallow etched



**Figure 2.5** Rib waveguide in SOI

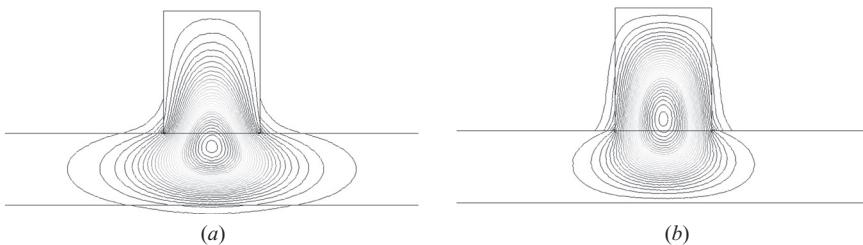


**Figure 2.6** SEM of a relatively small rib waveguide with  $r < 0.5$  ( $H = 1.20 \mu\text{m}$ ,  $W = 0.98 \mu\text{m}$ ,  $D = 0.76 \mu\text{m}$ ) [22]

ribs ( $r > 0.5$ ) and the waveguide dimensions were assumed to be larger than the operating wavelength. The analysis was based on the assumption that high-order vertical modes (i.e. modes other than the fundamental mode) confined under the rib, were coupled to the outer slab region during propagation, therefore yielding high propagation losses for the higher order modes. Thus the waveguides behave as single mode waveguides, as all other modes are lost.

Other authors have also considered the single mode condition for large waveguides, and produced similar expressions, e.g. [14–15]. However, as the current trend in silicon photonic circuits is to move to smaller device dimensions for improved cost efficiency and device performance, design rules for relatively small single mode and polarisation independent waveguides need to be established. One such small rib waveguide is shown in Figure 2.6.

The increased polarisation dependence in small waveguides is derived from the increasingly differing mode shapes of the TE and TM modes (Figure 2.7). The question arises as to whether the prime concern is to maintain similar losses for the TE and TM modes or to provide similar propagation constants in order to maintain similar phase performance for interferometric based devices, because it is not generally possible to maintain both.

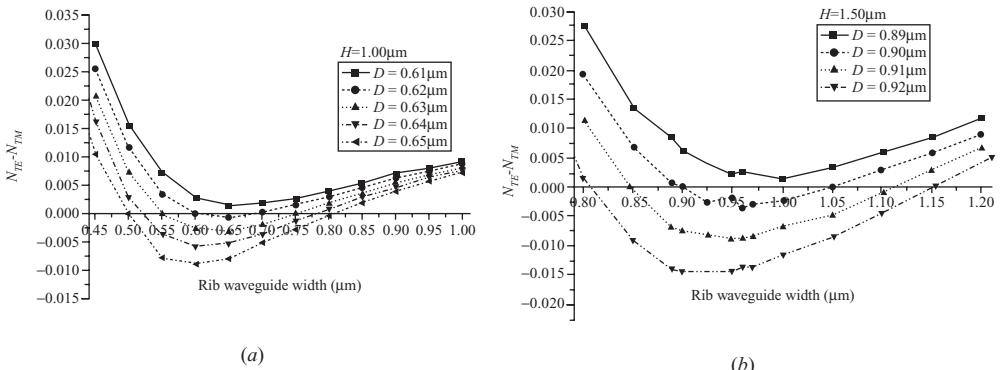


**Figure 2.7** (a) TE and (b) TM mode shapes for a rib waveguide in SOI with  $H = 1.35 \mu\text{m}$ ,  $D = 0.85 \mu\text{m}$ , and  $W = 0.70 \mu\text{m}$  [21]

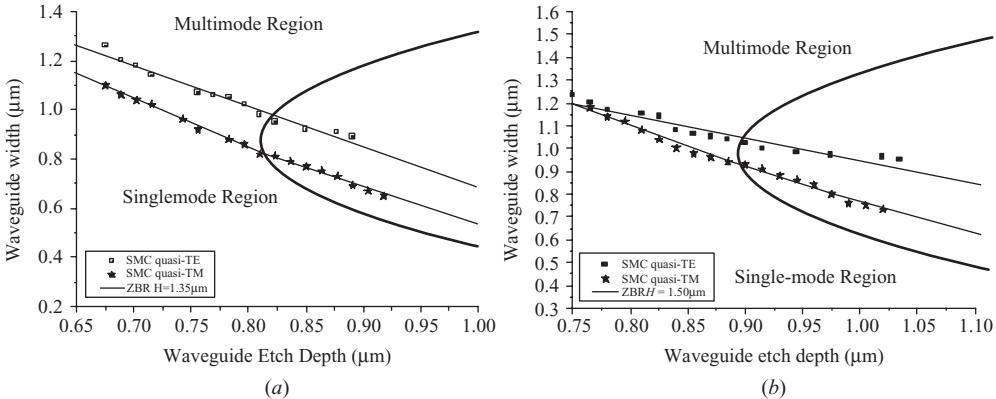
In order to maintain the consistency with other work [13, 15–17], the full-vectorial beam propagation method (BPM) [18] has been used to analyse deeply etched rib waveguide structures, mode propagation, and polarisation independence. Some results of the single mode condition (SMC) have also been verified by the finite element method (FEM) [19]. The simulations were carried out for rib waveguides of silicon ( $n_{\text{Si}} = 3.477$ ) on silica ( $n_{\text{SiO}_2} = 1.444$ ) and an upper cladding that is air ( $n_{\text{air}} = 1$ ) (Figure 2.5). SOI rib waveguides with an overall height in the range  $H = 1.00 \mu\text{m}$ , to  $1.50 \mu\text{m}$ , were analysed at a wavelength of  $1.55 \mu\text{m}$ .

The single mode cut-off condition has been evaluated by determining when the first mode of higher order than the fundamental mode begins to propagate. As the waveguides are relatively small, for high etch depths ( $r < 0.5$ ), the single mode condition becomes dominated by boundary conditions, and hence the conditions for quasi-TE and quasi-TM modes begin to diverge. We use ‘quasi-TE’ and ‘quasi-TM’ terms because there are no pure TE and TM modes in strip and rib waveguides. It has been shown that the single mode cut-off condition for deeply etched rib waveguides is different from design rule (2.41), but also that the work of Soref *et al.* [13] can be used as a basis for design of single mode waveguides with shallow-etched rib and large cross-sections [9].

The effective indices of the fundamental quasi-TE and quasi-TM modes have also been evaluated. Assuming an operating wavelength of  $1550 \text{ nm}$ , a graph of the variation of the TE/TM fundamental mode effective-index difference for various etch depths and waveguide widths can be produced. The condition when the effective indices are equal for both polarisations is defined as the zero birefringence condition, and there are up to two such events for each waveguide etch depth. Two examples are shown in Figure 2.8, for the rib height of  $1$  and  $1.5 \mu\text{m}$ . This indicates it is possible to produce birefringence-free waveguides of two different waveguide widths when a deep etch depth is employed. The wider waveguide width is highly preferable compared with its narrower counterpart when we consider that the parameters of wider devices will result in better accuracy during fabrication. If we consider  $H = 1.50 \mu\text{m}$ , and  $D = 0.90 \mu\text{m}$ , the birefringence-free waveguide widths can be determined from Figure 2.8(b) as  $0.90$  and  $1.05 \mu\text{m}$ . If we now introduce etch-depth uncertainties of (say)  $10 \text{ nm}$ , we see that the latter width is more desirable as the impact of the uncertainty on change in birefringence is reduced from  $6.03$  to  $3.28 \times 10^{-4}$ . The same trend can also be observed in Figure 2.8(a).



**Figure 2.8** Effective-index difference calculation between quasi-TE and quasi-TM polarized modes using the FEM, for waveguide heights of (a)  $1.00 \mu\text{m}$ ; (b)  $1.50 \mu\text{m}$  [9]



**Figure 2.9** The single mode condition and the polarisation independence locus plotted on the same graph, for rib waveguides with height of: (a) 1.35  $\mu\text{m}$ ; (b) 1.5  $\mu\text{m}$  [9]

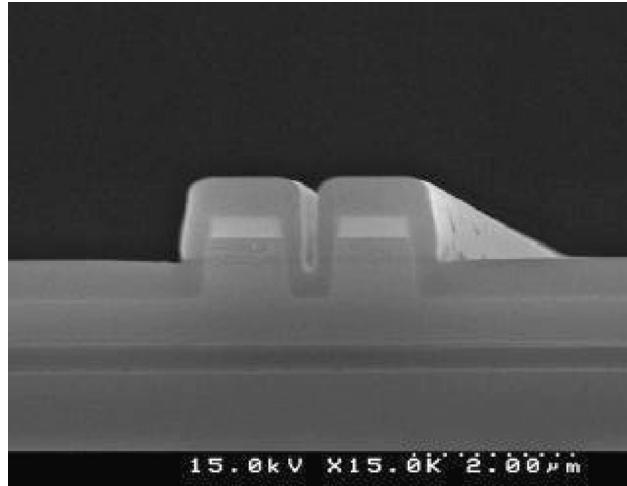
Both the single mode condition and the zero birefringence conditions can conveniently be plotted on the same curve to determine the waveguide parameters that allow both conditions to be satisfied simultaneously. For example, Figure 2.9 shows such plots for waveguide heights of  $1.35\text{ }\mu\text{m}$  (Figure 2.9a) and  $1.5\text{ }\mu\text{m}$  (Figure 2.9b). It can be seen that for truly single mode behaviour, it is the quasi-TM condition that is the limiting condition, because if this is satisfied, then the quasi-TE condition is automatically satisfied. Consequently, for both single mode behaviour and polarisation independence, the waveguide design should lie on the zero birefringence locus, below the quasi-TM single mode boundary (bottom right hand corner of Figures 2.9a and 2.9b). It is interesting to note that the crossing points for the two polarisations between the birefringence-free condition and both single-mode boundary lines are closer to each other as the waveguide height is increased (Figure 2.9). From the simulation data obtained, we can extract design guidelines for rib waveguide heights in the range  $H = 1.00\text{--}1.50\text{ }\mu\text{m}$  [9]:

$$\frac{W}{H} \leq 0.05 + \frac{(0.94 + 0.25H)r}{\sqrt{1-r^2}} \quad \text{for } 0 \leq r \leq 0.5 \quad \text{and} \quad 1.0 \leq H \leq 1.5 \quad (2.42)$$

$$D_{\min} = 0.06 \times 10^{-6} + 0.556H \quad (2.43)$$

Equation (2.42) defines the quasi-TM single mode boundary, and hence provides guidance on the geometric limitations to retain single mode behaviour, whilst Equation (2.43) defines the minimum etch depth required to obtain polarisation independence. If waveguide width or etch depth is increased, that can result in Equation (2.42) not being satisfied, and consequently in the introduction of higher-order modes.

The guidelines of Equations (2.42) and (2.43) are for waveguides with a surface cladding of air. However, an upper cladding layer is often deposited to reduce the influence of surface contamination, to passivate the surface, or to provide electrical isolation. For SOI, this layer is usually  $\text{SiO}_2$  (Figure 2.10), although nitride and polymer layers are also used in some cases. Therefore, the single mode and birefringence-free conditions also need to be determined for such surface clad waveguides, and therefore examples are analysed here. The



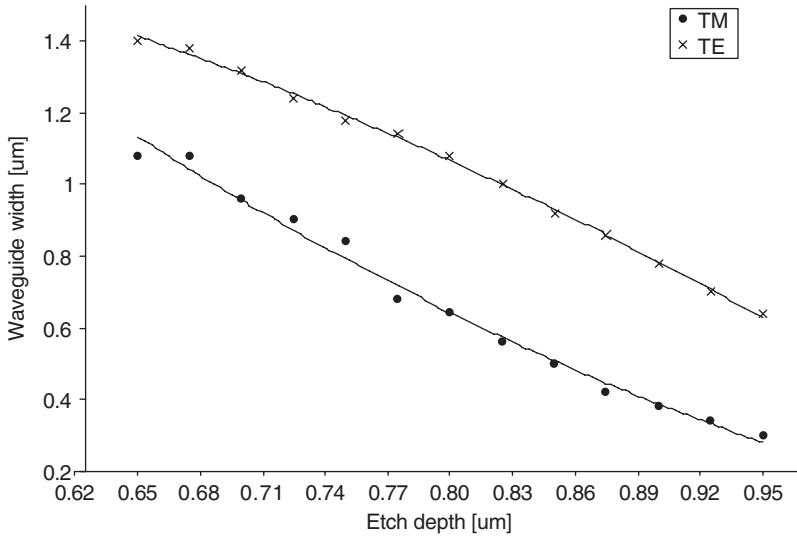
**Figure 2.10** SEM of a directional coupler made of two rib waveguides with oxide cover

structure we analyse is a rib waveguide with the height of 1.35  $\mu\text{m}$ , a variable waveguide width ( $W = 0.2\text{--}1.6 \mu\text{m}$ ) and etch depth ( $D = 0.65\text{--}0.95 \mu\text{m}$ ), which has the following refractive indices:  $n_{\text{core}} = 3.4764$ ,  $n_{\text{top cladding}} = n_{\text{buried oxide}} = 1.444$ . The simulations were performed by using the 3D semi-vectorial BPM method and verified by 2D FEM modelling. The method used was to calculate the effective indices for the fundamental and the first two higher-order modes, for both polarisations, over a range of waveguide dimensions.

Simulations confirm again that the TM mode is the most critical for multimode behaviour since the single mode (SM) curve for the TM polarisation is below the corresponding curve for the TE polarisation (Figure 2.11). Fitting the data points to a second order polynomial, the SM condition determined for the TM mode can be expressed as:

$$W \leq 2.86 \times D^2 - 7.42 \times D + 4.75 \quad (2.44)$$

It can be seen from Figure 2.11 that the SM conditions for TE and TM polarisations for the rib with oxide cladding are further apart than for the rib with air cladding (Figure 2.9a). It may be possible, therefore, that a single mode device that is polarisation independent (PI) could not be possible for certain dimensions/geometries of the rib. However, additional stress related refractive index changes derived from oxide claddings [21], and the influence of the side wall angle (Figure 2.10) [22] can result in a shift of the SM and ZBR curves and consequently in matching the SM and PI conditions simultaneously. Furthermore, use of the FEM simulation method results in less strict conditions suggesting, for example,  $W = 0.50 \mu\text{m}$  rather than  $W = 0.40 \mu\text{m}$  as the SM-boundary when  $D = 0.90 \mu\text{m}$ . This is not surprising considering that there is no sharp transition between the single-mode and multi-mode behaviour. Therefore, by allowing some flexibility for the SM characteristics, it should be possible to match PI and ZBR simultaneously, as has been demonstrated experimentally by Headley *et al.*, [23] for rib waveguides of 1.35  $\mu\text{m}$  in height.



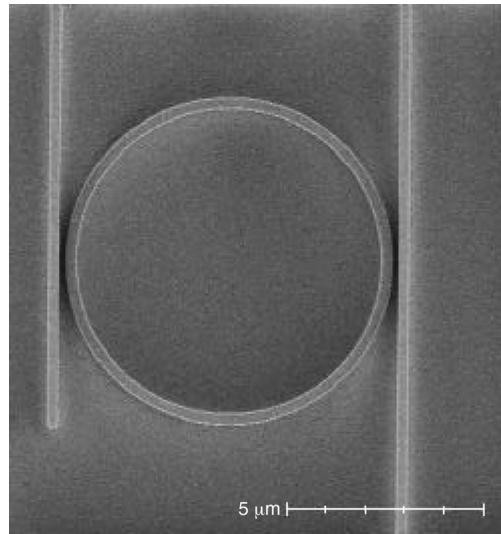
**Figure 2.11** Single mode condition for a rib waveguide with upper oxide cladding ( $H = 1.35 \mu\text{m}$ ,  $n_{\text{core}} = 3.4764$  and  $n_{\text{oxide}} = 1.444$ )

In addition, together with waveguide bends, side wall angle and oxide cover could be responsible for polarisation conversion [24], which has been observed experimentally [25], and therefore their influence needs to be investigated and incorporated into the design rules for small rib waveguides. It can be seen that meeting both single mode and polarisation independence can be a nontrivial task as it involves tailoring of several design parameters.

## 2.4 Strip Waveguides

SOI is a platform that offers very high refractive index contrast and consequently strong light confinement. This allows shrinking of the waveguide core to submicrometre dimensions. The core size for single mode propagation at the 1.3–1.5  $\mu\text{m}$  telecommunications wavelengths is a few hundred nanometres. Such extreme light confinement also allows the minimum bending radius to be reduced to the micrometre range (Figure 2.12), and therefore offers the possibility of realisation of ultra-dense photonic circuits, which can further decrease the cost of silicon photonics. It has been shown that small waveguides with small bending radii can improve the characteristics of photonic devices, for example, optical modulators and filters [26, 27]. Furthermore, such small waveguides or silicon wires as they are sometimes called, can realise an ultra-high optical power density, which can be as much as 1000 times that in a conventional single mode fibre, enhancing nonlinear optical effects [28].

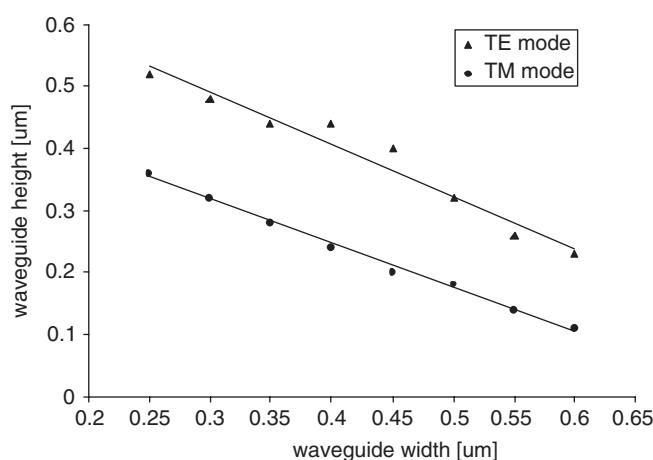
Unlike rib waveguides which can be monomodal, even for cross-sectional dimensions of several micrometres, strip waveguide dimensions must be significantly smaller than 1  $\mu\text{m}$  to suppress propagation of higher-order modes. Results obtained by 2D simulations for strip waveguides with refractive indices of 3.476 and 1.444 for the guiding structure and surrounding oxide respectively (Figure 2.13), suggest that strip waveguides can be multi-mode even for



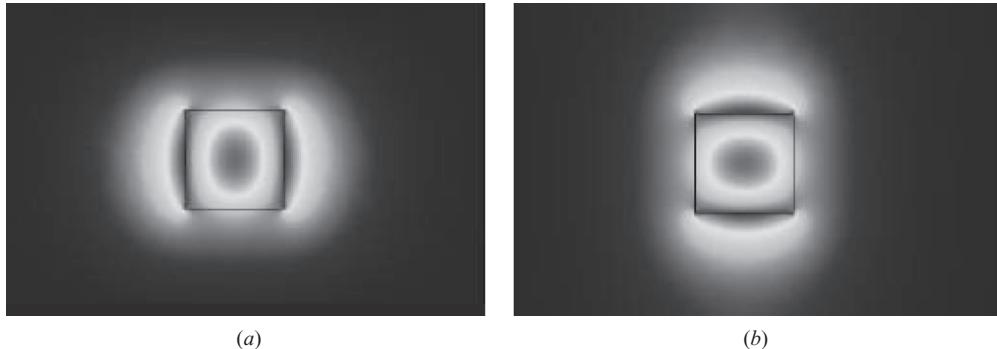
**Figure 2.12** Ring resonator structure based on strip waveguides with cross-section  $\approx 300 \times 300$  nm (SEM taken by Intel Corp.)

relatively small dimensions. This is particularly the case for the TM polarisation, which is the most critical condition for waveguide operation in the SM regime, as it is for rib waveguides (Figure 2.9).

Alto gave an approximate expression for the single mode condition for silicon strip waveguides:  $W \times H < 0.13 \mu\text{m}^2$  [29], where  $W$  and  $H$  are the width and height of the strip waveguide. This condition agrees reasonably well with our simulations for the TE polarisation, but is not sufficiently restrictive for TM polarisation (Figure 2.13). The line representing the boundary



**Figure 2.13** Single mode condition for strip waveguides



**Figure 2.14** Field profiles for: (a) TE; (b) TM polarisation in a small strip silicon waveguide

between the single-mode and the multi-mode regime can be approximated by the analytical expression

$$W \leq -1.405H + 0.746 \quad (2.45)$$

As an example, in order to prevent propagation of higher-order modes, a  $0.3\text{ }\mu\text{m}$  high waveguide should not be wider than  $\sim 0.35\text{ }\mu\text{m}$ . Clearly such a restriction has implications for fabrication resolution, which is why some authors suggest that aiming for polarisation independence in strip waveguides is impractical.

We can expect that the propagation constants for TE and TM polarisation modes will be identical for a square cross-section of the strip waveguide, and that polarisation independence can be achieved. Figure 2.14 shows field profiles for TE and TM modes in a  $300 \times 300\text{ nm}$  strip waveguide which confirm that the same propagation constants are possible. The TE field profile is characterised by much higher field intensity at the side walls, whilst the TM mode has a relatively small amplitude at the side walls, but much higher amplitude at the top and bottom interfaces (Figure 2.14).

This is the reason why propagation loss is typically higher for TE modes than for TM modes [30]. In particular, the roughness of the side walls is higher than the roughness of the top and bottom interfaces and that causes additional propagation loss for TE mode as its field is much higher at the side walls. The roughness of the side walls of the waveguide is a result of the imperfection of the resist pattern that can also be increased by the plasma etching. Therefore, lithography and etching are the key processes in achieving smoothly etched Si designs [31].

Propagation loss of TE modes is around  $3\text{ dB/cm}$  at the wavelength of  $1550\text{ nm}$  (e.g.  $3.6\text{ dB/cm}$  for  $220 \times 445\text{ nm}$  waveguides in [30], and  $2.8\text{ dB/cm}$  for  $200 \times 400\text{ nm}$  waveguides in [31]). These figures could be decreased further by sacrificial oxidation of the side walls, which smooths the sidewall further [32]. Vlasov and McNab also measured losses per  $90^\circ$  bend for  $220 \times 445\text{ nm}$  strip waveguides and obtained  $0.005\text{ dB/turn}$ ,  $0.013\text{ dB/turn}$ , and  $0.086\text{ dB/turn}$  for bending radius of  $5$ ,  $2$ , and  $1\text{ }\mu\text{m}$ , respectively [30].

In passing, we note that it can be seen in Figure 2.14(b), there is a large highly localised electric field distribution near the waveguide surface of small strip SOI waveguides. Therefore, shrinking silicon waveguides to these small dimensions, opens another application area,

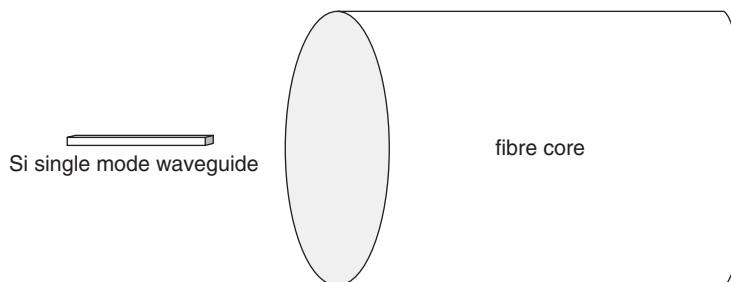
as these waveguides could be used as evanescent field sensors for biological or chemical applications [33].

## 2.5 Coupling to Small Silicon Waveguides

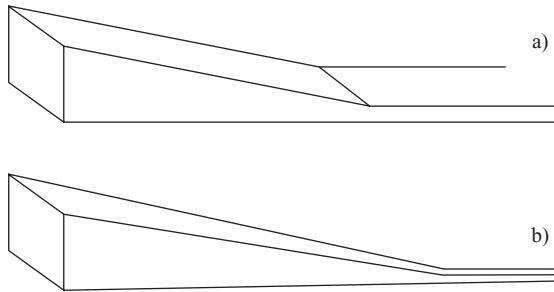
Whilst the reduction in waveguide dimensions enhances the performance of the photonic circuit and the packing density, it makes coupling of light to/from the circuit very difficult, particularly to/from standard optical fibres that typically have a core dimension of  $9\text{ }\mu\text{m}$ . As we could see from the previous section, to satisfy the single mode condition, the cross-section of a square Si waveguide core must be  $\sim 300 \times 300\text{ nm}$ . Direct coupling of the fibre and such a small silicon waveguide results in high coupling losses ( $\approx 20\text{ dB}$ ) due to very different thicknesses (Figure 2.15) and refractive indices.

Consequently, several approaches have been proposed for efficient coupling to/from the silicon photonic circuit, the most popular being tapers and grating-based couplers. Vertically (Figure 2.16a) (and laterally (Figure 2.16b)) tapered structures are intuitively the most obvious approach. These structures are integral waveguide extensions and act as classic adiabatic tapers that transform the input fundamental mode shape, which is matched to the mode shape of the fibre, to the waveguide mode shape.

The geometry of a vertically tapered mode converter is illustrated in Figure 2.17. The silicon dioxide layer should be at least  $1\text{ }\mu\text{m}$  thick to prevent power leakage to the substrate. An anti-reflection coating is implemented to reduce Fresnel reflection. Selective epitaxial growth of silicon is also involved to achieve the height of the converter to match the fibre mode diameter. The lateral taper can be obtained by DRIE (deep reactive ion etching), whilst grey-scale lithography is needed for the vertical taper [34]. Grey-scale techniques utilise a mask that is designed to project onto the photoresist, a photolithography light beam of variable intensity. The resist is designed so that its depth of removal during the developing step is dependent upon the exposure it receives. As a result, when the photoresist irradiated through the grey-scale mask is developed, it will have varying thickness. When the photoresist layer is subjected to the subsequent silicon etch step, the thinner regions of photoresist are fully removed, and thereby expose underlying silicon earlier than thicker regions of photoresist. The depth to which the underlying silicon is etched is therefore determined by the thickness of the developed photoresist, the selectivity, and the etch time. The result is that the depth of the silicon etch is made to vary across the silicon surface in a predetermined fashion.



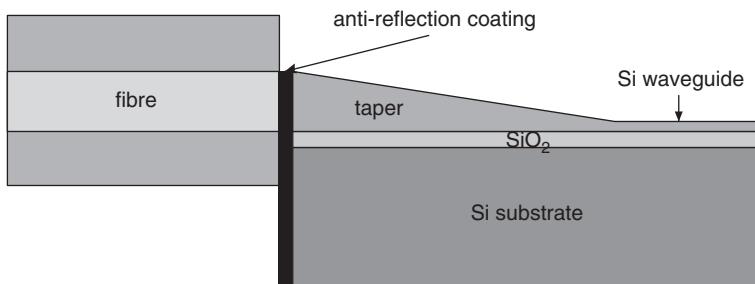
**Figure 2.15** Butt coupling of optical fibres with small silicon waveguides is inefficient



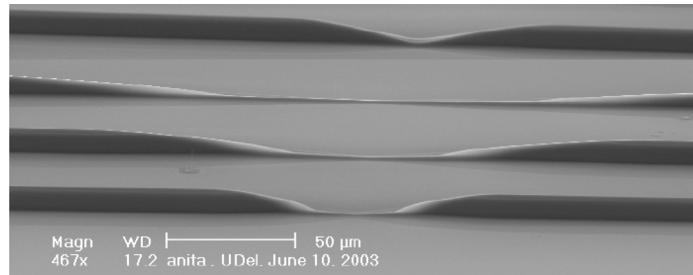
**Figure 2.16** Shapes of waveguide mode converters in silicon

Sure *et al.*, have improved grey-scale lithography and reported experimental results for a vertical taper in silicon (Figure 2.18) [35]. They employed a high-energy beam sensitive (HEBS) glass mask with UV grey-scale lithography. The desired device profile was encoded as an optical density (OD) profile within the mask, having determined the resist response to varying UV exposure. To pattern the mask blank with a continuous OD profile, they developed a technique that involved deconvolution of the desired OD profile with an experimentally determined electron beam (e-beam) point spread function. After the HEBS greyscale mask was patterned, a single lithographic step was used to transfer the mask pattern into the underlying substrate. The cross-sectional variation of the tapers was limited to vertical tapering because of the resolution constraints of the thick resist. Such couplers did not achieve horizontal mode conversion. The couplers had input and output facets, each 10  $\mu\text{m}$  high and a central waveguide 0.25–2  $\mu\text{m}$  high, depending on the mask design. With the careful choice of optical density in the mask, control of the lithography parameters and inspection during the development process, the repeatability of the central waveguide heights were within  $\pm 0.2 \mu\text{m}$ . For the coupler with the central waveguide height of 0.25  $\mu\text{m}$  the theoretical prediction for coupling efficiency was 82% while the measured value was 45%.

Recently, several inverted taper structures in SOI have been proposed and demonstrated, in which the waveguide is reduced in size to expand the mode size and reduce the effective refractive index. An inverse taper is shown in Figure 2.19. As the coupling is maximised when the overlap between the two waveguide modes is maximised, the mode in one waveguide



**Figure 2.17** Cross-section of mode converter and waveguide structure in silicon

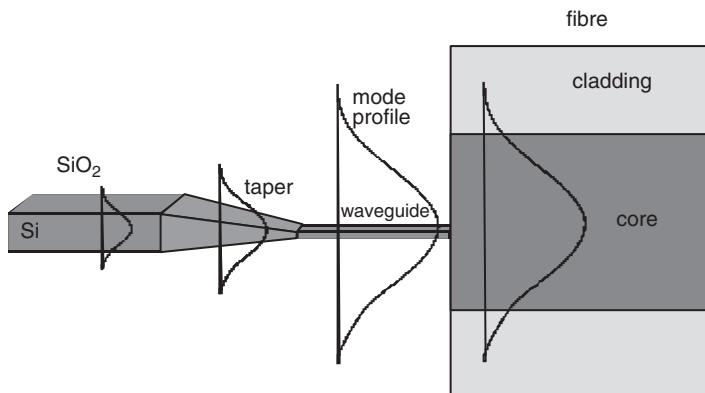


**Figure 2.18** SEM micrograph of vertical taper couplers in SOI [37]

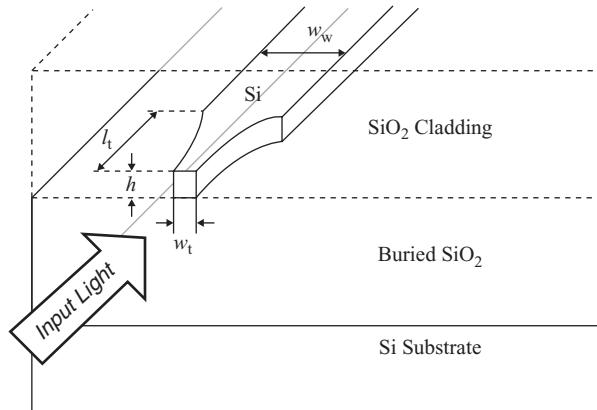
(Si waveguide) is transformed to be similar to the mode in the other waveguide (fibre). The inverted taper provides this mode transformation. If a mode in the Si waveguide propagates towards the taper, as the Si waveguide becomes smaller, the mode will become less well confined within the Si, and will start to expand in size. If the Si waveguide cladding is a low-index material such as  $\text{SiO}_2$  or polymer, at some point the mode will match the fibre mode. If the fibre is placed at that point in space, the overlap will be maximised ensuring strong coupling.

The group from Cornell University proposed and demonstrated a short inverted taper coupler shown in Figure 2.20 [36]. The nanotaper consists of a waveguide laterally tapered to a nanometre-sized tip at the facets in contact with the fibre. A short taper transition was employed by gradually varying both side walls in a symmetric parabolic transition toward the final waveguide width, where the parabola vertex was located at the nanotaper tip. At the tip, the mode field profile becomes delocalised from the waveguide core. This delocalisation of the mode field profile increases the mode overlap with the optical fibre mode. In addition, most of the mode field resides in the  $\text{SiO}_2$  cladding region at the tip, causing the effective index to be close to that of the fibre, which results in negligible back-reflections.

The coupler was fabricated on an SOI wafer with a 3  $\mu\text{m}$  buried oxide layer, using e-beam lithography, followed by inductively coupled plasma etching and the deposition of a 3  $\mu\text{m}$



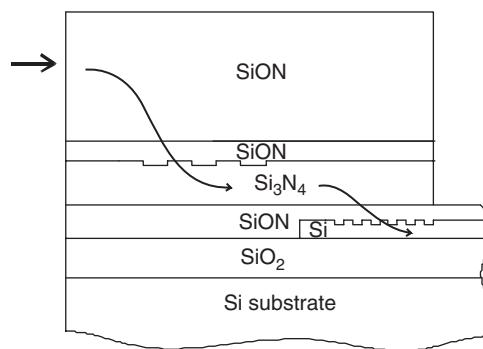
**Figure 2.19** Inverted taper in silicon technology



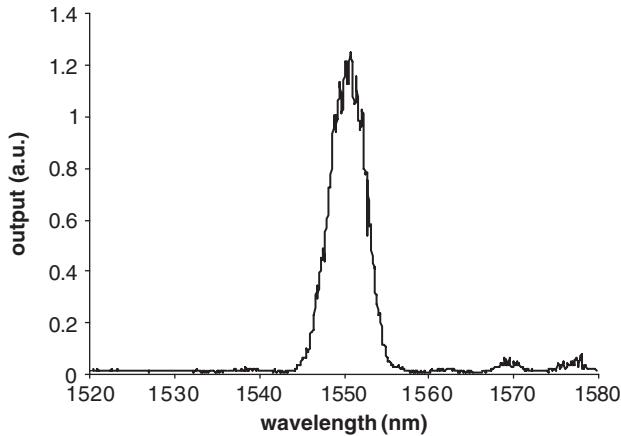
**Figure 2.20** Schematic of a waveguide with a short nanotaper proposed by Almeida *et al.* [36]

thick SiO<sub>2</sub> cladding by plasma enhanced chemical vapour deposition (PECVD). The average nanotaper and waveguide dimensions were  $l_t = 40 \mu\text{m}$ ,  $w_w = 470 \pm 20 \text{ nm}$ ,  $h = 270 \pm 10 \text{ nm}$ ,  $w_t = 100 \pm 10 \text{ nm}$ . The measured nanotaper insertion loss, which was  $3.3 \pm 0.3 \text{ dB}$  for the TM-like mode and  $6.0 \pm 0.4 \text{ dB}$  for the TE-like mode at  $\lambda_0 = 1550 \text{ nm}$ , originated from the mode mismatch loss between the optical fibre and tip facet modes and from mode conversion of the low-confined mode at the tip facet into the high-confined mode in the waveguide. Reference [37] from IBM also reports an inverted taper with a very impressive coupling loss of only 0.5 dB per interface. The waveguide loss was measured by a relatively inaccurate technique, with consequentially a very large uncertainty in the measured value, which means that the quoted loss could actually be significantly higher. However, similar results were reported by NTT [38] showing the viability of the technique.

The Surrey group have developed the dual grating-assisted directional coupler (DGADC) [39]. A schematic of an example design is shown in Figure 2.21. The top waveguide layer is 5  $\mu\text{m}$  thick with refractive index close to the refractive index of optical fibre, resulting in low



**Figure 2.21** DGADC in SOI [41]

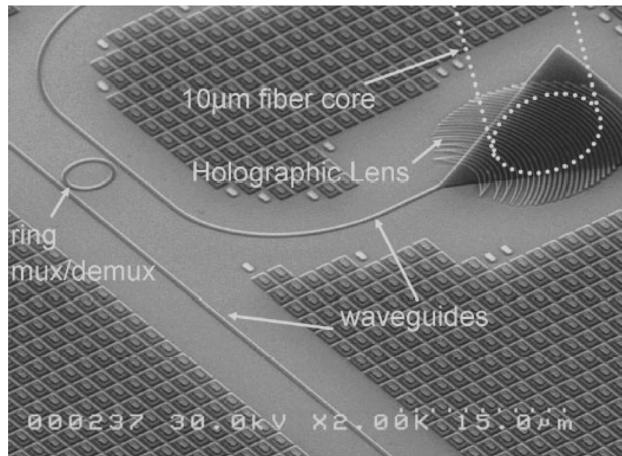


**Figure 2.22** Typical output from DGADC

insertion loss from the fibre to this waveguide. Therefore a fibre could be butt coupled to the thick SiON waveguide and subsequently the light coupled to the  $\text{Si}_3\text{N}_4$  waveguide using the first grating, and subsequently to the thin SOI waveguide via the second grating. The silicon nitride waveguide is crucial for the operation of the device, because it enables highly efficient coupling at both gratings, consequently forming an efficient DGADC. This waveguide bridges the gap between SiON and Si layers in both refractive index and thickness. The buried oxide layer serves as the lower cladding layer, for isolation from the substrate, hence removing any leakage loss.

Theoretical coupling efficiency is in excess of 90%, while the measured value was 55% [40]. As the thickness of the top SiON layer was not optimised this experimental value was within 5% of the theoretical value for the specific coupler design that was experimentally investigated. It is expected that the experimental value for optimised devices can be significantly higher ( $\sim 80\%$ ). Typical output of a double DGADC configuration [40] is shown in Figure 2.22. Light from an optical fibre was coupled to a 230 nm thick SOI waveguide via the first DGADC and then coupled back to another fibre via the second DGADC. For different grating periods, the resonant peak can be shifted towards longer or shorter wavelengths, still achieving coupling efficiency  $> 42\%$  for both C and L wavelength bands.

Surface grating couplers [41] provide a means of coupling to individual modes, and they are useful for coupling to waveguide layers of a wide range of thicknesses. The input beam must be introduced at a specific angle. These couplers can be used for testing of photonic integrated circuits. Previously, efficiencies in excess of 80% were achieved for coupling from silicon waveguides of thicknesses of 1  $\mu\text{m}$  [42]. Recently, Luxtera Inc. announced insertion loss of 1–1.5 dB over a wavelength range of 1537–1557 nm for coupling to single mode silicon waveguides [43, 44]. They used a two-dimensional grating design that they refer to as a holographic lens, shown in Figure 2.23. This design has the advantage of utilising a vertical height similar to that of the remainder of the photonic circuit, and hence no additional deposition or growth is required.



**Figure 2.23** Holographic lens coupler [46]

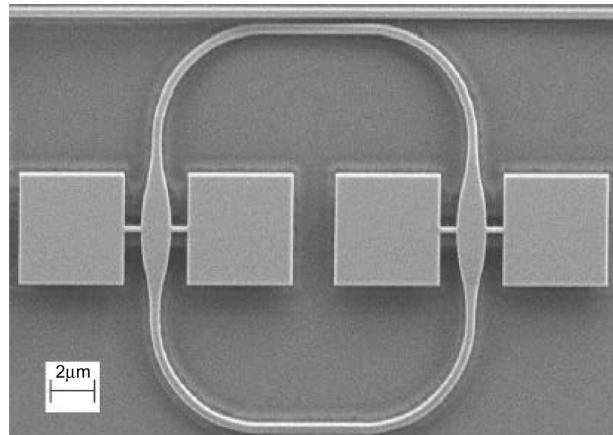
## 2.6 Novel Silicon Waveguide Structures

From the preceding sections of this chapter, we see that the most popular waveguide platform in silicon photonics is the silicon-on-insulator (SOI) structure, in the form of either a strip or a rib waveguide. This material structure, however, is not suitable for longer wavelengths (except in the 2.9–3.6  $\mu\text{m}$  range) due to the absorption spectra of silicon dioxide [45]. Silicon-based long-wave infrared photonics could find applications in several areas, including sensing, communications, signal processing, missile detection and imaging. To address these wavelengths, several different waveguide structures have been investigated such as: photonic crystals, plasmon optics, hollow waveguides, and nano-slotted rib or strip waveguides. In the next section, recent reports of the fabrication and propagation loss measurements of another different structure, the free-standing waveguide, are discussed, whilst in Section 2.6.2, another candidate for sensing and long wavelength applications, a hollow waveguide, is reviewed.

### 2.6.1 Free-standing Waveguides

Suspended structures, such as for example disk and ring resonators, have been recently demonstrated, e.g. [46, 47]. These structures were fabricated using relatively common patterning and undercutting methods. In [47], a micrometre-scale suspended ring resonator with a quality factor in excess of 15 000 has been reported. The losses due to the suspended arms were minimised by expanding the waveguide (Figure 2.24) so that the overlap between the waveguide mode and the suspension mechanism was minimal [48].

For the fabrication of a free-standing silicon waveguide shown in Figure 2.25 [49], an advanced technique, that of proton beam writing (PBW), has been used. As this waveguide has an air cladding, it is potentially very interesting for long-wavelength applications [45, 50]. Alternatively, any other suitable cladding could be deposited to make the structures suitable for specific wavelength ranges, and would also ruggedise the waveguides that are very fragile

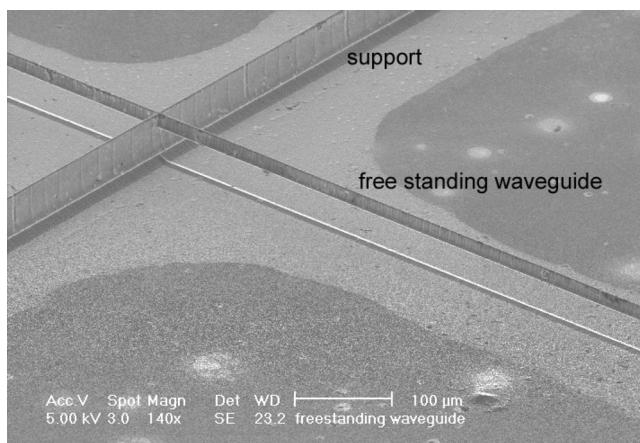


**Figure 2.24** SEM of the suspended ring resonator with expanded waveguide sections [49]

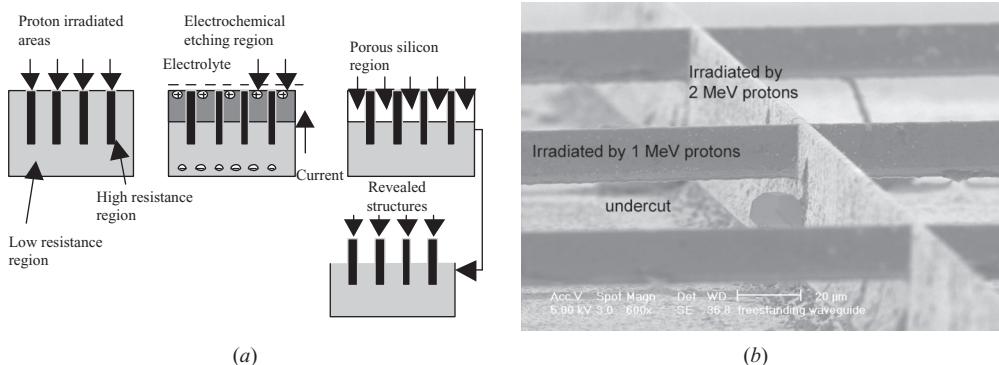
whilst in their free standing form. The waveguide was fabricated by fast-proton irradiation prior to electrochemical etching.

PBW is an advanced lithography technique for micromachining in the submicrometre scale. PBW utilises a high-energy ion microbeam to irradiate suitable resists (e.g. SU-8 and PMMA). Subsequent wet etching is then employed to develop the resist. Whilst it was initially investigated for producing latent microstructures in high molecular weight PMMA resist, Polesello *et al.* [51] and Mistry *et al.* [52] have demonstrated PBW as a method to fabricate three-dimensional structures directly in Si and GaAs, respectively, thereby eliminating the need for a resist.

There are several potential advantages of PBW. It is a direct write process which eliminates the need for a potentially costly mask. Furthermore, prototype devices which require small structural modifications can be produced by simply modifying the scan of the microbeam,



**Figure 2.25** Free-standing waveguide [51]



**Figure 2.26** (a) fabrication process, (b) fabricated waveguides and a support [51]

whereas modifications to a device using photolithographic methods would require the layout and fabrication of an additional mask. A second benefit is that the protons are deposited in the semiconducting material in a well-defined range, thereby allowing for good vertical control of the fabrication of devices. Furthermore, the high-energy protons deviate little from the ideal straight path due to multiple small-angle Coulomb scattering.

Proton beam writing has already demonstrated its potential to be a next-generation lithography technique in some silicon-based applications. Recently, free-standing bridges, multilevel structures, and high-aspect-ratio nano-tips fabricated in silicon have been demonstrated by Breese *et al.* [53], and Teo *et al.* [54, 55].

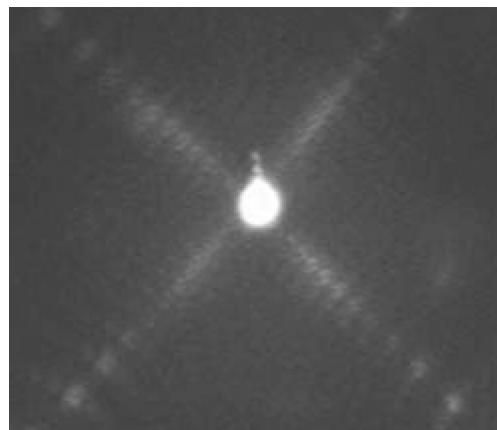
Figure 2.26(a) shows the basis of this fabrication method. A finely-focused proton microbeam is selectively scanned over the silicon surface. The high-energy protons penetrate the silicon substrate and stop within a well-defined range. The depth is controlled by the energy of the protons. As the protons propagate through the silicon, they undergo a series of collisions with electrons. Protons are deflected away from an ideal straight path, resulting in a beam cross-section increasing in size with depth. The energy deposited due to nuclear collisions per unit length is fairly constant during the first half of the penetration depth, but increases sharply at the end of range, because the speed of protons is much lower than their initial value, thereby increasing the probability of nuclear collisions with target atoms. High-energy protons lose their energy and eventually come to rest below the surface. Silicon vacancies are created along the path of the protons, with most of the damage occurring towards the end of their range.

After irradiation, the sample is electrochemically etched in a solution of HF:water:ethanol in a ratio of 1:1:2. Once the silicon becomes porous, it can be removed by a dilute KOH solution. In the irradiated regions, it is important that the density of defects at the surface of the silicon be high enough to significantly reduce the electrical hole current flow and therefore inhibit the electrochemical reaction at the damaged silicon–electrolyte interface. The result is a significant reduction in the rate of production of porous silicon in the proton irradiation regions compared with the unirradiated region. The final structure produced after etching is a three-dimensional representation of the scanned area. This area determines the dimensions of the final structure. The height of the structure is controlled by the etching time and current.

In order to fabricate free-standing waveguides, two different ion energies are required. A high-energy implant is used to create pillars upon which the waveguides will be supported. A lower-energy implant is subsequently used to create the waveguides. The pillars in [49] were created by irradiating the silicon with a dose of  $0.7 \times 10^{15}$  protons/cm<sup>2</sup> at an energy of 2 MeV and the waveguides with a dose of  $0.6 \times 10^{15}$  protons/cm<sup>2</sup> at an energy of 1 MeV (Figure 2.26b). Two scans of the microbeam were made over the same position in order to create the localized damage pattern. The doses are significant enough to inhibit the current flow in the irradiated regions, thereby limiting the formation process of porous silicon during electrochemical etching.

The electrochemical etch rate is mainly dependent on three important parameters: dopant type and concentration, HF concentration, and applied current density. Initially, the etching occurs in all regions except in the irradiated portions. As the etch continues, it will begin to undercut the unirradiated region that lies underneath the 1 MeV irradiated region. The 2 MeV regions will continue to remain unetched thereby creating the pillars to hold the waveguides (Figures 2.25 and 2.26b). In this way, the free-standing waveguide structures are fabricated in a single etch step.

The waveguide has a ‘tear-drop’ shape with a height of 17  $\mu\text{m}$ , and width of 3.7  $\mu\text{m}$  at the top and 8  $\mu\text{m}$  at the bottom. This shape is due to lateral spreading of the beam with depth, but it can be altered by alternative choices of irradiation energies. Propagation loss measurements were performed at a wavelength of 1550 nm in order to compare with the loss of standard (SOI) silicon waveguides prevalent in the literature. The light output from the waveguides was collected by an objective lens and imaged using an IR-sensitive camera (Figure 2.27). The optical power transmitted by the waveguide was also measured with a power meter, and propagation loss was determined using the cut-back method [11]. In order to polish back the waveguides, they were first potted in a transparent wax (Crystalbond), which prevented the waveguides from moving and hence from being damaged during polishing. The propagation loss was determined to be  $13.4 \pm 0.7$  dB/cm for TE and  $14.6 \pm 0.6$  dB/cm for TM polarised light respectively. The sidewall roughness of the waveguide and irradiation damage of the waveguide itself during the PBW are probably two main reasons for the high propagation losses obtained. The authors expected that



**Figure 2.27** Output from a free standing waveguide [51]

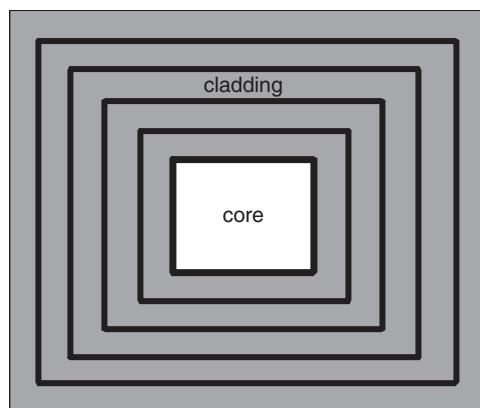
the irradiation damage could be reduced by modification of the fabrication process to include an annealing step [49]. Oxidising the waveguides may also help to alleviate the roughness issue of the waveguide side walls. Additional loss may be a result of the shape and dimensions of the waveguide, and the intersection with the support pillars. Therefore, fabrication of square cross-section waveguides with smaller dimensions, and minimisation of interaction with the supports are currently being investigating. The propagation loss values are rather high, they are nonetheless promising especially when compared with the propagation loss of 25 dB/cm that was measured for early SOI waveguides, e.g. [56].

### 2.6.2 Hollow Waveguides

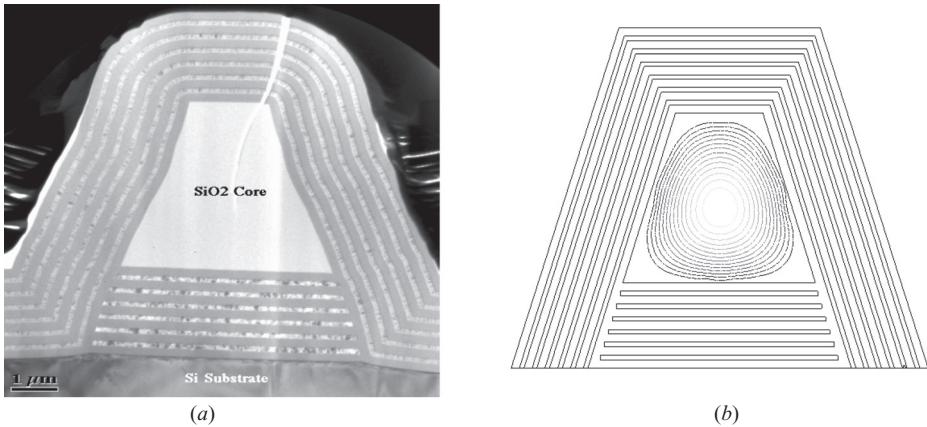
So far we have been investigating conventional waveguides, in the sense that their core had higher refractive index than the cladding. Due to this fact, and the high confinement in SOI-based waveguide structures, their sensor applications are somewhat limited. If it is desirable for light to propagate through small volumes of liquids or gases on a chip, for example, in chemical sensing, fluorescence spectroscopy in biology, biochemistry, biomedicine, flow cytometry, pollution monitoring, etc. [57], then there is very limited range of available cladding materials. Currently, the only available material suitable for liquid core waveguides is Teflon AF with refractive index that is larger than 1.29 [58]. As the refractive index of water is around 1.33 and air has a refractive index of 1, this means that waveguides with Teflon AF cladding will have large dimensions for an aqueous core, and are not possible for an air core.

However, a hollow channel waveguide whose walls consist of multilayer coatings can be very useful for sensor applications as its core can be infilled with gases or liquids. Other potential applications of these waveguides include tight turning radii, high power transmission, and dispersion compensation.

The walls of the waveguide consist of alternating layers: material A and material B (Figure 2.28), where A (shown as a dark layer) has a higher refractive index than B. The hollow waveguides operate in two categories: those with a Bragg-mirror cladding [59] and those



**Figure 2.28** Cross-section of a hollow waveguide



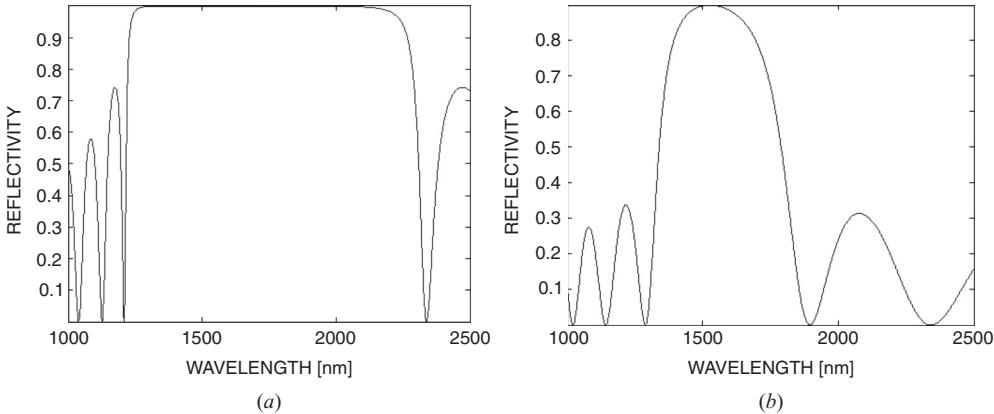
**Figure 2.29** (a) TEM image of the Bragg cladded channel waveguide [61]; (b) mode profile for trapezoidal shape omnidirectional waveguide with SiO<sub>2</sub> core and Bragg mirror made of six Si/Si<sub>3</sub>N<sub>4</sub> double layers [63]

with an ARROW (anti-resonant reflecting optical waveguide) cladding [60]. In the Bragg case, light is totally reflected at the multilayer, for any incident angle or polarisation. For the ARROW, the layered wall is similar to a Fabry–Perot interferometric reflector. Whilst in the Bragg case there are multiple A/B layers ( $N$ ), in the ARROW case this number can vary from only 1 to  $N$ .

There are a number of papers in the literature about ARROW waveguides, e.g. [57, 60–62], but only a few have reported results for Bragg-mirror claddings in silicon waveguide technology. The MIT group reported omnidirectional waveguide with  $N = 6$  bilayers of Si<sub>3</sub>N<sub>4</sub>/Si in the Bragg-mirror and SiO<sub>2</sub> core [59]. Each layer in the mirror had a quarter-wavelength thickness at the targeted wavelength of 1550 nm. The low-pressure chemical vapour deposition (LPCVD) method was used to deposit the Si and Si<sub>3</sub>N<sub>4</sub> cladding layers and the low-temperature oxide (LTO) method was used for the oxide core. The waveguide had a trapezoidal shape and dimensions of 4–6 μm (Figure 2.29a). The minimum propagation loss for this waveguide was 6 dB/cm. Figure 2.29 (b) shows the theoretical fundamental optical mode for this waveguide structure.

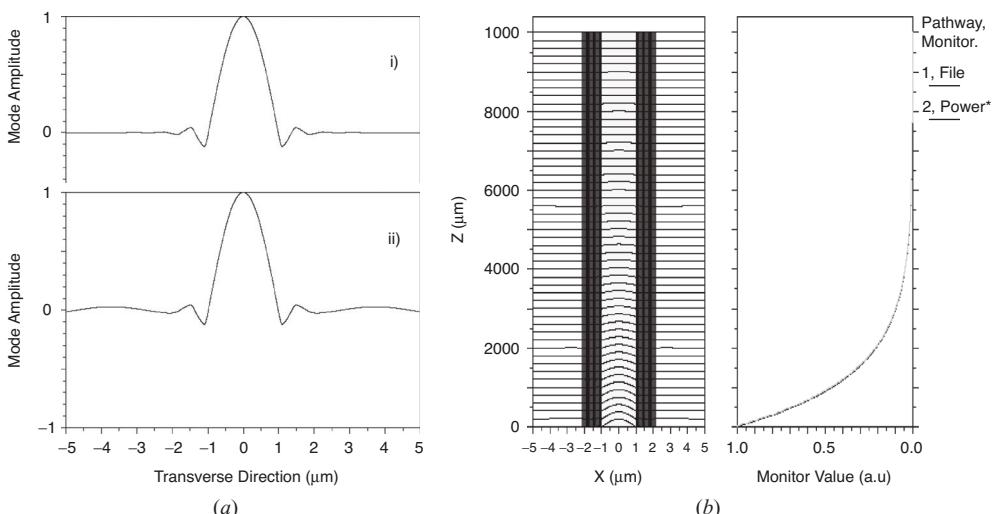
As in ARROW waveguides, for effective reflection, the refractive index difference between the core and the first cladding material should be as large as possible, whilst the refractive index of the second cladding material should be as close as possible to the core refractive index [60]. Therefore, a waveguide which consists of a Si/SiO<sub>2</sub>, rather than a Si/Si<sub>3</sub>N<sub>4</sub> reflector, could have a lower propagation loss for the same number of layers. Another advantage of a Si/SiO<sub>2</sub> reflector over the Si/Si<sub>3</sub>N<sub>4</sub> one, is a lower deposition temperature for the PECVD method compared with the LPCVD.

For a Si/SiO<sub>2</sub> reflector where the refractive index difference is almost 2, a reflectivity of 100% can be achieved for a broad range of wavelengths, as shown in Figure 2.30 (a). On the other hand, if a Si<sub>3</sub>N<sub>4</sub>/SiO<sub>2</sub> reflector is used, for example, where that difference is only 0.5, the 100% reflectivity range becomes much narrower (Figure 2.30b) and more dependent on structure parameters and fabrication tolerances.

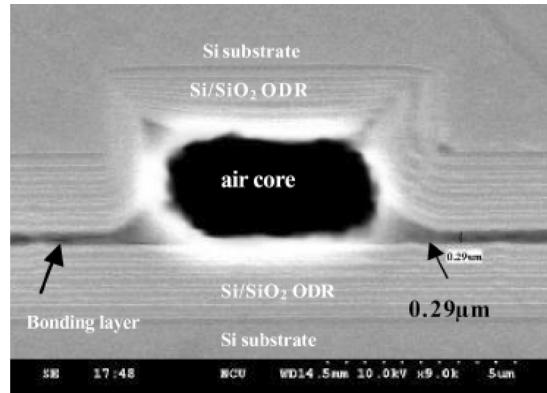


**Figure 2.30** Reflectivity of six pairs of: (a) Si/SiO<sub>2</sub>; (b) Si<sub>3</sub>N<sub>4</sub>/SiO<sub>2</sub> layers

For a smaller number of Si/SiO<sub>2</sub> bi-layers, the spectrum also becomes narrower and the reflectivity at 1550 nm can drop below 100%. That can, in turn, cause significant propagation loss, as shown in Figure 2.31. The modelled fundamental mode profiles for a waveguide with 2 μm thick air core, and Si/SiO<sub>2</sub> reflector are shown in Figure 2.31 (a). Figure 2.31 (a)(i) and Figure 2.31 (a)(ii) show the mode profile for a waveguide with a six bi-layer reflector and a three bi-layer reflector, respectively. It is clear that the optical power is well confined within the core in the former case, whilst it leaks into the substrate in the later case. Therefore, the propagation loss for the three bi-layer structure is predicted to be 24 dB/cm (Figure 2.31b),



**Figure 2.31** (a) Mode profile for a waveguide that consists of an air core and a Si/SiO<sub>2</sub> reflector with six (i) and three bi-layers (ii); (b) propagation of light in a three-bilayer Si/SiO<sub>2</sub> reflector waveguide [63]



**Figure 2.32** SEM image of the cross-section of the hollow waveguide fabricated by wafer bonding [64]

and only 0.07 dB/cm for the six bi-layer one. These results were obtained by 2D BeamPROP simulations [18]. Comparative predictions of the propagation loss for hollow waveguides with Si/SiO<sub>2</sub> and Si/Si<sub>3</sub>N<sub>4</sub> reflectors have also been made, and it was clear that Si/SiO<sub>2</sub> outperforms Si/Si<sub>3</sub>N<sub>4</sub> structure. For example, propagation loss for a hollow waveguide with 2 μm core, and a six bi-layer Si/Si<sub>3</sub>N<sub>4</sub> mirror is 2.5 dB/cm, significantly higher than the loss for the Si/SiO<sub>2</sub> six bi-layer mirror (0.07 dB/cm).

The fabrication process for the waveguide with Si/SiO<sub>2</sub> Bragg-mirrors can be based on the plasma-enhanced chemical vapour deposition (PECVD) as in [63], or on the combination of the PECVD and wafer bonding [64]. Lo *et al.* [64], fabricated the hollow waveguide shown in Figure 2.32, using the following recipe. The trench was defined and etched via photolithography and inductive coupled plasma (ICP) etching. Six Si/SiO<sub>2</sub> bi-layers were then deposited by PECVD on the trenched silicon wafer and another blank silicon wafer. The thicknesses of the Si and SiO<sub>2</sub> layers were 111 nm and 258 nm, respectively. A 2% dilute KOH solution was then applied, before pre-bonding at a temperature of 70° C for 10 min. Finally, the pre-bonded wafer was annealed at 200° C for 2 h. The authors measured the propagation loss to be 1.0 ± 0.5 dB/cm. Such a low loss was attributed to the significantly reduced air gap between the two wafers (Figure 2.32).

The previously described waveguides are not suitable for long-wavelength infra red (LWIR) applications due to absorption loss of the SiO<sub>2</sub> layers, as discussed earlier. Therefore, alternative claddings need to be used. One possible solution may be SiGe/Si.

## 2.7 Summary and Conclusions

Among several platforms considered for silicon photonics, silicon-on-insulator has become by far the most popular platform. In the early days, silicon on insulator waveguides had rather high propagation loss, which has been reduced to the order of 0.1 dB/cm owing to improvements in design and fabrication. At first silicon waveguides also had rather large dimensions in order to reduce the insertion loss when coupled to/from an optical fibre. The current trend is, however, to move to smaller device dimensions for improved cost, efficiency and device performance.

CMOS compatibility is another reason for this move and therefore small strip waveguides are becoming a very interesting building block for silicon photonic circuits. However, these waveguides have higher propagation losses of 2–3 dB/cm. Furthermore, it is difficult to achieve polarisation independence in devices based on these small waveguides. As the cross-section of these waveguides is much smaller than the cross-section of an optical fibre, coupling to/from fibres remains a very important issue.

The question of whether to pursue polarisation independence in silicon waveguides is related to the application in question. For example, it is likely that for optimum modulation speed, a strip waveguide-based modulator placed directly in front of a laser would not be required to exhibit polarisation-independence because the laser is inherently polarised. Alternatively, a modulator fed via an optical fibre is much more likely to be required to provide polarisation independent performance due to the random polarisation of light emerging from a fibre. Fortunately, the flexibility of silicon photonics means there is room for both approaches, and each approach has advantages in some application areas.

Mid- and long-wave infrared photonic applications are becoming more interesting and therefore new waveguiding structures suitable for those wavelength ranges need to be found. In this chapter, two such promising structures have been considered. One is a free-standing waveguide which has an air cladding and hence is suitable for the long-wave infrared region, whilst the other is a hollow waveguide which can be useful for sensor applications.

## References

- [1] R. A. Soref and J. P. Lorenzo, ‘All-silicon active and passive guided-wave components for  $\lambda = 1.3$  and  $1.6 \mu\text{m}$ ’, *IEEE J. Quantum Electron.*, **QE-22**, 1986, 873–879.
- [2] R. A. Soref and J. P. Lorenzo, ‘Single-crystal silicon: a new material for  $1.3$  and  $1.6 \mu\text{m}$  integrated-optical components’, *Electron. Lett.*, **21**, 1985, 953–954.
- [3] D. J. Albares and R. A. Soref, ‘Silicon-on-sapphire waveguides’, *Proc. SPIE*, **704**, 1987, 24–25.
- [4] R. A. Soref, F. Namavar and J. P. Lorenzo, ‘Optical waveguiding in a single-crystal layer of germanium silicon grown on silicon’, *Opt. Lett.*, **15**, 1990, 270–272.
- [5] E. Cortesi, F. Namavar and R. A. Soref, ‘Novel silicon-on-insulator structures for silicon waveguides’, *IEEE SOS/SOI Technology Conference (Cat. No.89CH2796-1)*, 3–5 October 1989. Stateline, NV, USA: IEEE, 1989, p. 109.
- [6] B. N. Kurdi and D. G. Hall, ‘Optical waveguides in oxygen-implanted buried-oxide silicon-on-insulator structures’, *Opt. Lett.*, **13**, 1988, 175–177.
- [7] A. Rickman, G. T. Reed, B. L. Weiss and F. Namavar, ‘Low-loss planar optical waveguides fabricated in SIMOX material’, *IEEE Photon. Technol. Lett.*, **4**, 1992, 633–635.
- [8] J. Schmidtchen, A. Splett, B. Schuppert and K. Petermann, ‘Low loss integrated-optical rib-waveguides in SOI’, *1991 IEEE International SOI Conference Proceedings (Cat. No.91CH3053-6)*, 1991, pp. 142–143.
- [9] S. P. Chan, C. E. Png, S. T. Lim, V. M. N. Passaro and G. T. Reed, ‘Single mode and polarisation independent SOI waveguides with small cross section’, *J. Lightwave Technol.*, **23**, 2005, 1573–1582.
- [10] B. Timotijevic, G. T. Reed, R. Jones, A. Liu, A. Michaeli and G. Mashanovich, ‘Optical filters in silicon-on-insulator: Design considerations for devices based upon strip and rib’, *Proc SPIE*, **6350**, 2006, 63500K.
- [11] G. T. Reed and A. P. Knights, *Silicon Photonics: An Introduction*, John Wiley, Chichester, England, 2004.
- [12] MATLAB by Math works, <http://www.mathworks.com>.
- [13] R. A. Soref, J. Schmidtchen and K. Petermann, ‘Large single-mode rib waveguides in Ge-Si and Si-on-SiO<sub>2</sub>’, *IEEE J. Quantum Electron.*, **27**, 1991, 1971–1974.
- [14] S. Pogossian, L. Vescan and A. Vonsovici, ‘The single-mode condition for semiconductor rib waveguides with large cross section’, *J. Lightwave Technol.*, **16**, 1998, 1851–1853.
- [15] O. Powell, ‘Single-mode condition for silicon rib waveguides’, *J. Lightwave Technol.*, **20**, 2002, 1851–1855.

- [16] L. Vivien, S. Laval, B. Dumont, S. Lardenois, A. Koster and E. Cassan, ‘Polarization-independent single-mode rib waveguides on silicon-on-insulator for telecommunication wavelengths’, *Opt. Commun.*, **210**, 2002, 43–49.
- [17] J. Lousteau, D. Furniss, A. Seddon, T. M. Benson, A. Vukovic and P. Sewell, ‘The single-mode condition for silicon-on-insulator optical rib waveguides with large cross section’, *J. Lightwave Technol.*, **22**, 2004, 1923–1929.
- [18] BeamPROP by Rsoft Design Group, Inc., [www.rsoftdesing.com](http://www.rsoftdesing.com).
- [19] FEMLAB by COMSOL, Inc., [www.comsol.com](http://www.comsol.com).
- [20] S. P. Chan, V. M. N. Passaro and G. T. Reed, ‘Single mode condition and zero birefringence in small SOI waveguides’, *Electron. Lett.*, **41**, 2005, 528–529.
- [21] D. X. Xu, W. N. Ye, A. Bogdanov, D. Dalacu, A. Delage, P. Cheben, S. Janz, B. Lamontagne and M.-J. Picard, ‘Design of polarization-insensitive components using geometrical and stress-induced birefringence in SOI waveguides’, *Proc. SPIE*, **5730**, 2005, 158–172.
- [22] W. N. Ye, D.-X. Xu, S. Janz, P. Cheben, M.-J. Picard, B. Lamontagne and N. G. Tarr, ‘Birefringence control using stress engineering in silicon-on-insulator (SOI) waveguides’, *J. Lightwave Technol.*, **23**, 2005, 1308–1318.
- [23] W. R. Headley, G. T. Reed, M. Paniccia, A. Liu, O. Cohen and D. Hak, ‘Polarization-independent optical racetrack resonators using rib waveguides on silicon-on-insulator’, *Appl. Phys. Lett.*, **85**, 2004, 5523–5525.
- [24] H. Deng, D. O. Yevick, C. Brooks and P. E. Jessop, ‘Design rules for slanted-angle polarization rotators’, *J. Lightwave Technol.*, **23**, 2005, 432–445.
- [25] W. R. Headley, ‘Optical ring resonators in silicon-on-insulator’, PhD Thesis, University of Surrey, 2005.
- [26] F. Y. Gardes, G. T. Reed, N. G. Emerson and C. E. Png, ‘A sub-micron depletion-type photonic modulator in silicon on insulator’, *Opt. Express*, **13**, 2005, 8845–8854.
- [27] G. T. Reed, W. R. Headley, F. Y. Gardes, B. D. Timotijevic, S. P. Chan and G. Z. Mashanovich, ‘Characteristics of rib waveguide racetrack resonators in SOI’, *Proc. SPIE*, **6183**, 2006, 61830G.
- [28] K. Yamada, H. Fukuda, T. Watanabe, T. Tsuchizawa, T. Shoji and S.-I. Itabashi, ‘Functional photonic devices on silicon wire waveguide’, *Proc. LEOS 2nd Group IV Photonics*, Antwerp, Belgium, September 2005, 186–188.
- [29] T. Alto, *Microphotonic silicon waveguide components*, (ISBN 951-38-6422-7) VTT, Finland, 2004.
- [30] Y. A. Vlasov and S. J. McNab, ‘Losses in single-mode silicon-on-insulator strip waveguides and bend’, *Opt. Express*, **12**, 2004, 1622–1627.
- [31] T. Tsuchizawa, K. Yamada, H. Fukuda, T. Watanabe, J.-I. Takahashi, M. Takahashi, T. Shoji, E. Tamechika, S.-I. Itabashi and H. Morita, ‘Microphotonics devices based on silicon microfabrication technology’, *IEEE J. Select. Topics Quantum Electron.*, **11**, 2005, 232–240.
- [32] T. Tsuchizawa, T. Watanabe, E. Tamechika, T. Shoji, K. Yamada, J. Takahashi, S. Uchiyama, S. Itabashi and H. Morita, ‘Fabrication and evaluation of submicron-square Si-wire waveguides with spot-size converter’, *Proc. 15th Annual IEEE LEOS Meeting*, 2002, pp. 287–288.
- [33] G. J. Veldhuis, O. Parriaux, H. J. W. M. Hoekstra and P. V. Lambeck, ‘Sensitivity enhancement in evanescent optical waveguide sensors’, *J. Lightwave Technol.*, **18**, 2000, 677–682.
- [34] J. J. Fijol, E. E. Fike, P. B. Keating, D. Gilbody, J. J. LeBlanc, S. A. Jacobson, W. J. Kessler and M. B. Frisch, ‘Fabrication of silicon-on-insulator adiabatic tapers for low loss optical interconnection on photonic devices’, *Proc. SPIE*, **4997**, 2003, 157–170.
- [35] A. Sure, T. Dillon, J. Murakowski, C. Lin, D. Pustai and D. W. Prather, ‘Fabrication and characterization of three-dimensional silicon tapers’, *Opt. Express*, **11**, 2003, 3555–3561.
- [36] V. R. Almeida, R. R. Panepucci and M. Lipson, ‘Nanotaper for compact mode conversion’, *Opt. Lett.*, **28**, 2003, 1302–1304.
- [37] S. J. McNab, N. Moll and Y. A. Vlasov, ‘Ultra-low loss photonic integrated circuit with membrane-type photonic crystal waveguides’, *Opt. Express*, **11**, 2003, 2927–2939.
- [38] K. Yamada, M. Notomi, I. Yokohama, T. Shoji, T. Tsuchizawa, T. Watanabe, J. Takahashi, E. Tamechika and H. Morita, ‘SOI-based photonic crystal line defect waveguides’, *Proc. SPIE*, **4870**, 2002, 324–338.
- [39] G. Z. Masanovic, V. M. N. Passaro and G. T. Reed, ‘Dual grating-assisted directional coupling between fibres and thin semiconductor waveguides’, *IEEE Photon. Technol. Lett.*, **15**, 2003, 1395–1397.
- [40] G. Z. Masanovic, G. T. Reed, W. Headley, B. Timotijevic, V. M. N. Passaro, R. Atta, G. Ensell and A. G. R. Evans, ‘A high efficiency input/output coupler for small silicon photonic devices’, *Opt. Express*, **13**, 2005, 7374–7379.
- [41] T. W. Ang, G. T. Reed, A. Vonsovici, A. G. R. Evans, P. R. Routley and M. R. Josey, ‘Effects of grating heights on highly efficient Unibond SOI waveguide grating couplers’, *IEEE Photon. Technol. Lett.*, **12**, 2000, 59–61.
- [42] T. W. Ang, G. T. Reed, A. Vonsovici, A. G. R. Evans, P. R. Routley and M. R. Josey, ‘Highly efficient Unibond silicon-on-insulator blazed grating couplers’, *Appl. Phys. Lett.*, **77**, 1999, 4214–4216.

- [43] C. Gunn, ‘10 Gb/s CMOS photonics technology’, *Proc. SPIE*, **6125**, 2006, 612501.
- [44] C. Gunn, ‘CMOS photonics for high-speed interconnects’, *IEEE Micro*, **26**, 2006, 58–66.
- [45] R. A. Soref, S. J. Emelett and W. R. Buchwald, ‘Silicon waveguided components for the long-wave infrared region’, *J. Opt. A: Pure Appl. Opt.*, **8**, 2006, 840–848.
- [46] M. Borselli, K. Srinivasan, P. E. Barclay and O. Painter, ‘Rayleigh scattering, mode coupling, and optical loss in silicon microdisks’, *Appl. Phys. Lett.*, **85**, 2004, 3693–3695.
- [47] L. Martinez and M. Lipson, ‘High confinement suspended micro-ring resonators in silicon-on-insulator’, *Opt. Express*, **14**, 2006, 6259–6263.
- [48] T. Fukazawa, T. Hirano, F. Ohno and T. Baba, ‘Low loss intersection of Si photonic wire waveguides’, *Jap. J. Appl. Phys.*, **43**, 2004, 646–647.
- [49] P. Y. Yang, G. Z. Mashanovich, I. Gomez-Morilla, W. R. Headley, G. T. Reed, E. J. Teo, D. J. Blackwood, M. B. H. Breese and A. A. Bettoli, ‘Free standing waveguides in silicon’, *Appl. Phys. Lett.*, **90**, 2007, 241109.
- [50] B. Jalali, V. Raghunathan, R. Shori and O. M. Stafsudd, ‘Mid-infrared silicon photonics’, *Proc. MRS Fall Meeting 2006*, 27 November – 1 Decembr 2006, Boston, USA, L11.1.
- [51] P. Polesello, C. Manfredotti, F. Fizzotti, R. Lu, E. Vittone, G. Lerondel, A. M. Rossi, G. Amato, L. Boarino, S. Galassini, M. Jakšić and Z. Pastuovic, ‘Micromachining of silicon with a proton microbeam’, *Nuclear Instr. Meth. Phys. Res. B*, **158**, 1999, 173–178.
- [52] P. Mistry, I. Gomez-Morilla, G. W. Grime, R. Webb, C. Jeynes, R. Gwilliam, A. Cansell, M. Merchant and K. J. Kirkby, ‘Proton beam lithography at the University of Surrey’s Ion Beam Centre’, *Nuclear Instr. Meth. Phys. Res. B*, **242**, 2006, 387–389.
- [53] M. B. H. Breese, F. J. T. Champeaus, E. J. Teo, A. A. Bettoli and D. J. Blackwood, ‘Hole transport through proton-irradiated p-type silicon wafers during electrochemical anodisation’, *Phys. Rev. B*, **B73**, 2006, 035428.
- [54] E. J. Teo, M. H. Liu, M. B. H. Breese, E. P. Tavernier, A. A. Bettoli, D. J. Blackwood and F. Watt, ‘Fabrication of silicon microstructures using a high energy ion beam’, *Proc. SPIE*, **5347**, 2004, 264–270.
- [55] E. J. Teo, M. B. H. Breese, E. P. Tavernier, A. A. Bettoli and F. Watt, ‘Three-dimensional microfabrication in bulk silicon using high-energy protons’, *Appl. Phys. Lett.*, **84**, 2004, 3202–3204.
- [56] B. L. Weiss, G. T. Reed, S. K. Toh, R. A. Soref and F. Namavar, ‘Optical waveguides in SIMOX structures’, *IEEE Photon. Technol. Lett.*, **3**, 1991, 19–21.
- [57] D. Yin, D. W. Deamer, H. Schmidt, J. P. Barber and A. R. Hawkins, ‘Integrated optical waveguides with liquid cores’, *Appl. Phys. Lett.*, **85**, 2004, 3477–3479.
- [58] H. Schmidt, D. Yin, J. P. Barber and A. R. Hawkins, ‘Hollow-core waveguides in 2-D waveguide arrays for integrated optics of gases and liquids’, *IEEE J. Select. Topics Quantum Electron.*, **11**, 2005, 519–527.
- [59] Y. Yi, S. Akiyama, P. Bermel, X. Duan and L. C. Kimerling, ‘On-chip Si-based Bragg cladding waveguide with high index contrast bilayers’, *Opt. Express*, **12**, 2004, 4775–4780.
- [60] R. Bernini, S. Campopiano and L. Zeni, ‘Silicon micromachined hollow optical waveguides for sensing applications’, *IEEE J. Select. Topics Quantum Electron.*, **8**, 2002, 106–110.
- [61] M. A. Duguay, Y. Kokubun, T. L. Koch and L. Pfeiffer, ‘Antiresonant reflecting optical waveguides in SiO<sub>2</sub>-Si multilayer structures’, *Appl. Phys. Lett.*, **49**, 1986, 13–15.
- [62] J. P. Barber, D. B. Conkey, J. R. Lee, N. B. Hubbard, L. L. Howell, D. Yin, H. Schmidt and A. R. Hawkins, ‘Fabrication of hollow waveguides with sacrificial aluminium cores’, *IEEE Photon. Technol. Lett.*, **17**, 2005, 363–365.
- [63] G. Z. Mashanovich, G. Pucker, C. Kompocholis, A. Liu, S. Stannovic, J. Crnjansni, V. M. N. Passaro, P. Matavulj and G. T. Reed, ‘Omnidirectional silicon photonic waveguides,’ *Proc. 14 Telecommunications Forum, TELFOR*, 2006, 357–360.
- [64] S. S. Lo, C.-C. Chen, S.-C. Hsu and C.-Y. Liu, ‘Fabricating a hollow optical waveguide for optical communication applications’, *J. Microelectromechanical Systems*, **15**, 2006, 584–587.

# 3

## Silicon-based Photonic Crystal Structures: From Design to Realization

Dennis W. Prather, Shouyuan Shi, Janusz Murakowski, Garrett Schneider, Ahmed Sharkawy, Caihua Chen and BingLin Miao

### 3.1 Theory of Photonic Crystals and Photonic Bandgap Structures

Throughout the last two decades there has been significant activity in the development of photonic devices that can confine, control, and route light on a scale comparable to modern electronic devices, namely the nanometer scale. A key motivation for this is to realize photonic circuits having a density approaching that of modern electronic circuits. However, in order for this to be done, such devices would need the ability to confine light on a subwavelength scale and exist in a material compatible with the microelectronics manufacturing infrastructure. While the latter requirement was readily satisfied through a proper choice of materials, i.e., silicon, the former one was more illusive. The reason for this arose from the fact that reflective, or conducting, devices are very lossy at optical wavelengths and refractive, or total internally reflective, devices do not offer mode confinement on a small enough scale. For this reason, researchers turned to the field of photonic crystal (PhC) devices, and their associated photonic bandgap devices, which offer both low loss, high confinement, and can be readily fabricated in silicon. However, before we begin discussing the various aspects of silicon based PhCs, we first present a brief perspective on their current status.

The physical phenomenon that clearly describes the operation of a photonic crystal is the localization of light, which is achieved from the scattering and interference produced by a coherent wave in a periodic structure. Upon an incident radiation, the periodic scatterers constructing a photonic crystal could reflect an incident radiation at the same frequency in all directions. Then, wherever in space the radiated radiation interferences constructively,

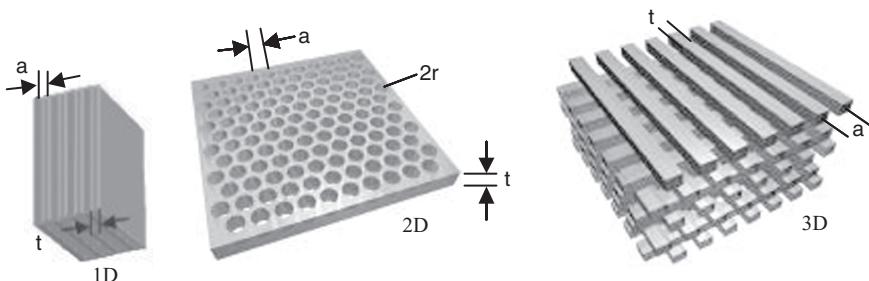
sharp peaks would be observed. This portion of the radiation spectrum is then forbidden to propagate through the periodic structure and this band of frequencies is what was called a stop band, or a photonic bandgap. On the other hand, wherever in space an incident radiation destructively interferes with the periodic scatterers in a certain directions, this part of the radiation spectrum will propagate through the periodic structure with minimal attenuation, and this band of frequencies was called a pass band.

For an electromagnetic wave propagating within a dielectric material, scattering takes place on a scale much larger than the wavelength of light. The localization of light occurs when the scale of the coherent multiple scatterers are reduced to the wavelength itself. In this case, a photon located in a lossless dielectric media provides an ideal realization of a single excited state in a static medium, at room temperature. Unlike electron localization, which requires an electron-electron interaction and electron-phonon interactions, photon localization offers a unique possibility of studying the angular, spatial, and temporal dependence of wave field intensities near localized transitions.

Light localization has fundamental consequences at the quantum level. This can be seen for a periodic array of high dielectrics that have dimensions comparable to the wavelength of light, by exhibiting a complete photonic bandgap in certain range, analogous to the electronic energy bandgap in a semiconductor material. In a photonic crystal, there are no allowed electromagnetic states in the forbidden frequency range.

To a large extent, the field of PhCs can be thought of as having its origins in the rather conventional one-dimensional (1D) thin-film stack, wherein a quarter wave thickness of alternating materials is arranged in a periodic fashion. In this device, it is well known that certain frequencies are transmitted while others are not. As the index contrast between the alternating layers is increased, the selectivity of the transmitted and reflected wavelengths is also increased.

When extended to two and three dimensions, this selectivity is what gives rise to high mode confinement. Figure 3.1 illustrates photonic crystal structures in one, two and three dimensions. While 1D thin-film stacks have been known for over a century, their generalization to higher dimensions was not proposed until the 1970s by Bykov [1, 2] as a possible way of inhibiting spontaneous emission. In essence, Bykov proposed the use of a periodic structure that served to inhibit certain electromagnetic frequencies, thereby disallowing spontaneous emission. These devices ultimately became known as photonic bandgap (PBG) structures, as they suppressed a band of frequencies from existing, the so-called photonic bandgap. In



**Figure 3.1** Periodic structure in one, two, and three dimensions is the origin where photonic crystals emerged

similar vein, Yablonovitch [3] proposed a structure where an electronic and photonic gap overlapped, thereby making it possible to enhance the performance of lasers, heterojunctions bipolar transistors, and solar cells. Subsequent to this work, and John [4] proposed using such structures for the localization of light in strongly scattering dielectric structures. In each of these cases, the basic idea was to tailor the properties of photons in a PhC in a way directly analogous to how atomic crystals tailor the properties of electrons. That is to say, in the electronic case the wave functions of electrons interact with the periodic potential of the atomic lattice and for a certain range of energies (similar to frequencies for photons) electronic states cease to exist, thus, giving rise to an electronic bandgap. For PhCs, the analogue of the electronic potential in an atomic crystal is the dielectric constants of the constituent materials of the PBG structure. In addition, due to the periodic interaction, certain photonic bandgaps appear wherein certain modes, or frequencies, are forbidden. In such a structure, one can then introduce a line, or point, defect, which accounts to the absence of the periodic lattice, wherein a mode is localized by virtue of being suppressed within the lattice. For this reason, these devices offer extreme mode confinement as well as the ability to control and route light very efficiently.

With these advances, the field of PhCs and in particular their realization in silicon-based materials has been a very active field of research over the last two decades. During this time, great success was achieved in identifying suitable periodic structures, dielectric materials and both theoretical as well as experimental demonstration and characterization of one-, two-, and three-dimensional PhC and bandgap structures. However, the challenge remains to develop suitable structures for novel device applications and systems that are economically feasible. This difficulty arises from the challenging aspects of their fabrication, which often requires high-resolution lithography and high-aspect-ratio etching. While progress has been, and is being, made in the areas significant market opportunities have yet to arise. Nonetheless, the field of PhCs offers significant potential and is poised to find its place in the realm of high-technology applications.

Two-dimensional photonic crystals can be realized using either a periodic array of dielectric rods of any shape and/or geometry, or by using perforated dielectric slab of air holes. Such structures can be further optimized to achieve either a wider, or a narrower, bandgap based on the desired application. Two-dimensional photonic crystals impose periodicity in two dimensions while the third dimension is either infinitely long (photonic crystal fiber) or has a finite height (photonic crystal slabs). Three-dimensional photonic crystals impose periodicity in all three dimensions.

Since they are easier to fabricate and analyze, two-dimensional photonic crystals have attracted the attention of a large number of researchers and engineers. Planar photonic crystal circuits such as splitters [5–9], high Q-microcavities [10–16] and channel drop/add filters [17–20] have been investigated both theoretically [21–23] and experimentally [24–42].

Once a photonic crystal has been designed, its properties can be engineered in a manner similar to that which is done to an electronic crystal. For bandgap-related applications this is attained through the process of doping. In a photonic crystal doping is achieved by either adding or removing dielectric material to a certain area. The area, which we have added, or removed, dielectric material then acts as a defect region that can be used to localize an electromagnetic wave. Doping a photonic crystal opens a broad range of possibilities for optical device development through the localization of light. For non-bandgap-related applications, the dispersion properties of the host photonic crystal structure can be engineered in a similar fashion to implement various functionalities.

In this chapter, we will highlight the major milestone achievements which have highly contributed to the progress attained thus far to the research and development of photonic crystal-based applications. Major contribution to develop the appropriate design and analysis tools, in addition to fabrication and characterization tools, will be discussed. Bottle neck challenges that have limited the progress in various areas, will be highlighted and solutions to resolve them will be also presented.

### 3.1.1 Analogy Between Photonic and Semiconductor Crystals

In a semiconductor crystal, electron localization can be described using Schrödinger equation for an electron with an effective mass  $m^*$

$$\left[ \frac{-\hbar^2}{8\pi^2 m^*} \nabla^2 + V(x) \right] \varphi(x) = E\varphi(x) \quad (3.1)$$

where  $\hbar$  is Planck's constant,  $m^*$  is the effective mass of electron,  $V(x)$  is the potential function,  $\varphi(x)$  is the wavefunction, and  $E$  is the total energy. The probability of finding an electron at  $x$  is given by  $|\varphi(x)|^2$ . The electron can be trapped by a random potential  $V(x)$  in deep local potential fluctuations if the energy  $E$  is sufficiently negative. As the energy increases, the probability for the trapped electron to tunnel to a nearby potential fluctuation also increases.

In the case of monochromatic electromagnetic waves of frequency  $\omega$  propagating in an inhomogeneous, but nondissipative dielectric medium, Maxwell's equations are used to describe the wave propagation through space.

Starting with four macroscopic Maxwell equations

$$\nabla \cdot \mathbf{B} = 0 \quad (3.2)$$

$$\nabla \cdot \mathbf{D} = \rho \quad (3.3)$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad (3.4)$$

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} \quad (3.5)$$

and using the two constitutive equations

$$\mathbf{D}(\mathbf{r}) = \epsilon(\mathbf{r})\mathbf{E}(\mathbf{r}) \quad (3.6)$$

$$\mathbf{B}(\mathbf{r}) = \mu(\mathbf{r})\mathbf{H}(\mathbf{r}) \quad (3.7)$$

while keeping in mind that for a dielectric material

$$\mu(\mathbf{r}) = 1.0, \quad (3.8)$$

we can substitute Equations (3.6), (3.7) and (3.8) into Equations (3.4) and (3.5) write them in frequency (steady-state) domain form. Doing so results in the following equations

$$\nabla \times \mathbf{E}(r) + j\omega\mu(r)\mathbf{H}(r) = 0 \quad (3.9)$$

$$\nabla \times \mathbf{H}(r) - j\omega\epsilon(r)\mathbf{E}(r) = 0 \quad (3.10)$$

In deriving Equation (3.10) we have assumed that there are no sources of current ( $\mathbf{J} = 0$ ). Taking the curl of Equation (3.9) and using Equation (3.10) to eliminate  $\mathbf{H}(r)$  we get

$$\nabla \times [\nabla \times \mathbf{E}(r)] = \omega^2 \mu(r) \varepsilon(r) \mathbf{E}(r) \quad (3.11)$$

$$\nabla \times [\nabla \times \mathbf{H}(r)] = -\omega^2 \mu(r) \varepsilon(r) \mathbf{H}(r) \quad (3.12)$$

The right-hand side of Equation (3.11) can be further expanded using vector identities

$$-\nabla^2 \mathbf{E}(r) + \nabla (\nabla \cdot \mathbf{E}(r)) = \omega^2 \mu(r) \varepsilon(r) \mathbf{E}(r) \quad (3.13)$$

The total dielectric constant  $\varepsilon(r)$  can be separated into two parts as

$$\varepsilon(r) = \varepsilon_0 + \varepsilon_{\text{spatial}}(r) \quad (3.14)$$

where  $\varepsilon_0$  is the average value of the dielectric function and  $\varepsilon_{\text{spatial}}(\mathbf{r})$  is the spatial component of the dielectric function, which is analogous to the potential  $V(x)$  in Schrödinger's Equation (3.13) can be then written as

$$-\nabla^2 \mathbf{E}(r) + \nabla (\nabla \cdot \mathbf{E}(r)) = \omega^2 [\varepsilon_0 + \varepsilon_{\text{spatial}}(r)] \mu(r) \mathbf{E}(r) \quad (3.15)$$

The quantity  $\varepsilon_0 \omega^2$  is similar to the total energy  $E$  in Schrödinger's equation.

For an electronic system, lowering the electron energy usually enhances the electron localization. For a photonic crystal, lowering the photon energy leads to a complete disappearance of the scattering mechanism itself, where at a high photon energy, geometric and ray optic theory becomes more valid and interference corrections to optical transport become less and less effective.

Equation (15) can be formulated in the form

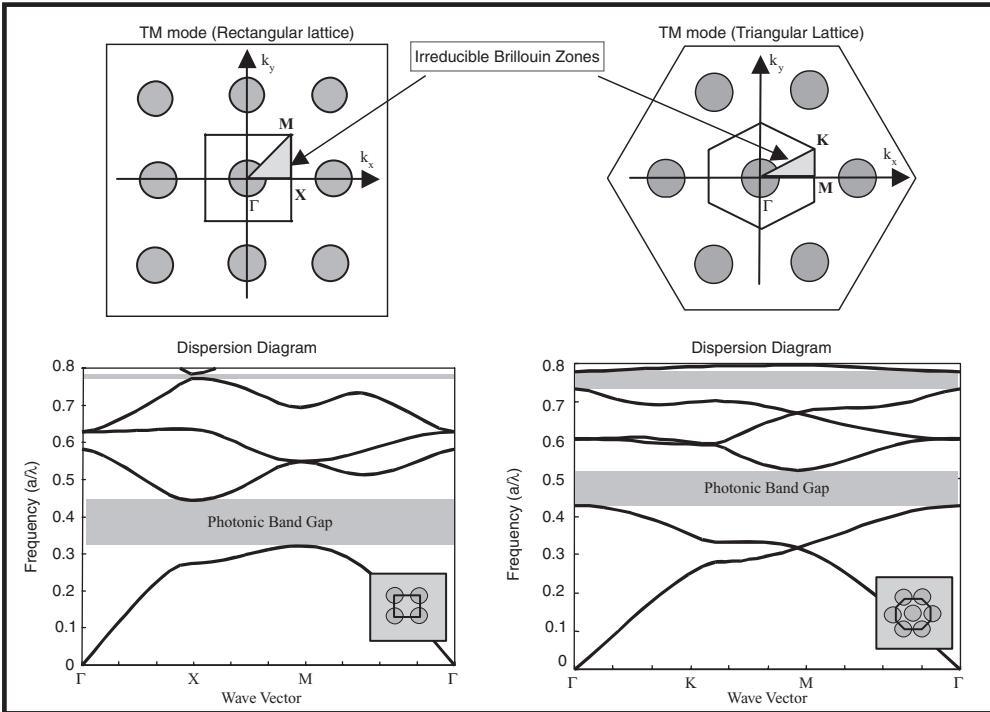
$$-\nabla^2 \mathbf{E}(r) + \nabla (\nabla \cdot \mathbf{E}(r)) - \omega^2 [\varepsilon(r) - 1] \mathbf{E}(r) = \omega^2 \mathbf{E}(r) \quad (3.16)$$

which is another form of the Schrödinger equation. By comparison, it can be seen that positive dielectric scatterers are analogous to regions of negative potential energy in a quantum system.

We can also see from Equation (3.16) that, because the increase in dielectric strength is analogous to an increase in the potential well depth of a quantum mechanical system, the overall effect is to lower the frequency of all modes of the system; hence the band edges will move downward in frequency with a general frequency dependence of  $1/\sqrt{\varepsilon_r}$ .

### 3.1.2 Analyzing Photonic Bandgap Structures

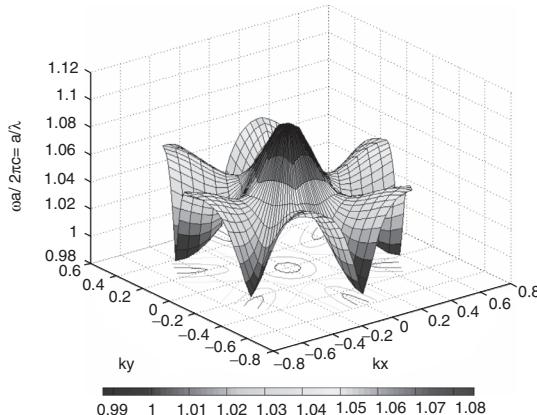
The method initially used for the theoretical analyses of PBG structures is the plane-wave expansion method (PWM), which makes use of an important principle: that normal modes in periodic structures can be expressed as a superposition of a set of plane waves, which is also known as Floquet's theorem [43]. The photon dispersion relations inside a photonic crystal (PhC) have been calculated using the plane-wave expansion method [44–47], where Equation (3.16) is solved as an eigenvalue problem with  $\mathbf{E}(\mathbf{r})$  as its eigenfunctions, and  $\omega^2$  is its eigenvalues. Solution over an irreducible Brillouin zone is plotted in a form of a dispersion



**Figure 3.2** Plane wave method (PWM) used to analyze different two-dimensional PBG structures of either rectangular or triangular lattice over an irreducible Brillouin zone. Result is a dispersion diagram, showing the possible eigenmodes for different wavevectors within the PBG lattice

diagram, as shown in Figure 3.2. A dispersion diagram is a two-dimensional plot of different eigenmodes for different wave vectors, or propagation angles, within a photonic crystal lattice. While a two-dimensional dispersion diagram is sufficient to show whether or not a bandgap may, or may not, exist for a certain PBG structure, it may not be sufficient for applications where nonlinear behavior of photonic crystals is being analyzed. An example is the negative refractive index phenomenon and its applications to the super-prismatic effect [48–50] in photonic crystals. For these applications a three-dimensional dispersion diagram, or a dispersion surface, will provide a more detailed view on a photonic crystal spatial response for various bands of frequencies both inside and outside the bandgap. Shown in Figure 3.3 is a plot of a dispersion surface of the first sub-band. Even though the PWM produces an accurate solution for the dispersion properties of a PhC structure, it is still limited because transmission spectra, field distribution, and back reflections cannot be easily extracted.

The plane-wave expansion method was limited to simulating infinitely periodic structures, which was constrained by multiple symmetries and assumed the structure to be lossless. It also assumed the structure to be perfectly periodic and, hence, fabrication tolerances, which highly modulate the spatial and temporal response of a periodic structure, could not be easily simulated in the PWM technique. In addition, it was not capable of calculating the transmission and/or reflection spectra. Nevertheless, plane wave expansion method remains to be a useful



**Figure 3.3** Dispersion surface. The horizontal plane gives both Bloch wave vectors  $k_x$  and  $k_y$ . The vertical axis gives the normalized frequency  $\omega a / 2\pi c = a/\lambda$

platform to quickly determine whether a periodic structure does or does not have a bandgap for a specific polarization, and remains to be the platform for extracting the highly complex dispersive properties of such periodic structures.

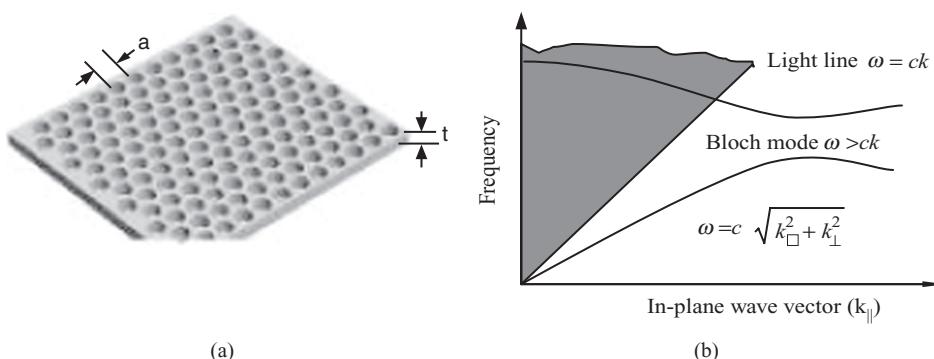
The FDTD method [51] is widely used to calculate transmission and reflection spectra for general computational electromagnetic problems, and it is generally considered to be one of the most applicable for the PhCs. In this case, a wave propagating through the PhC structure is found by a direct discretization of Maxwell's equations in point form, wherein the partial difference equations are discretized in both time and space on a staggered grid. In addition to discretization, the proper boundary conditions, i.e. absorbing and periodic boundary conditions, can be applied. If one defines the input signal as a continuous wave (CW) or pulse, the excitation can be propagated through the structure by time stepping through the entire grid repeatedly. Using this approach, several algorithms have been developed to calculate PhC band structures. Though powerful and flexible, the basic FDTD implementation on a single computer is extremely time consuming since its computational requirements grow exponentially with problem size. Various approaches have been taken to overcome this problem. The initial solution was to parallelize the FDTD algorithm over a Beowulf cluster of tens or hundreds of PC nodes [52]; this, however, provided only a short-term resolution to the ongoing performance issue. The mean time to failure of the number of nodes constituting the cluster grew exponentially and the maintenance cost and the physical space necessary to host such clusters soon became limiting factors. Recently, the industry has provided a more powerful and scalable alternative. By using commodity hardware devices, a single workstation can be created that outperforms clustered solutions while being cheaper to purchase and maintain. The two most popular alternative platforms for accelerated FDTD computations are graphics cards (GPUs) and field-programmable gate arrays (FPGAs). This approach relies on implementing the FDTD algorithm on a hardware accelerator-based workstation, [53–55] where dedicated commodity hardware is programmed to execute the FDTD algorithm at computational speeds up those seen by a 150 PC node cluster. FDTD solvers based on both GPUs [56] and FPGAs [57] are commercially available to provide the computational power required by current applications.

In a manner similar to the FDTD method, the transfer matrix method is implemented by discretizing Maxwell's equations. However, in this approach the initial excitation is limited to a monochromatic wave. The structure under consideration is divided into a set of layers with the same number of grid nodes in each layer. Then, using the discretized form of Maxwell's equations, the field  $\mathbf{E}_i$  in the nodes of one layer may be connected to the field  $\mathbf{E}_{i+1}$  in the nodes of the neighboring layers via the transfer matrix  $\mathbf{E}_{i+1} = \mathbf{T}_i \mathbf{E}_i$ . Thus, by integrating all layers, the output field is connected to the input field by the transfer matrix, which is a product of individual layer-to-layer transfer matrices. As in the case of the FDTD method, proper boundary conditions need to be used. While the transfer matrix method is less universal due to numerical instabilities during the integration, it is generally more computationally efficient than the FDTD method.

The finite element method (FEM) is a frequency domain method used to solve Maxwell's equations. In fact, it is also based on a variational principle, as in the case of the plane wave method. However, instead of using a plane wave expansion basis, which is defined over the extent of an entire unit cell, FEM uses a sub-domain basis to discretize within the computational unit cell. As such, FEM more efficiently takes into account material discontinuities in the dielectric structure, which helps overcome the slow convergence of the plane wave expansion method. To solve the resulting matrix eigenvalue problem, a preconditioned subspace iteration algorithm may be applied to find the most relevant set of eigenvalues within the large system of equations.

For silicon-based PhCs, the most promising class of PhC structures is the PhC slab, which has two-dimensional (in-plane) periodicity and a height that is comparable to the wavelength of light. The PhC slab is much easier to fabricate than a corresponding three-dimensional PhC structures and, consequently, more attractive for chip-level integration of different optical devices [23, 58, 59]. However, being finite in height requires another mechanism for light confinement in the third dimension, namely total internal reflection. Therefore, for these devices it is a combination of these two phenomena that serve to localize the in-plane light within the slab.

For these devices, the boundary between the guided and radiation modes is described as the light cone, where the radiation modes are the states that extended to infinity in the clad region outside the slab, and the guided modes are those localized to the plane of the slab, as shown in Figure 3.4(a). States that lie below the light line in the band diagram, Figure 3.4(b), cannot



**Figure 3.4** (a) PhC slab; (b) dispersion diagram for in-plane wave vector of a periodic structure overlapped with light line

couple with modes in the bulk background, shaded area in Figure 3.4(b). Such states are bound to the slab region by virtue of the multiple Bragg reflections within the slab plane or simply due to the existence of the photonic crystal structure. Thus, the discrete bands below the light cone are confined. Mathematically, we express the wave vector  $k$  as:  $\mathbf{k} = \mathbf{k}_{xy} + \mathbf{k}_z$ , where  $\mathbf{k}_{xy}$  is the in-plane wave vector, and  $\mathbf{k}_z$  is the out-plane wave vector. If the guided modes have an imaginary  $k_z$  component, then their modes decay in the cladding. However, if the radiation modes have a real  $k_z$  component, then they will leak to the cladding, or radiate to infinity.

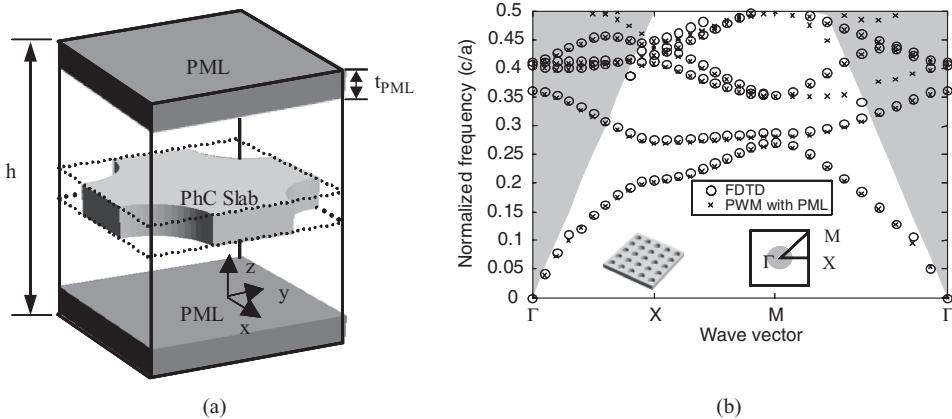
PhC slab structures have many potential applications and most of them rely on their corresponding band structures. To employ the plane wave method for the band structure calculations of PhC slab, which has only two-dimensional periodicity, a third dimensional periodicity was imposed by introducing a periodic sequence of slabs separated by a sufficient amount of background region to ensure electromagnetic isolation, which is commonly referred to as the supercell technique. In this way, the guided modes are localized within the slab so that the additional periodicity of periodic slabs having a large separation will not affect their eigenfrequencies. However, for the radiation (leaky) modes, which lie above the light cone, this technique is no longer appropriate due to the artificial periodicity in the out-of-plane direction.

To overcome this and determine the leaky modes above the light cone requires the application of a perfectly matched layer (PML) in the  $z$  direction to absorb the radiation from the slab. The PML absorbing boundary condition [60], which was first introduced by Berenger as a means to truncate the computational region in the FDTD method, has fast become one of the most proficient ways to absorb waves for any frequency and angle of incidence. This technique is based on an anisotropic material-based formulation that offers special advantages in that it does not require modification of Maxwell's equations [61]. As such, using PMLs the artificial periodicity in the  $z$  direction can be used without affecting the accuracy as any wave propagating out of the supercell will be absorbed and not give rise to an artificial resonance within it. Thus, combining the PMLs with the PWM to cast Maxwell's equations into a generalized complex eigenvalue problem [62] has been shown to be a very accurate tool for determining the band diagrams for PhC-slab structures.

While the introduction of the PMLs into the PWM does sufficiently suppress spurious modes, they do give rise to so-called PML modes, which are generated due to the periodic boundary conditions applied along the  $z$  direction. Therefore, an additional tool is required to distinguish the guide modes, leaky modes, and PML modes. Along these lines, two concepts can be used to distinguish those modes: one is based on the evaluation of the Q-factor of complex resonance modes and the other is based on the fact that guided modes are characterized by a relatively high power concentration within the PhC slab.

To see this, consider a square lattice with air holes embedded in slab with a high dielectric constant of 12.25 and a thickness of  $0.6a$ , as shown in Figure 3.5(a). The air holes are of a circular cross-section with radius  $r = 0.3a$ . As shown in Figure 3.5(b), there is a good agreement between the modified PWM with PML method and a 3D FDTD method. For a PhC slab, another possible technique is the effective index method, which can reduce a full 3D problem to a 2D one. For structures with a relatively low index contrast between the slab and cladding layers, the conventional effective index scheme offers a good approximation over a wide frequency range [63].

However, for those structures with a high index contrast, the effective index method is only valid within a very narrow frequency range since the effective index varies significantly over a



**Figure 3.5** (a) Unit cell for the band structure calculations; (b) comparison of the dispersion diagram between the FDTD and the proposed methods for even modes

typical frequency range of interest. For instance, for a perforated silicon slab with a permittivity of 12.25 and a slab thickness of  $0.6a$ , the effective index of the fundamental mode varies from 1 to 3.25 within the frequency range between 0 and  $0.5c/a$ . To obtain a good approximation of the effective index, a stepwise effective index method combined with the PWM has been proposed to determine the dispersion diagram [64].

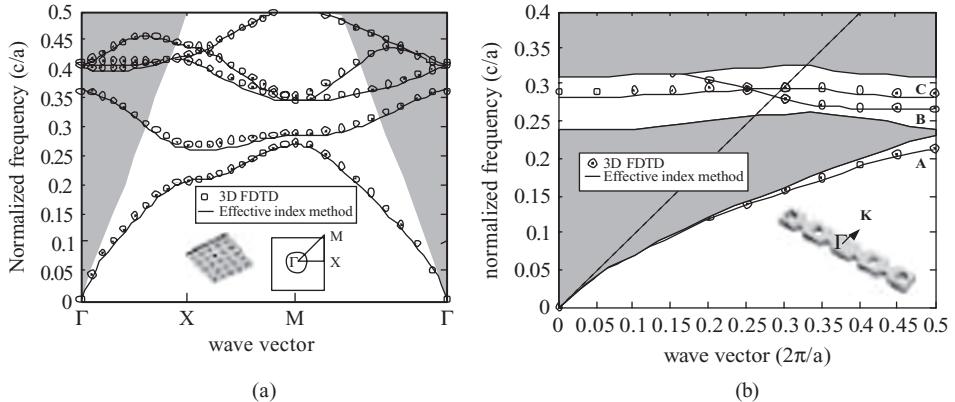
This revised PWM method starts with a 2D PhC, for which the problem can be decoupled into two sets of polarizations, namely, TE and TM. The eigenproblem can be obtained by following a similar procedure as in the conventional PWM derivation, which for the case of TE modes is [65]:

$$\frac{1}{k_0} \begin{bmatrix} -k_0[G_y] & k_0^2 - (k_x + [G_x])[\kappa]^{-1}(k_x + [G_x]) \\ k_0^2[\kappa] & -k_0[G_y] \end{bmatrix} \begin{bmatrix} [e_x] \\ [h_z] \end{bmatrix} = k_y \begin{bmatrix} [e_x] \\ [h_z] \end{bmatrix} \quad (3.17)$$

As the above equation shows, instead of solving for the eigenfrequencies, as done in the conventional PWM, the revised PWM solves for the wave vectors as eigenvalues for a given frequency.

Figure 3.6(a) shows agreement for the TE mode between the FDTD results and those obtained by the proposed method, within the normalized eigenfrequency range between 0 and 0.5.

In addition, a comparison of the computational cost between the two algorithms is worth noting: it took 128 min to obtain the eigenfrequencies for each given wave vector using the FDTD method while it took only 0.05 s to calculate the eigenwave vectors for each given frequency using the proposed algorithm, which is a significant reduction in computational cost. This method was also used to examine a PhC line defect waveguide, in which the waveguide is formed by removing a row of air holes along the  $\Gamma$ -K direction of the triangular lattice. The suspended slab had a thickness of  $0.6a$  and dielectric constant of 11.56. The radii of air holes were  $0.3a$  and the results are shown in Figure 3.6(b).

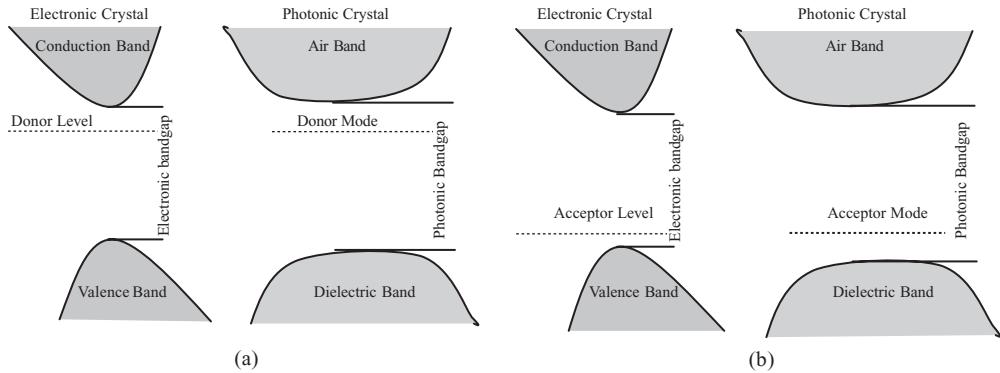


**Figure 3.6** Comparison of the dispersion diagrams for the TE modes of a PhC slab between the FDTD and the revised PWM: (a) square lattice; (b) triangle line defect

### 3.1.3 Doping of Photonic Crystals

Doping a semi-conductor material was achieved by either adding a donor or an acceptor atom. Both result in a changing of the electrical properties of an atomic crystal by either having a *p*-type or an *n*-type material. In a similar fashion, the optical properties of a photonic crystal can be changed by introducing point defects, which can be of two kinds, a donor type and an acceptor type. A donor type defect corresponds to increasing the effective refractive index of a local point defect, while an acceptor type defect corresponds to decreasing the effective refractive index of a local point defect, i.e. either adding or removing a certain amount of dielectric material. It was previously found that low-frequency modes concentrate their energy in the high-dielectric regions, and hence the bands below the photonic bandgap were called dielectric bands, while high-frequency modes concentrate their energy in the low-dielectric regions, and hence the bands above the photonic bandgap were called air bands [66].

When adding dielectric material to a unit cell it behaves like a donor atom in an atomic crystal, which corresponds to a donor mode and has its origin at the bottom of the air band of the photonic crystal. Alternatively removing dielectric material from a unit cell behaves like an acceptor atom in an atomic crystal, which corresponds to an acceptor mode and has its origin at the top of the dielectric band of the photonic crystal, which can be seen in Figure 3.7. In general, acceptor modes are preferable for making single mode laser nano-cavities, since they allow a single localized mode to oscillate in the cavity. By adding or removing a certain amount of dielectric material to the photonic crystal, we are disrupting the symmetry of the photonic lattice. By disrupting the symmetry we are allowing a single state or a multiple of closely separated states to exist within the bandgap, as shown in the transmission spectra in Figure 3.8 for both an undoped PhC sample as well as a doped sample. Transmission results were obtained using FDTD method with PML absorbing boundary conditions. This phenomenon of localizing states, by introducing point defects, can be useful in designing high Q-value nano-cavities in photonic crystals.

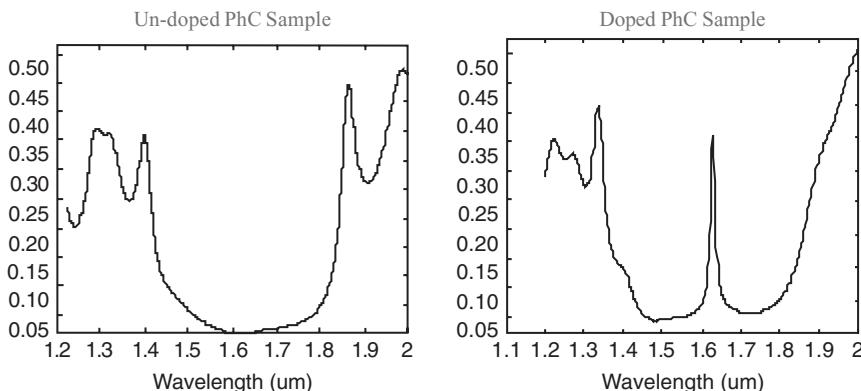


**Figure 3.7** Analogy between doping in electronic crystals (right column) and in photonic crystals (left column) for two cases: (a) donor doping; (b) acceptor doping

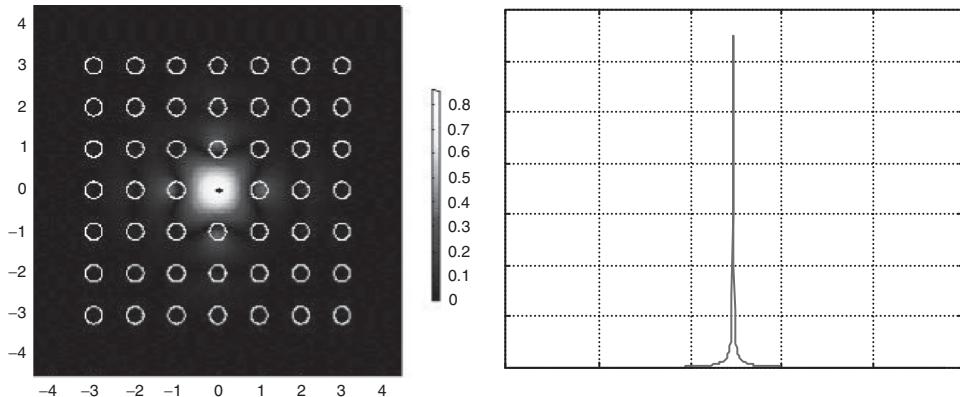
### 3.1.4 Waveguides and Nano-cavities in Photonic Crystals

A nano-cavity can be made by introducing a point defect in a photonic crystal. As such, the defect can have any shape, size or dielectric constant. By varying any of these parameters, the number of modes and the center frequency of the localized mode or modes inside the cavity can be changed. If we consider the case of a square lattice of cylindrical rods with a difference in dielectric constant much greater than 2, between the host material and the lattice material, we can introduce a point defect by simply changing one of the parameters of a given rod within the crystal. For example, a point defect consisting of a rod with a radius smaller than those surrounding it, will guarantee a single mode to be localized at the point defect. Alternatively, as we increase the radius of the defect, to be equal to or greater than those surrounding it, we will introduce a multiple of closely separated modes localized within the cavity.

To examine this we used the FDTD method to calculate the center frequency of a localized defect mode in a two-dimensional rectangular photonic crystal of circular dielectric rods in



**Figure 3.8** Transmission spectra for both an undoped photonic crystal sample as well as a doped photonic crystal sample

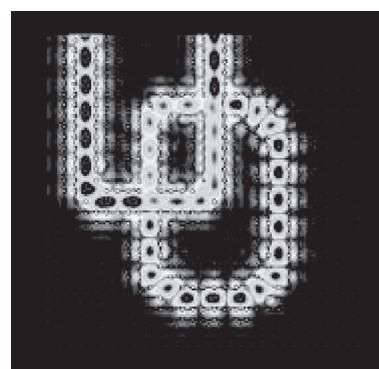


**Figure 3.9** A microcavity built in a PBG by removing a single rod from center of a rectangular unit lattice. FDTD simulation (left) and spectral result for the field inside the cavity (right)

an air background. The results are shown in Figure 3.9, where we placed a broadband point source within the defect region with a detector to record the time varying electric field which we then took the fast Fourier transform FFT for the recorded field at the detector as shown on the right, and on the left is the steady-state electric field distribution.

The quality factor of the nano-cavity plays a major rule in designing a high density WDM system. The quality factor depends mainly on the size of the crystal, as was previously shown [18]. For high Q values the size of the crystal surrounding the cavity needs to be large. It was also shown that the spectral widths of the defect modes decrease rapidly with an increasing number of lattice layers [19], which is more favorable in WDM since it maximizes the selectivity of the available bandwidth. In the next section we discuss the results of our computational analysis.

When line defects are introduced into a photonic crystal lattice, an electromagnetic wave having a frequency within the bandgap of the structure can be guided through the crystal. In this case the line defect resembles a waveguide, as shown in Figure 3.10. In this way line defects



**Figure 3.10** A photonic crystal waveguide created by introducing a line defect resembling the letters ‘UD’ was used to guide the light through the defect region

can be formed by either adding or removing dielectric material to a certain row or column along one of the directions of the photonic crystal. To this end, photonic crystal waveguides can be used as an optical wire to guide an optical signal between different points, or devices, within an optical integrated circuit or an optical network.

Note that for the case of a perforated dielectric slab, elimination of a single row or column will not be sufficient to have a single mode of propagation through the line defect and further design considerations must be taken to achieve that goal [67, 68]. By removing a column, or a row, we can confine the optical beam to the waveguide in a very similar fashion to the total internal reflection (TIR) concept, which is used to confine the optical signal in optical fibers. However, in photonic crystals the mechanism of in-plane optical confinement for a wave propagating through the defect is through multiple Bragg reflections, or distributed Bragg reflections (DBR). For finite height photonic crystal structures (slabs) vertical confinement is achieved through TIR at the interface between the PhC slab and lower dielectric constant material, e.g. air [69, 70].

The main idea of operation for this kind of waveguide is that an incident beam with a frequency within the bandgap of the structure will not propagate through the structure, but will propagate through the waveguide with minimal field leakage. Using this approach a throughput efficiency as high as 100% can be achieved through the waveguide [71–80].

Over the past 15 years, numerous modeling and simulation tools to design and analyze complex PhC structures were introduced to the commercial market. These tools reflect an ongoing effort within the research community to continue to develop and refine suitable and efficient algorithms and simulation tools for the future development of PhC devices and applications. With a range of such tools available, the community next moved towards realization of the various devices and application in terms of fabrication and experimental demonstration. Therefore, in the following section we present a detailed description of the efforts and techniques developed and optimized for this aspect of PhC development.

## 3.2 Fabrication and Characterization of Photonic Crystals

Modeling methods such as those discussed in the previous section had, by the mid-1990s, elucidated the fundamental properties of photonic crystals and validated, in principle, the operation of numerous useful devices based upon them. In addition, they established the analogy between the behavior of photons/electromagnetic waves in photonic crystals and that of electrons/deBroglie matter waves in crystalline materials, such as semiconductors. Whereas crystalline materials are formed in nature, and the technology to create highly purified forms of semiconductor crystals was well developed by the second half of the 20th century, naturally occurring materials which exhibit photonic bandgaps are exceedingly rare. Silica opals and biologically originated iridescent materials such as certain butterfly wings are examples, but they cannot easily be integrated into devices, and they lack the full bandgap required for many proposed devices to operate. Therefore, photonic crystals must be ‘synthesized’ using micro- or nano-scale assembly/machining technology. However, the practical realization of photonic crystals with visible or infrared bandgaps requires fabrication capabilities that lay at or beyond the limits of the tools available at the end of the last decade. The promise of photonic crystals was the motivation for a great deal of original work in micro- and nano-scale fabrication. The University of Delaware has been exploring this area for several years and has developed several

original methods. The following sections will describe these methods within the context of the work of the larger research community.

### 3.2.1 Planar/2D Photonic Crystal Fabrication

#### 3.2.1.1 Electron-beam Lithography

As explained previously, photonic crystals are periodic structures in which the periodicity is of the order of, and typically smaller than, the wavelength of light for which they are designed. A common embodiment is the planar slab photonic crystal, depicted in Figure 3.1. In case of high index materials such as silicon, which are necessary to take full advantage of the unique features of photonic crystals, including the photonic bandgap or unusual dispersion properties, the periodicity is typically about a third of the wavelength. This means that for photonic crystals operating in the near-infrared telecommunication band, around 1550 nm, the periodicity has to be about 500 nm, with the minimum feature sizes pushing the 100 nm mark. Patterning such small features in high-volume production lines is a challenge for commercial lithography systems. However, research tools are available that can tackle lithography at such small scales. Notable in this respect are electron-beam, or e-beam, lithography systems available from several vendors. The least expensive option for university research labs is a scanning electron microscope (SEM) converted to a lithography tool by fitting it with suitable electronics, sample stage, software, and a computer system capable of controlling the electron column and patterning the fine features of photonic crystals. For our research, we use Raith50, available from Raith USA, which is based on the XL30 FEI column with LaB<sub>6</sub> electron source. The system can routinely pattern 50 nm features, which is more than adequate for realizing photonic crystals in the near infrared region.

While capable of high resolution, e-beam lithography is not free of drawbacks. Chief among them is the relatively slow speed, which is the result of the serial nature of the writing process, where patterns are drawn sequentially. This is in contrast to photolithography employed in large-scale semiconductor manufacturing, where the entire chip is exposed all at once by optical projection of the image onto a die coated with photosensitive polymer (resist). As a result, e-beam lithography is typically used only for prototyping or in research environments where frequent changes in the design make the commercial approach of projection photolithography less practical and more expensive. Stemming from the serial nature of the e-beam lithography process is another problem associated with writing photonic crystals: the number of such small features that would have to be patterned for a device with respectable functionality. For example, patterning an array the size of an Athlon64 chip, 193 mm<sup>2</sup>, would require nearly a billion individual cells of a planar photonic crystal. As a result, vast amount of data would need to be stored, handled, and transferred from the computer system to the electron column for writing.

Furthermore, such a large area with small features cannot be exposed using just electron beam deflection due to unavoidable noise in the electric current of deflection coils and finite addressing accuracy of digital-to-analog converters employed to translate digital designs to analog deflection signals. For this reason, a single writing field is limited to few hundreds of micrometers where the deflection noise is small compared to photonic crystal feature size, and the addressing resolution of the pattern generator is sufficient for accurate feature placement. Larger fields are achieved by combining the electron beam deflection with the movement of the sample stage to create a patchwork of single exposure fields. For the photonic crystal to

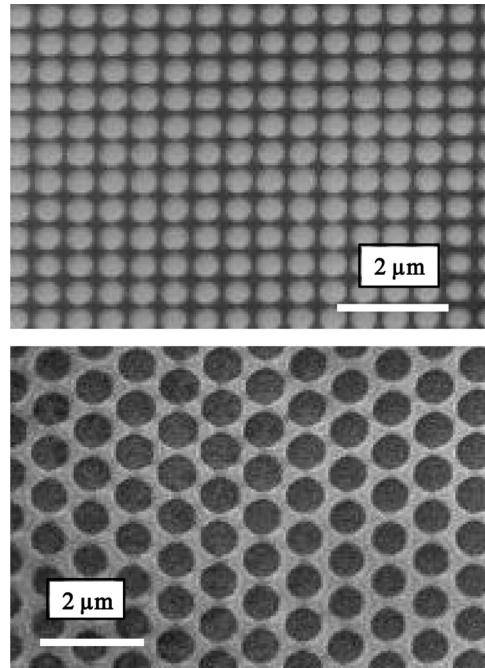
work as designed, it is necessary that the stitching accuracy between the exposed fields be considerably smaller than the period of the photonic crystal, otherwise phase errors may occur at the field boundaries and compromise the optical performance of the device. The stitching accuracy of state-of-the-art tools with interferometrically controlled stages is typically better 100 nm, and down to 50 nm, which is excellent for electronic applications where only wire contact is required, but insufficient for photonic crystal applications. As a result, research devices exposed using e-beam lithography are typically limited to a single write field of, at most, a few hundred micrometers and thus contain few functional structures.

To address these difficulties, we have been working on alternative methods to pattern photonic crystals, which rely on the fundamental nature of photonic crystals, i.e. their periodicity. They include interferometric, or holographic, lithography, and self-assembly. These methods will be described in subsequent sections.

### 3.2.1.2 Holographic Lithography/Combination Lithography

While e-beam lithography offers the best possible resolution in present-day lithographic methods and has the ability to define arbitrary structures, it is limited with regard to the fabrication of large-area planar photonic crystals. An alternative to e-beam lithography that can address this limitation is holographic/interferometric lithography. This method, first proposed for general use in photonic-crystal fabrication by Berger *et al.* [81], exploits the periodic nature of photonic crystals and the availability of coherent UV laser sources for exposing resists. By simultaneously exposing a layer of resist with several coherent laser beams, a periodic array of nodes and antinodes is formed. This interference pattern is a simple hologram, and if the orientations of the interfering beams are properly chosen, the symmetry of the interference pattern can be designed to match that of a desired photonic crystal lattice. Interference of two beams is widely used to pattern gratings, for example in the fabrication of fiber Bragg gratings. Planar photonic crystals require three or more beams to generate a 2D interference pattern. The spatial frequencies contained within the interference pattern are equal to the differences between the mutually interfering beams' wavevectors. Figure 3.11 shows examples of photonic crystal lattices patterned using this method.

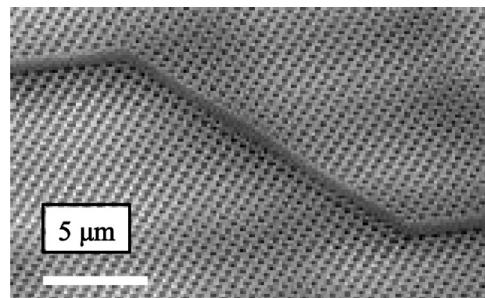
The great advantage of the holographic technique is its ability to expose large areas, limited only by the size of the laser beams (which can be expanded using lenses or mirrors), with nearly perfect periodicity and in a parallel fashion. With even modestly powerful lasers and commercial UV resists, photonic crystal lattices can be patterned in fractions of a second, over areas that would require many hours (or days) to pattern with e-beam lithography. The disadvantage comes from the inability to arbitrarily incorporate defects into the photonic crystal, which is key to the realization of many proposed devices to date. We have developed a hybrid fabrication methodology called combination lithography to exploit the best of both e-beam and holography and simultaneously address the limitations of each [82]. In combination lithography, we use a resist that is sensitive to both e-beam and UV exposure, using UV to expose the photonic crystal lattice holographically, and patterning defects such as waveguides with e-beam. The e-beam exposure is performed first, and renders the exposed resist insensitive to UV exposure. Then the sample is exposed over a wide area with interfering UV beams that define the photonic crystal. Subsequently, the resist thus patterned can be used as an etch mask to transfer the pattern into silicon by anisotropic dry etching (described in the following section). A sample fabricated using this process is shown in Figure 3.12.



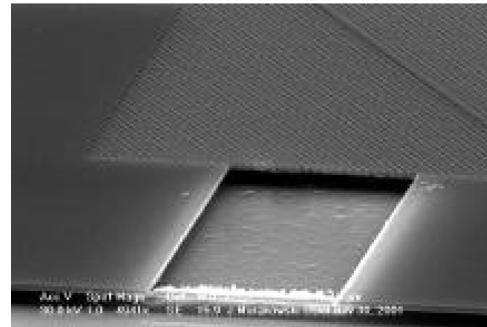
**Figure 3.11** Photonic crystal lattices patterned by holographic method. Top: exposure with interfering laser beams originating from the corners of a square yields a square lattice. Bottom: when the beams originate from the corners of an equilateral triangle, a triangular lattice results

### 3.2.1.3 Anisotropic Dry Etching for Planar Photonic crystals

The patterning of a photonic crystal is but a first step in their realization. To obtain functional devices, the pattern has to be transferred from the low-index polymer used in lithography to high-index material such as silicon. A conventional way of doing this for planar photonic crystals is by etching crystalline silicon using the patterned polymer as a mask. Planar photonic



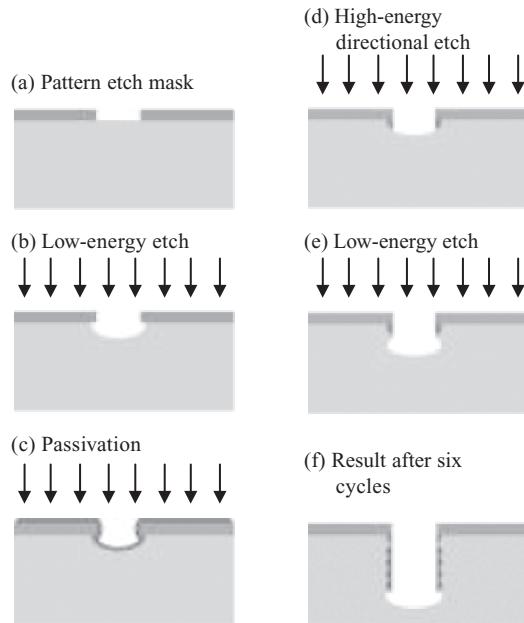
**Figure 3.12** Silicon photonic crystal lattice with defect waveguide. This sample was patterned with a combination of e-beam (waveguide) and UV holographic (PhC lattice) lithography



**Figure 3.13** Planar photonic crystal realized in SOI. The structure is suspended above the handle wafer by removing the underlying oxide layer with hydrofluoric acid

crystals consist of a thin layer of high-index material perforated with an array of holes as shown in Figure 3.13. The thickness of the silicon layer depends on the wavelength of light to be used with the device, with the goal of having a single-mode slab waveguide. For near infrared, this means less than 300 nm. Certainly, silicon layer this thin cannot be easily handled without a mechanical support of a thicker wafer. To provide this support, a few-hundred-micrometer-thick wafer is employed. For optical isolation of the device layer, either the handle wafer has to exhibit low refractive index or a spacer of lower index be provided. The latter strategy is used in the case of silicon-on-insulator (SOI) wafers, which consist of a several-hundred-micrometer-thick silicon handle, and a thin device layer separated from it by a layer of oxide. The advantage of such a structure for photonic applications is the possibility to remove the oxide layer from under the photonic crystal devices at the end of the process, as shown in Figure 3.13, to achieve symmetric structure and thus improved in-plane confinement. The former strategy is exemplified by silicon-on-sapphire (SOS), where a silicon device layer is grown on a sapphire wafer. With its refractive index lower than that of silicon, sapphire allows confinement of light to the device layer. Unfortunately, in this case, one ends up with an asymmetric structure prone to increased losses in photonic crystal devices, and no easy way to undercut the planar photonic crystals since sapphire is chemically inert.

In light of this, it is no surprise that most researchers working with planar silicon photonic crystals choose SOI as their platform. Such is also the case in our group where the majority of results have been obtained using SOI [83–90], although we also developed fabrication processes for and realized functional devices in SOS [91]. To achieve symmetric planar photonic crystal structures, it is necessary that the etch into the device layer yields vertical side walls. At the same time, in order to maintain high fidelity of pattern transfer from the resist to silicon, it is desirable that the resist itself is used as a ‘soft’ etch mask for etching silicon without intermediate hard mask material. To satisfy these goals, highly anisotropic (directional) and selective etch of silicon is required. Such an etch can be realized by alternately etching and passivating the patterned wafer [85]. The process flow is presented in Figure 3.14. After patterning the etch mask using, for example e-beam or holographic lithography, the sample is loaded into a reactive ion etching (RIE) reactor where a sequence of steps is performed.

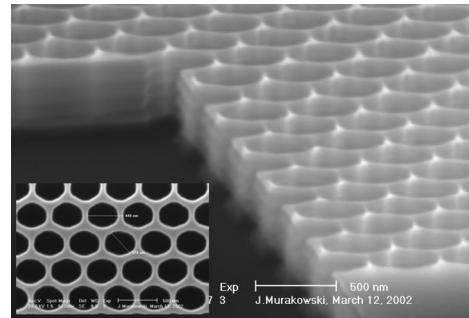


**Figure 3.14** Fabrication process for realizing vertical sidewalls with soft etch mask

First, low-energy isotropic etch in fluorine-based plasma chemistry is performed as shown in Figure 3.14(b). Fluorine reactive species selectively attack silicon, leaving the polymer mask layer intact, thus allowing high-selectivity etching in the absence of high-energy ion bombardment. Then, a passivation layer is deposited, Figure 3.14(c), which consists of Teflon-like material synthesized in the plasma discharge with octofluorocyclobutane ( $C_4F_8$ ) flowing into the reaction chamber. Following the passivation, high-energy ion bombardment is used to selectively remove the Teflon layer from horizontal surfaces. Due to high directionality of the ions accelerated toward the sample surface, the passivation on the side walls remains intact whereas the bottom of the etched trench is opened for subsequent steps. Low-energy etch similar to that of step (b) follows, as shown in Figure 3.14(e).

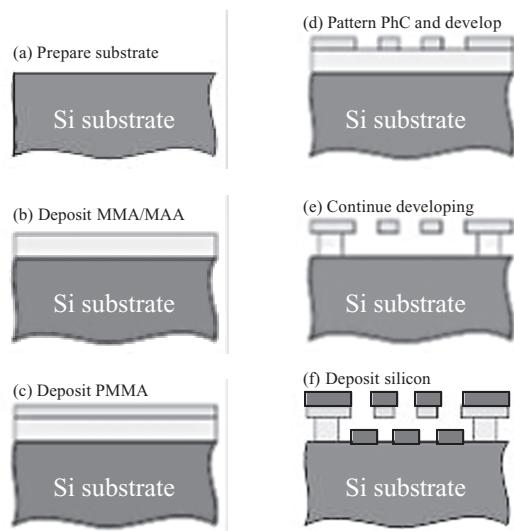
The sequence is repeated until the desired depth is reached. The result of the process is shown in Figure 3.15, which demonstrates the ability to achieve vertical side walls and sub-100 nm feature size, as required for high-performance photonic crystal devices. Side wall ripples, which are the result of the cyclic process, are about 80 nm wide and have little effect on the propagation of light with 1550 nm wavelength. On the other hand, as described in the next section, the ripples can be exaggerated by increasing the time of the low-energy isotropic etch in each cycle to efficiently realize three-dimensional photonic crystals in bulk silicon solely from a 2D etch mask.

The etching of devices in a wafer is an example of a top-down approach where material is selectively removed to obtain desired structures. An alternative relies on selective addition of material, referred to as a bottom-up approach, and can also be employed to realize

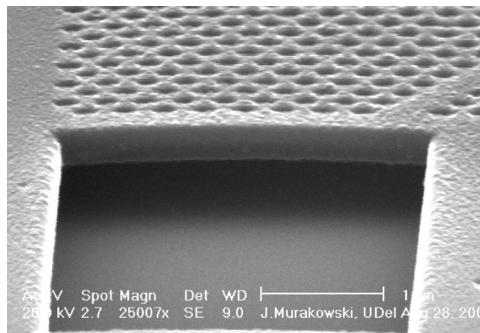


**Figure 3.15** Planar photonic crystal realized in SOI using cyclic process of alternate etching and passivation

photonic crystals, as reported in [92]. The process flow as shown in Figure 3.16 relies on difference in dissolution rates of polymethylmethacrylate (PMMA) polymer and methylmethacrylate/methacrylic acid (MMA/MAA) copolymer in methylisobutyl ketone (MIBK), which is used as a developer after e-beam lithography. Thus, if PMMA is deposited on top of MMA/MAA as shown in Figure 3.16(b) and (c), then exposed with e-beam, and developed in MIBK, the faster-dissolving layer of MMA/MAA will produce undercut, and suspended perforated PMMA membrane will be created, as in Figure 3.16(e). Since the refractive index of PMMA is low, such a perforated membrane does not constitute a useful photonic crystal device. To remedy that, silicon is sputtered over the entire structure. Where there is an opening in the membrane, the flux of silicon atoms passes through and settles on the substrate. Otherwise, it is deposited on the membrane. The result of this process, shown in Figure 3.17, demonstrates



**Figure 3.16** Etchless fabrication of planar photonic crystals in silicon



**Figure 3.17** Silicon planar photonic crystal created with the etchless process described in Figure 3.16

its viability for the realization of planar photonic crystals. It shows that the PMMA membrane is sufficiently robust to withstand the silicon sputtering process, and allow the planar silicon photonic crystal to form.

### 3.2.2 3D Photonic Crystal Fabrication

The challenge of fabricating three-dimensional photonic crystals in silicon has inspired a wide variety of solutions, but the majority fall into one of several broad categories. First is the most direct type of approach, the repeated application of mature planar lithographic methods in a layer-by-layer manner [93, 94]. Such approaches have the advantage of compatibility with existing CMOS microelectronics infrastructure and ease of optical-electronic integration. The disadvantage is the high cost of the processing equipment involved and the tediousness and difficulty of precisely aligning repeated deposition, lithography and etching steps with sub-10-nm accuracy. While modern microprocessors and memories often involve numerous layers for devices and interconnects, photonic crystal devices may require many more, particularly so-called ‘woodpile’ structures, in which four processed layers are required to realize each unit cell of the final crystal structure.

The second family of approach involves the self-assembly of uniform particle dispersions to form templates which are subsequently back-filled with silicon (or, more generally any other material as application requires) by, e.g. chemical vapor deposition (CVD). The self-assembly of uniform particles is a well-known and well-studied process, commonly known for resulting in the natural formation of opals. Such structures are termed ‘artificial opals’ when they are formed in a laboratory under controlled conditions. An example of how this approach can be used to realize silicon photonic crystals is provided by the work of Vlasov *et al.* [95], who used low-pressure CVD to back-fill a silica template formed by convective assembly. The advantages of approaches based on self-assembled templates are low cost and speed, as the particle suspensions can be made cheaply or obtained commercially and the self-assembly process is highly parallel, potentially yielding hundreds or more layers in a matter of hours. The great disadvantages are the randomness of the self-assembly, which leads to undesired defects, the limited types of lattices that can be obtained from close-packed arrays of spherical particles, and the lack of compatibility with standard CMOS platforms for opto-electronic

integration. Additionally, self-assembly does not readily allow for the controlled placement of defects in the PhC lattice. A notable effort to overcome this limitation is that of Lee *et al.* [96], whose method combines self-assembly with the third family of patterning discussed below, photopatterning.

The third family of approaches is similar to the second in that it involves the fabrication of a template structure, but rather than self-assembled particles, it uses volumetric patterning of photosensitive materials, i.e. resists. Within this family are two basic strategies, namely serial writing and parallel holographic fabrication. The holographic or interferometric approach, pioneered by Campbell *et al.* [97] is the 3D generalization of the 2D holographic method described in the previous section, wherein four or more coherent laser beams are combined within a transparent resist material to form a three-dimensional interference pattern. The serial strategy uses a tightly focused ultrafast infrared laser and multi-photon absorption to write a structure within the volume of a resist point-by-point. This method was first demonstrated for 3D photonic crystal fabrication by Sun *et al.* [98] and Cumpston *et al.* [99]. These two strategies are somewhat complementary in their advantages and disadvantages; specifically, the holographic approach is rapid, parallel, with near perfect lattices achievable over large volumes, but does not readily allow the incorporation of defects, while the serial approach is slower and subject to positioning errors, but can generate arbitrary defects with ease.

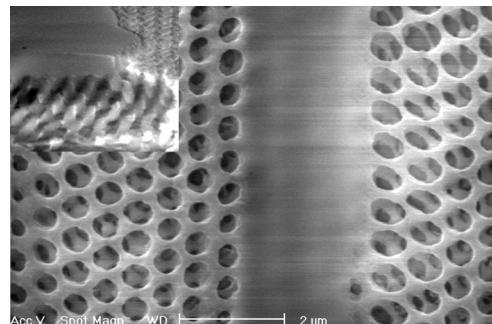
We have pursued several 3D fabrication methods falling into the categories above, with the particular method dictated by the application. These efforts have included self-assembly [100], layer-by-layer photolithography [101, 102], and holography [103, 104]. In this chapter, we limit our discussion to those methods most compatible with integrated opto-electronic platforms in silicon.

### 3.2.2.1 Holographic Lithography

Extending the combination lithography process described in the previous section to 3D is possible because SU-8 photoresist, which is chosen for use in 3D interference lithography for its thickness, transparency, and sensitivity, is also an excellent e-beam resist. We have developed a process based on SU-8 that is capable of embedding a wide variety of defects within 3D PhCs. A preliminary experimental result is pictured in Figure 3.18. This sample was obtained by exposing a straight waveguide with a low-kV e-beam exposure, such that the exposed region did not extend through the depth of the resist film. This is confirmed by the cross-sectional image in the inset. Note that the earlier patterning of the defect by e-beam lithography did not disrupt the formation of the 3D interference pattern, even in the region directly beneath the defect. We are working to embed such defects within the lattice by applying a second layer of photoresist on the top of the sample after e-beam patterning, but before the interferometric exposure.

### 3.2.2.2 Layer-by-layer Photolithography

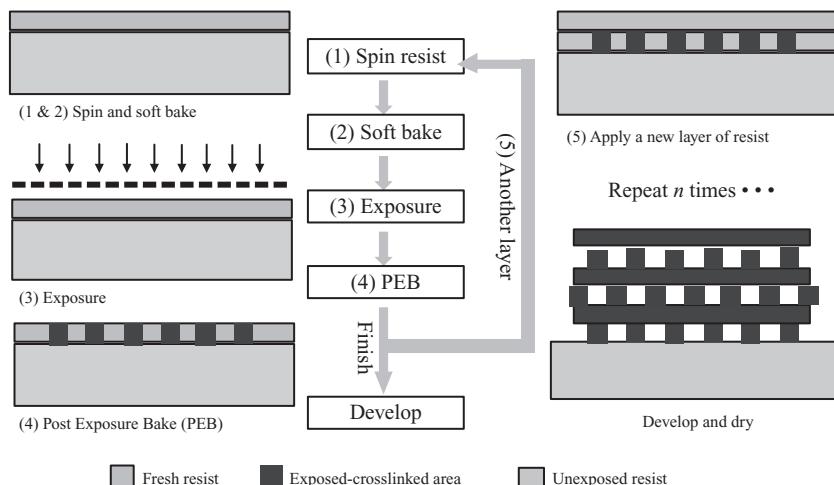
We have also developed a 3D layer-by-layer method to fabricate PhCs by photolithography. Like most layer-by-layer methods, this technology promises greater control and makes simpler the introduction of arbitrary defects into photonic crystals compared with holographic lithography, at the expense of throughput. Figure 3.19 depicts the process, which is comprised of



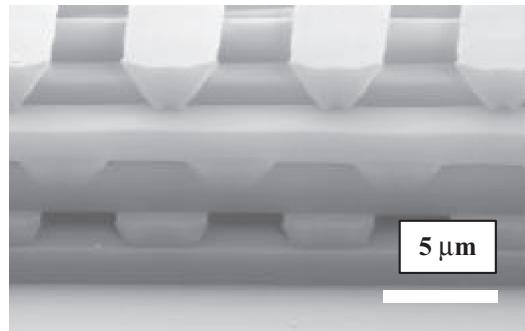
**Figure 3.18** Demonstration of 3D combination lithography. This sample consists of SU-8 photoresist that was exposed with a pattern of lines using e-beam lithography, and subsequently flood-exposed interferometrically, creating a 3D photonic crystal with an embedded defect. Inset: a cross-section showing that the depth of the e-beam-patterned defect is limited and that the formation of the PhC structure was not disrupted by the defect

many repetitions of a four-step process of: (1) applying resist; (2) soft bake to remove solvents; (3) patterning by UV exposure with a photomask; and (4) post-exposure bake to cross-link the exposed resist. After all repetitions, the final structure is developed all at once.

The key to the layer-by-layer photolithography process lies in the ability to confine the UV dose in each exposure step to the top layer of freshly applied photoresist. Without such confinement, subsequent exposures would lead to unintended re-patterning of the lower layers of resist. Two methods have been developed to accomplish this: desensitization of the resist



**Figure 3.19** Process steps for the layer-by-layer 3D photolithography process. The key to such a process lies in the method used to confine the UV dose of later exposures to the freshly applied resist layer, preventing undesired repatterning of lower layers



**Figure 3.20** A six-layer woodpile photonic crystal achieved using the layer-by-layer 3D photolithography process

by baking using a positive resist with image reversal [102], and intentional mismatch of the UV source and the negative photoresist [105]. The source–resist mismatch method offers a greatly simplified process compared to image reversal, and has yielded superior results. In this method, a negative photoresist, which is designed for use at 365 nm, is exposed by deep-UV radiation at 248 nm. At this shorter wavelength, the photoresist is strongly absorbing. For a properly calibrated combination of resist thickness and dose, this absorption just prevents the deep UV light from reaching layers beneath the top layer.

Figure 3.20 depicts a sample fabricated by the layer-by-layer photolithography with source–resist mismatch. This structure is called a woodpile photonic crystal, and is known to possess a complete photonic bandgap for silicon. The layer-by-layer nature of the method offers the ability to introduce defects arbitrarily simply by modifying the photomask for a given layer, or set of layers. In fact, completely arbitrary structures are possible; the structure merely needs to be ‘sliced’ into discrete layers, and a photomask designed to pattern each layer.

The flexibility of the layer-by-layer process, in which the introduction of defects into a periodic structure does not require modifying the process (only the photomasks), is gained at the expense of speed and convenience compared with interference lithography, which patterns the entire volume of the PhC simultaneously.

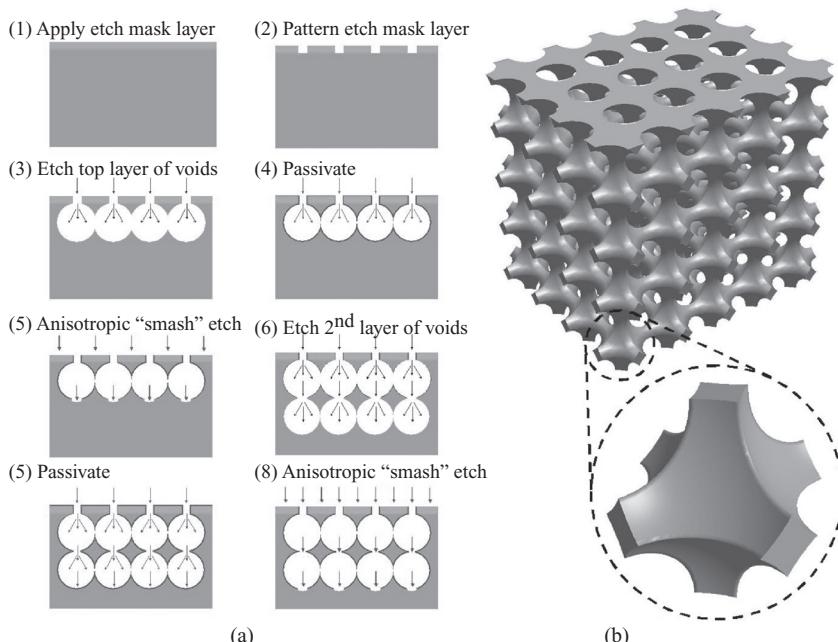
To realize silicon photonic crystals, templates fabricated using layer-by-layer photolithography or holographic lithography can be back-filled with silicon via CVD.

### 3.2.2.3 Time-modulated Dry Etching of Silicon

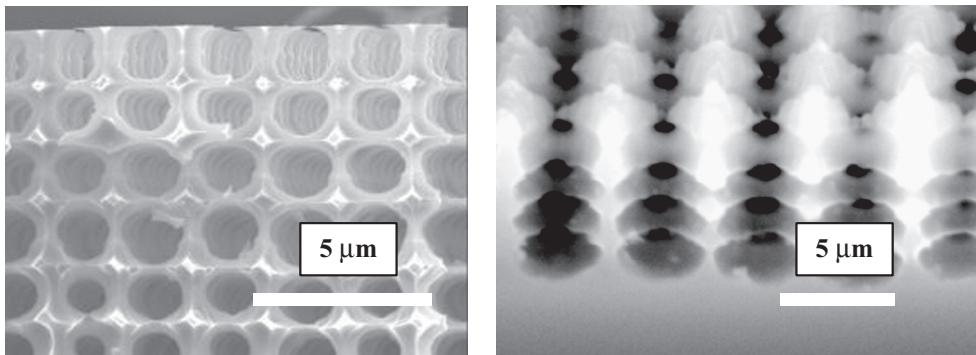
A novel process has been developed at the University of Delaware to realize 3D PhCs in silicon by processes compatible with standard CMOS processing, utilizing only planar lithography combined with a time-modulated deep reactive ion etch [106]. It possesses the advantage of a high degree of parallelism, alleviating the tedious repetition of deposition, polishing, alignment, lithography, and etching that is typically necessary in CMOS-compatible layer-by-layer approaches. It also avoids the need for back-filling, creating a buried PhC lattice directly in a silicon substrate. Limited types of defects can be introduced, either as local variations in the planar etch mask that are carried through the depth of the lattice, or as a horizontal plane by modifying the etch process for a given layer or set of layers.

The development of this fabrication process began with our work on planar photonic crystals, as follows. Optimization of the performance of finite-height 2D planar photonic crystal structures often requires high-fill-factor lattices ( $r/a > 0.4$ , where  $r$  is the hole radius and  $a$  the lattice constant). In near-IR slab photonic crystals in SOI, the lattice constant is approximately 450 nm. During our investigations, we fabricated several high-fill-factor slab PhCs and noticed that when the diameter of the holes in the mask is nearly equal to the lattice constant, the side wall ripples from the etch process started to overlap, leading to perforated side walls. These ripples can be controlled by adjusting the cycle time and extremity of etching during that cycle of the process. However, by looking at these ripples, it seems that it might be possible to exacerbate this effect to the point where they become nested spheres, and indeed, it is. We have shown that it is possible to realize a buried PhC lattice within a silicon substrate using only a planar etch mask. Assuming that the mask is patterned with a square lattice of holes, this would give rise to a 3D simple-cubic lattice. The etch mask can be patterned by any of several lithographic methods discussed earlier, such as e-beam or 2D holographic lithography. In addition, gray-level masks can be used to introduce vertical offsets between layers of spheres, thereby enabling the realization of lattices such as bcc and fcc.

The modification to the DRIE process that would be required to realize the proposed method of 3D PhC fabrication is depicted in Figure 3.21. The essence of the process is the alternation of nearly isotropic etches to create the spheres, with highly effective surface passivation and anisotropic etching to protect each layer of spheres from further etching and to allow the next



**Figure 3.21** (a) Process sequence schematic for a modified TM-DRIE process that would yield 3D simple cubic PhCs; (b) rendering of the ideal structure that would result from the etch process depicted in (a). *Inset* shows the contents of a single unit cell.



**Figure 3.22** SEM images of overlapping voids etched in silicon: (a) early sample created using a mask with a 3  $\mu\text{m}$  pitch; (b) recent sample scaled down to telecomm length scales: lattice pitch is 0.5  $\mu\text{m}$

isotropic etch, respectively. While the initial attempts used polymer passivation similar to the planar process described in the previous section, we have found that better results are obtained with the use of thermally grown oxide for passivation. Figure 3.22 contains SEM images of samples realized using this process possessing 4–6 photonic crystal unit cells along the vertical direction.

Samples thus fabricated have potential applications as a dense optical interconnect fabric. By integrating such a photonic crystal lattice in the handle of a CMOS wafer, it may be possible to realize on-chip optical interconnects by the self-collimation phenomenon. Research indicates that as few as six layers of a simple cubic photonic crystal in silicon may be capable of sustaining a self-collimated optical beam [107]. We have also obtained experimental evidence of self-collimation in 3D PhCs at microwave frequencies [108]; these results should readily scale to telecommunications wavelengths in silicon.

### 3.3 Overview of Photonic Bandgap Applications in Silicon Photonic Integrated Circuits

Due to their capability of manipulating photons in an easy-to-engineer manner, photonic crystals has been of great interest as a platform for the development of photonic integrated circuit. In 1994, Meade *et al.* [109] first proposed the concept of utilizing PhCs with PBG(s) as waveguides for optical interconnection by introducing a line defect (perturbing the crystal along one line). They also pointed out that this type of waveguide is useful for compact, low-loss bends because there is no limitation on the bend radius imposed by the radiation loss as occurred in dielectric waveguides. Mekis *et al.* [74] subsequently showed that highly efficient transmission of light can indeed be achieved thorough sharp corners in waveguides created by taking one row away in square arrays of infinitely long dielectric rods with a refractive index of 3.4, wherein 100% transmission through sharp bends was shown at several frequencies. Lin *et al.* [110] experimentally validated these results using alumina rods in the microwave region. Since light propagates in air or low-index material for waveguides made in arrays of dielectric rods, and there is no confinement along the vertical direction, leakage of waves would be expected for rods with finite length. This renders this type of waveguide impractical for most applications.

For this reason, El-Kady *et al.* [111] proposed using high-index slab structures periodically patterned with air cylinders and sandwiched by low-index material for light guiding. Since in this case light propagates in a material having higher index than the surrounding materials, TIR along the vertical direction resulting from this index difference, in combination with the in-plane distributed Bragg reflection (DBR) effect, results in full 3D light confinement possible. This prediction was experimentally validated by Lin *et al.* [112]. In their experiments, highly efficient transmission at a wavelength of around 1.5 mm in a high-index-contrast PhC slab waveguide was observed.

It should be noted that although it was proved possible to achieve high transmission efficiency in PhC slab waveguides, it was also found [111] that there exist significant losses due to the coupling of the guided mode with radiation modes and with the backward guided mode. To actually bring PhC slab waveguides into practical use, it is very important that propagation loss be eliminated or at least reduced to an acceptable level. Therefore, tremendous effort has been made to quantify [113] and minimize [80] the propagation loss in PhC slab waveguides. Smith *et al.* [80] fabricated, tested, and quantified a PhC slab waveguide consisting of three missing rows with a propagation loss of 20 dB/mm for the first time. Notomi *et al.* [114] subsequently brought this number to 6 dB/mm by squeezing the waveguide width to 0.7 times one missing row. Later, Ohtera *et al.* [115] experimentally demonstrated a new optical waveguide consisting of a Si/SiO<sub>2</sub> autocloned PhCs with modulated lattice structures. This waveguide was found to have a net absorption and scattering loss of around 4.2 dB/mm. Very recently, McNab *et al.* [116] reported a PhC slab waveguide with a measured propagation loss of as low as 2.4 dB/mm at the wavelength of 1.55 mm. If this is true, it would be sufficient to apply in millimeter-scale PICs.

On the other hand, if PhC slab waveguides are to take the role of metal wires in the contemporary EICs for future PICs, bends in waveguides will be required to achieve the necessary functionality. However, it was found that although highly efficient and compact bending can be achieved in principle in PhC waveguides without radiation loss if a full PBG exists, sharp bends restricted by the PhC lattices normally introduce significant backreflection and thus bending loss. For a device to be practical, the transmission loss through the bends will have to be minimized in a wide frequency band. Some early experiments [50] demonstrated that low-loss bending is possible to achieve in PhC slab waveguides. However, the bandwidth of high transmission was limited to a narrow range, which greatly restricts the capability of PhC sharp bends in practical applications. To overcome this limitation, significant efforts have been made to improve the performance of PhC bends. Among them, the most representative works are rearranging PhC lattices and changing the size of holes to control the positions of guided modes within PhC slab waveguides [117], adjusting missing holes at bend corners to form open cavity-resonant bends [118], smoothing bends and changing the local width of waveguides [118], reducing bend angles [119], and adding appropriate defects at bend corners [120].

Microcavities are also essential components for many important optical devices, such as lasers and filters. Since metal materials are extremely lossy in the optical region and can not be used as reflective walls for optical applications, researchers have long sought a material functioning at optical frequencies, like a metal does in the microwave region. PhC structures with full PBG(s) seem to be a perfect material candidate for building optical microcavities because they can be made highly reflective with low-loss materials. The concept of PhC microcavities was first proposed by Yablonovitch and Gmitter [67], where they pointed out that if a perfect 3D periodicity is broken by a local defect, local electromagnetic modes can

occur within its PBG. This makes it possible to build cavities capable of scaling from millimeter waves to ultraviolet wavelengths with high quality factor(s). This vision sparked tremendous interest in PhC microcavities. Many excellent results have been reported both theoretically and experimentally since then. Most works are concentrated on three aspects in planar structures. First, since most applications require specific working wavelength, tremendous effort has been put in to find easy and practical methods of tuning resonant wavelengths [12]. It was found that resonant wavelengths of PhC microcavities are extremely sensitive to fill factor, dielectric contrast, cavity size, properties of object(s) in the cavity, and slab thickness in the case of planar structures. Second, the mode volume and the quality factor [121] are two very important properties of microcavities. Because the high finesse demanded by lasers and filters [122] and the highly desired Purcell factor for controlling the spontaneous emission in the cavity [121] both require high quality factors and small mode volumes, properties of PhC microcavities have been extensively studied and many methods have been proposed to improve the quality factor and reduce the mode volume. Among them, the most representative ones are adjusting geometry symmetry [123], tailoring the defect geometry in Fourier space [124], engineering PhC walls to taper the guided modes into the PhC Bloch modes to decrease the radiation loss [125, 126], recycling the radiated photons [14, 125, 126], and varying the cavity to gently confine light [124, 127] in order to reduce radiation. Third, since the final goal of theoretical studies is for making functional devices, it is therefore necessary to actually fabricate these designed PhC microcavities to validate theoretical results in order to provide concrete information for fabricating real devices. In this regard, although theoretically ideal PhC microcavities are very hard to experimentally achieve, due to their high sensitivity on the cavity geometry, much progress has been made. For instance, Foresi *et al.* [128] demonstrated a PBG airbridge microcavity with a resonance at a wavelength of 1560 nm, a quality factor  $Q$  of 265, and a mode volume  $V$  of  $0.055 \mu\text{m}^3$ . Recently, Akahane *et al.* [127] successfully demonstrated a nanocavity with  $Q = 45\,000$  and  $V = 7.0 \times 10^{-14} \text{ cm}^3$  using a silicon-based PhC slab.

Along with the establishment of theory and fabrication of PhC microcavities, tremendous progress has been made on studies of PhC lasers. Lee *et al.* [129] reported the first working pulsed PhC cavity laser in 1999. Painter *et al.* [130] demonstrated the first single mode pulsed PhC cavity laser with the threshold optical pump power of 6.75 mW and the quality factor of 250 at a cooled substrate temperature of 143 K. They were then able to improve the quality factor to 500 and lower the threshold optical pump power to 1.5 mW at room temperature [131]. Park *et al.* [132] further lowered the threshold optical pump power to 0.3 mW and increased the quality factor to 1900. Ryu *et al.* [133] reported a photonic band-edge pulsed PhC laser with a very low threshold optical pump power of 35  $\mu\text{W}$  at 80 K. Painter *et al.* [134] demonstrated a  $10 \times 10$  array of optically pumped 2D PhC defect lasers with varying lattice parameters. By adjusting the PhC interhole spacing as well as the hole diameter they were able to tune the laser wavelength from 1500 to 1625 nm on a monolithic InP-InGaAsP wafer. A wavelength resolution of 10 nm from device to device was obtained, limited only by the lithography and etching tolerances of the fabrication method. In 2000, Hwang *et al.* [135] reported the first continuous PhC cavity laser with the threshold optical pump power of 9.2 mW. In the meantime, Zhou *et al.* [136] demonstrated the first electrically pumped PhC cavity laser with a threshold electrical pump current of 300  $\mu\text{A}$ .

With the success of PhC applications on waveguides, cavities, and lasers, researchers began to consider PICs using PhCs as platforms. Most passive and active devices required by PICs,

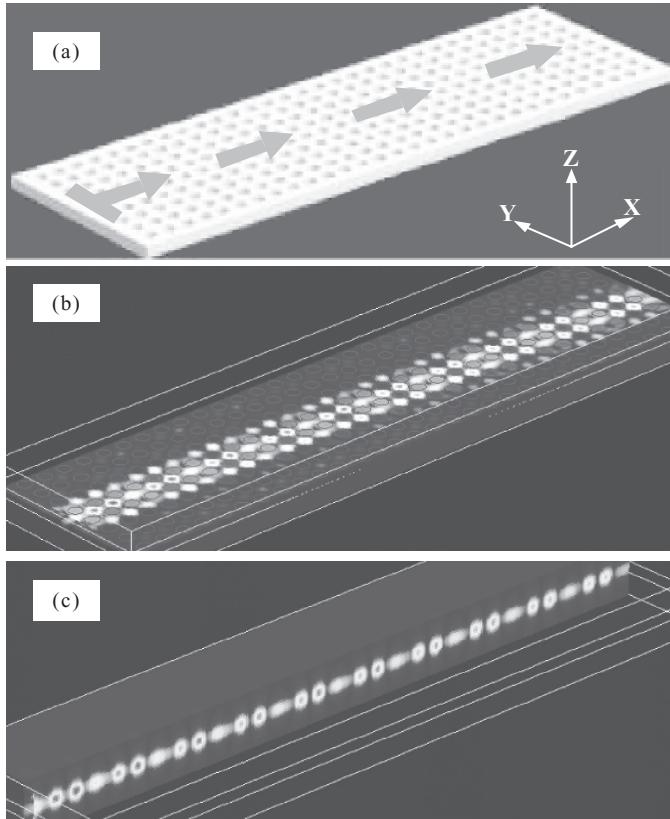
such as splitters [6], add/drop filters [137], switches [138], waveguide branches [8], transistors [139], limiters [140, 141], modulators [142–146], amplifiers [147, 148], and optical delay lines [149], have been proposed using PBG materials. Some of them have been experimentally demonstrated with very good performance. For instance, Noda *et al.* [150] demonstrated a drop filter, which drops photons passing by in a PhC waveguide through two PhC defect cavities with a  $Q$  value of around 400 at the wavelength of 1563 nm. Qiu *et al.* [137] fabricated an optical filter based on a contra-directional PhC waveguide coupler. This filter has a measured dropped bandwidth of 10 nm centered at 1585 nm, while the corresponding theoretical values are 6 and 1582 nm, respectively. Ohtera *et al.* [151] designed, fabricated, and tested a PhC polarization splitter. This splitter has a measured insertion loss of 0.4 dB and an extinction ratio of larger than 40 dB at a wavelength of 1.55  $\mu\text{m}$ . Lin *et al.* [152] first successfully demonstrated a Y-splitter with a splitting angle of 120 and a splitting loss of 0.5–1dB at wavelength 1640–1680 nm. In the meantime, Sugimoto *et al.* [153] experimentally observed strong optical beam propagation along a PhC Y-splitter. Wison *et al.* [154] subsequently demonstrated a PhC single-defect Y-splitter based a GaAs/AlGaAs heterostructure. This splitter was showed to have a measured transmission of around 40% for each arm relative to the comparable single-defect PhC waveguide. Very recently, ShiH *et al.* [155] fabricated and tested several PhC Mach–Zehnder interferometers, which can be used as filters, modulators, switchers, and multiplexers/demultiplexers.

### 3.4 Silicon Photonic Crystal Dispersion-based Applications

While the existence of PBG in PhC structures provides an excellent way to manipulate photons and thus allows for a variety of applications as overviewed above, another aspect of PhCs, namely the highly anisotropic dispersion property, is also very important despite being less well known. This property has been widely utilized to build highly dispersive prisms [156]. In 1999, Kosaka *et al.* [157] observed another outside-gap phenomenon exhibited by a 3D PhC showing collimated light propagation insensitive to the divergence of the incident beam. This was called the self-collimating phenomenon. Subsequently, Witzens *et al.* [158] theoretically investigated self-collimating phenomena in planar PhCs. Later, Wu *et al.* [159] experimentally demonstrated self-collimating phenomena in a planar PhC. They also combined the self-collimating phenomenon with the superprism phenomenon to make a beam deflection device. Later, Chigrin *et al.* [160] studied self-collimating in 2D PhCs. Inspired by these works, we have extensively investigated Silicon photonic crystal outside-gap applications.

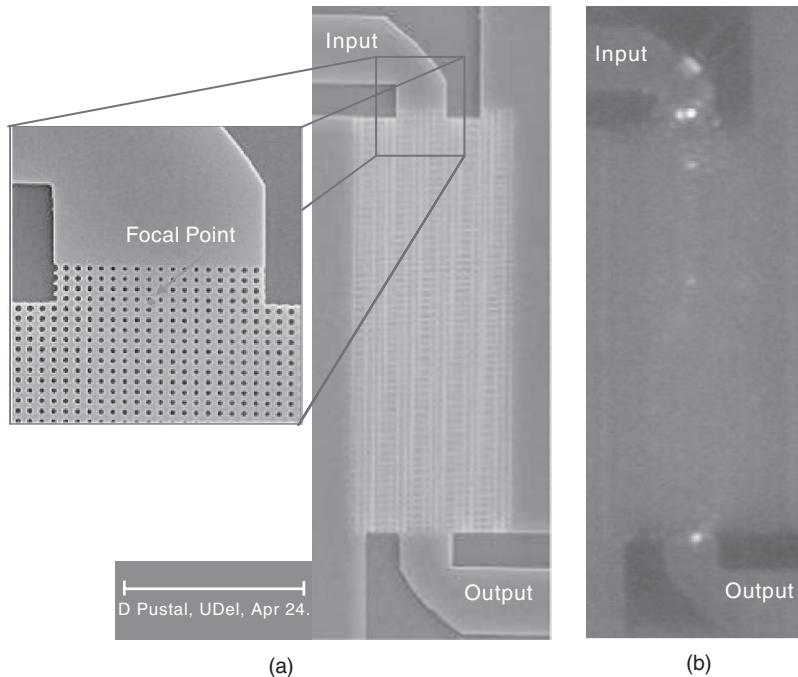
#### 3.4.1 Non-channel Waveguide

The first silicon photonic crystal outside-gap application we developed is a non-channel waveguide on planar photonic crystal with a square-like equi-frequency dispersion contour [161]. This type of waveguide has no physical boundary to confine light. Instead, it employs the self-collimation mechanism for lateral light confinement and the total internal reflection (TIR) mechanism for vertical light confinement. To show the self light guiding, we applied the 3D FDTD method to simulate the propagation of a Gaussian beam launched on the PhC from the unpatterned silicon, as shown in Figure 3.23(a). Figure 3.23(b) shows a horizontal cross-section of steady-state amplitudes of the  $H_z$  component along the central plane of the PhC slab,



**Figure 3.23** The 3D FDTD simulation of the propagation of a Gaussian beam launched on the PhC: from the unpatterned silicon: (a) the illustration of the structure and the source excitation; (b) a horizontal cross-section of steady-state amplitudes of  $H_z$  component on the central PhC plane; (c) a vertical cross-section of steady-state amplitudes of  $H_z$  component along the center of waveguide

which clearly indicates that the beam is well confined laterally within 3–4 lattice constants. Figure 3.23(c) shows a vertical cross-section of steady-state amplitudes of  $H_z$  components along the center of the waveguide. From this figure, one can see that the field is well confined in the slab vertically as well. The light propagation with good lateral and vertical confinement within this planar PhC was also experimentally validated. In his experiment, a tunable laser source was used to end-fire couple light into a wide input ridge waveguide, within which a J-coupler [162] was then fabricated to focus light into the PhC. As shown in Figure 3.24, the light enters the PhC lattice via the top J-coupler where some of the light is scattered by the silicon/PhC interface. Another scattered point of light at the opposite end of the PhC, which demonstrates the lateral confinement of the initially divergent Gaussian light. The conspicuous absence of a light trail in Figure 3.24(b) suggests low out-of-plane losses in this guiding structure. This is in contrast to conventional PhC line defect waveguides, which, for a wide frequency band, typically show a brightly lit path along the linear defect and thus out-of-plane radiation. From the above discussion, it is

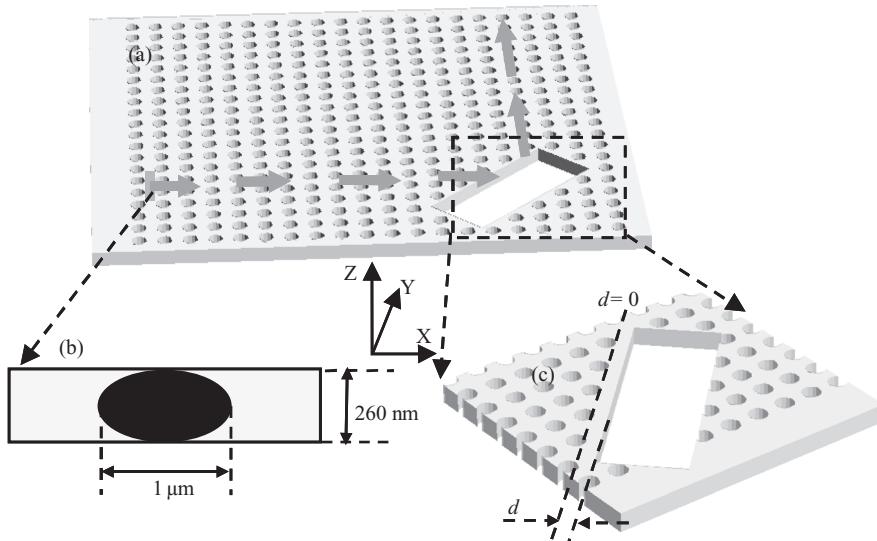


**Figure 3.24** Experimental characterization of PhC dispersion-based waveguide: (a) SEM image of the fabricated PhC dispersion-based waveguide; (b) image captured by a near-IR camera of the scattered light at the PhC/silicon boundaries. The point located at the output shows how the light is confined laterally within the PhC lattice

clear that such a structure can be used to efficiently guide electromagnetic waves within a planar PhC without the use of channel defects or structural waveguides. Since this type of waveguides does not require physical boundaries to confine light, it offers many advantages over dielectric waveguides or PhC line defect waveguides. For instance, it releases the strict alignment requirements imposed by the coupling efficiency in the case of dielectric or PhC line defect waveguides. As such, it enables a highly efficient in-plane coupling. Furthermore, the lack of waveguide structures allows multiple beams to be directed across one another in a very high-density fashion without imposing limitations due to structural interactions or crosstalk.

It should be noted that although the above mentioned non-channel waveguides allow highly-efficient coupling and arbitrary beam crossing, further implementation of this type of non-channel planar PhC waveguides, especially in high-density PICs, is greatly hindered due to their inability to efficiently bend and redirect light because no physically curved boundary can be implemented.

To this end, we also realized a device [163] to bend light in a non-channel planar PhC waveguides, which is illustrated in Figure 3.25. In doing so, we introduce a small rectangular air area that is etched away along the  $\Gamma$ -M direction, which behaves as  $45^\circ$  mirror to achieve  $90^\circ$  light bending. The 3D-FDTD method was used to characterize the bending efficiency of this mechanism. It was found that bending efficiency is very sensitive to the mirror location.



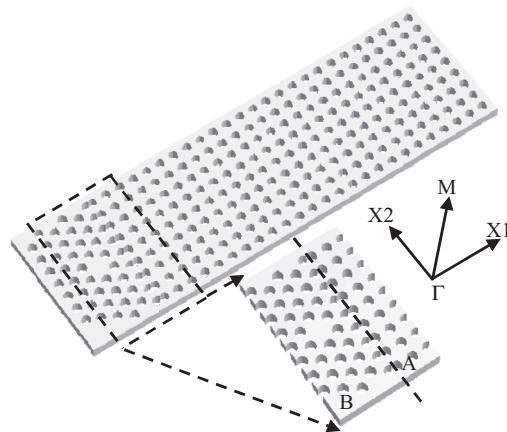
**Figure 3.25** The electromagnetic wave bending structure with a single 90° bend in a non-channel planar PhC waveguide: (a) the entire device; (b) the launched Gaussian beam seen from the propagation direction; (c) the zoom-in display of the bending part of the device.  $d$  is defined as the distance from the nearest edge of air holes to the mirror edge

Since the radius of air holes is  $0.6a$ , thus, when  $-0.6a < d < 0$ , the etched part overlaps with air holes and the mirror edge is corrugated. These results had been validated experimentally by fabricating the bending device in the 260 nm thick device layer of a silicon-on-insulator substrate.

### 3.4.2 A Unidirectional Emitter

We have also proposed a mechanism [164] that extracts spontaneous emission from the in-plane direction instead of the vertical direction. This simplified spontaneous emission model is illustrated in Figure 3.26. The model combines two different types of PhC structures, namely structure A and structure B. Structure A is designed to be a self-collimating structure. Structure B is a regular PhC defect cavity, which is tailored to control the spontaneous emission. While the lattice of the PhC cavity contains enough layers to ensure high reflectivity, the layers on the side of the self-collimating structure has been reduced to only two layers. In this way, the emission in the PhC cavity is prone to couple through the thin PhC lattice into the collimation structure A and be self-collimated as a narrow beam. In our simplified model, the spontaneous emission is taken at 1.5  $\mu\text{m}$ .

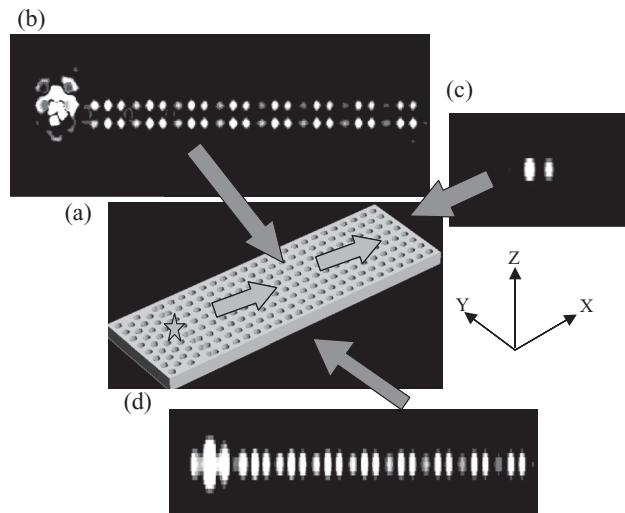
With the self-collimating structure, we desire to extract the spontaneous emission within the PhC cavity into a narrow and highly directional self-collimation beam. To enhance the coupling of photons to the self-collimating structure, we oriented the cavity structure with respect to the self-collimating structure in a way that the air holes in both structures aligned with each other. We then took this structure as our initial structure and varied the relative position of the



**Figure 3.26** A unidirectional PhC dispersion-based Emitter. The emitter consists of two structures, structure A and structure B. Structure B is the micro-cavity part which is created by taking one air hole away in a triangular array of air holes in the slab

two structures and the radius of the air holes comprising the structure to reduce the vertical loss and the leaky loss along the interface of the two structures. The purpose of doing this is to increase the power flux coupled into the self-collimation beam at the wavelength of  $1.5 \mu\text{m}$ .

The 3D FDTD method is used in the above optimization. Figure 3.27 shows the resulting steady-state distribution of the  $H_z$  component. By calculating the vertical power flow  $w_1$  on the



**Figure 3.27** The steady-state distribution of  $H_z$  component: (a) schematic of light propagation in the structure; (b)  $x-y$  horizontal cross-section along the central slab plane; (c)  $y-z$  end-on cross-section view; (d)  $x-z$  vertical cross-section passing through the center of the PhC cavity

top of the cavity and the interface, the leaky power flow  $w_2$  along the interface, and the power flow  $w_3$  in the self-collimation beam using the field distributions of the six electromagnetic components obtained with the 3D FDTD method, it is found that  $w_3$  is 2.7 times  $w_1$  and 28 times  $w_2$  if a detector with an area of  $1.75a$  by the slab thickness is used.

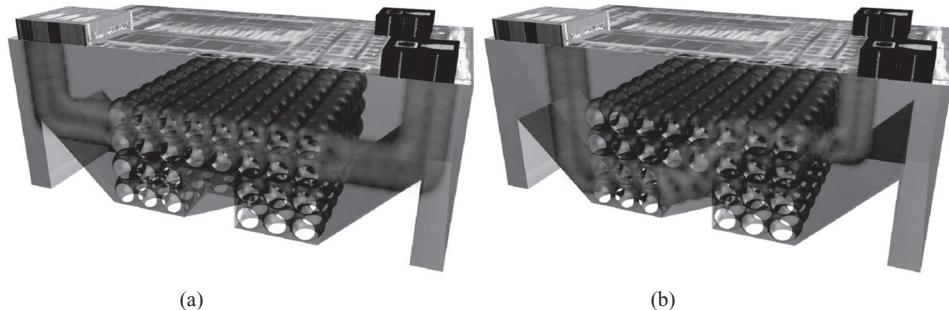
### 3.4.3 Sub-surface Silicon Optical Bus Based on Three-dimensional Photonic Crystals

The 2003 International Technology Roadmap for Semiconductors highlights key interconnect challenges for next-generation microprocessors and computing systems. The roadmap suggests that the most difficult challenges in the near term include the rapid introduction of interconnect processes compatible with device roadmaps, coupled with fine dimensional control and providing good mechanical stability and thermal budget. The continued push towards finer geometries, higher frequencies and larger chip sizes increasingly exposes the disparity between interconnect needs and projected interconnect performance. Further, the interconnect technologies should be able to meet performance requirements and manufacturing targets by leveraging low-cost conventional mass fabrication techniques and provide solutions to address global wiring scaling issues. The economies of scaling achieved by extending Moore's Law have lead to the dominance of silicon in the microelectronics industry; conversely only modest progress in silicon-based optoelectronic circuits has been achieved in recent decades. However, recent progress towards a silicon light source by erbium doping, nanostructures that enhance quantum effects, novel approaches such as dislocation engineering, optical amplification through Raman effect and free-carrier dispersion based optical modulation, coupled with need for alternate mechanisms to solve the interconnect bottleneck, has renewed the interest in silicon microphotonics. One of the other major obstacles to the realization of silicon microphotonics is chip-scale optical interconnects due to incompatibility of optical device materials and disparate integration scales with electronic devices and ICs. For this reason, we propose a buried silicon optical interconnect technology [89], the *sub-surface silicon optical bus* (S3B) with the ability to meet the challenges cited earlier by the semiconductor roadmap, specifically process compatibility and mass fabrication. Our approach towards this demonstration is by engineering the dispersion properties of embedded silicon three-dimensional photonic crystals to create sub-micrometer routing channels and control light using the technique discussed above. The idea is shown in Figure 3.28. The conceptual rendering in Figure 3.29 depicts a buried photonic-crystal lattice used to route optical signals beneath surface-patterned CMOS electronics.

In this way, CMOS circuits could be fabricated on the topside, sources and emitters can be flip-chip bonded directly on top to emit down through the silicon layer, and optical signals would be subsequently confined in the buried PhC lattice. This would enable the realization of a complete monolithic optical interconnect medium and would only require the flip-chip integration of the active devices, shown (not to scale) atop the structures in Figure 3.28.

To achieve this, we investigate self-collimation in a 3D simple cubic silicon PhC structure, as shown in Figure 3.29(a).

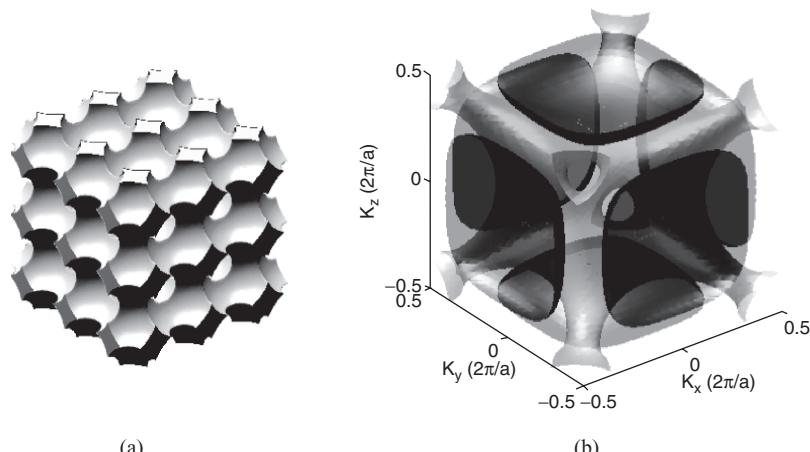
The dispersion equi-frequency surfaces (EFS) shown in Figure 3.29(b) of this photonic crystal structure is obtained by casting Maxwell's equations into an eigenvalue problem using Bloch wave boundary conditions, which is then solved using the plane wave expansion method



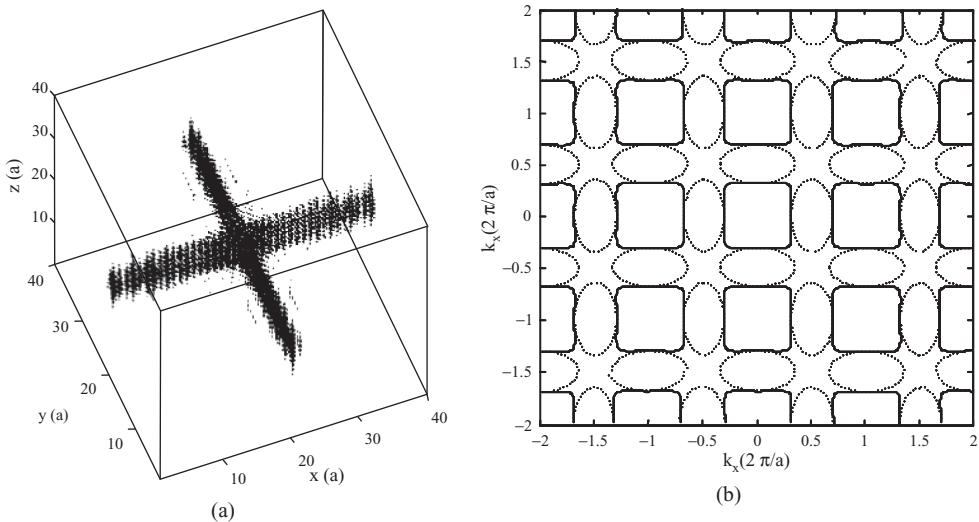
**Figure 3.28** Conceptual rendering of a monolithic integrated optoelectronic device containing a PhC-based sub-surface silicon bus (S3B) for optical routing. The optical bus is comprised of a buried three-dimensional photonic crystal, embedded beneath the electronic device layer. (a) and (b) show different optical paths followed by beams launched into different layers of the S3B, by a source with multiple emitters. The beams are coupled into the bus by angled mirror facets, and the lower beam in (b) is redirected by vertical 45° mirrors

(PWEM) [165]. The solid contour is the third band and dash contour is the fourth band. It is clearly shown that the EFS of the band 3 at frequency  $f = 0.34c/a$  has a cubic shape, which can be used to guide light, while the fourth band doesn't have such a capability.

To illustrate the self-collimation behavior, we simulated a dipole source located in the center of this simple cubic structure in silicon using **3D FDTD method**. Figure 3.30(a) shows the intensity distribution of the total  $H$  field  $H_{\text{total}} = \sqrt{H_x^2 + H_y^2 + H_z^2}$ . In addition, we also calculate the spatial spectrum of the propagation of a dipole within a simple cubic PhC structures. To do this, we took 3D spatial Fourier transform of the  $H_z$  field component. The spatial spectrum at



**Figure 3.29** (a) Simple cubic silicon photonic crystals with close packing spheres; (b) band diagram along the edge of Brillouin zone. The insets show the schematic view of the high-symmetry points in the first irreducible Brillouin zone.

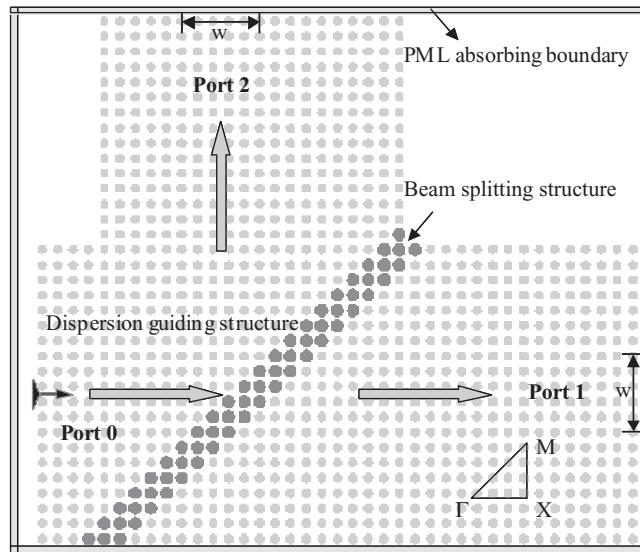


**Figure 3.30** (a) 3D-FDTD simulation result of a magnetic dipole source located in the center of the simple cubic structures. The figure shows the intensity of  $H$  field; (b) spatial spectrum of  $H$  field and comparison with dispersion diagram by PWEM

wavevector of  $k_z = 0$  is plotted in Figure 3.30(b), and compared with the dispersion diagram we obtained by using PWEM, which clearly shows that spatial spectrum of field matches very well with our expectations. Therefore, magnetic current source dominantly excites the third band and leads to the self-collimation phenomenon.

### 3.4.4 Dispersion-based Variable Beam Splitter

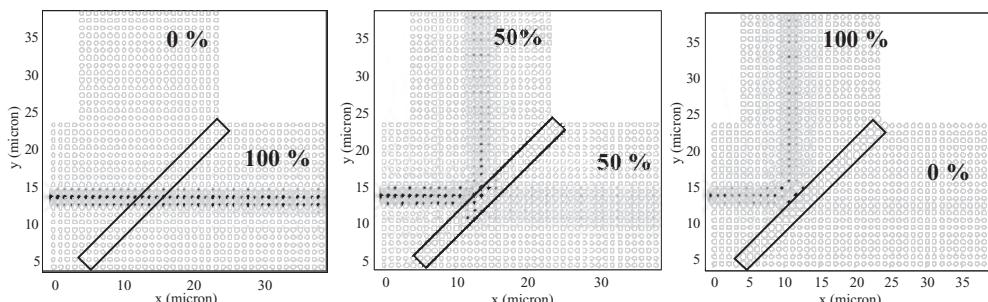
If the dielectric mirror in Figure 3.25 was replaced with a photonic crystal structure, which may take the same geometry as the host photonic crystal structure or may be different. In either case, controlling the parameters of such structure can be used to control the percentage transmitted through such structure. By doing so a hybrid photonic crystal structure, combining dispersion-based and a confinement-based structure as shown in Figure 3.31. Such a structure was used as a variable beam splitter. To this end, the optical beam splitter consists of two sections, as shown in Figure 3.31: a dispersion guiding PhC structure, and a beam splitting structure. These two kinds of PhC structures with different radii are both arranged on the same square lattice in a high-index background of  $n = 3.5$ . The PhC guiding structure has an air hole with a ratio  $r_g = 0.3a$ , where  $a$  is the lattice constant. The  $45^\circ$  rotated splitting structure has an air hole of radius  $r_s$ , which varies from  $0.3a$  to  $0.435a$ . If the light is launched from port 0, then it will propagate through the dispersion guiding structure and arrive at the splitting structure. While passing through the splitting structure, the signal will be split between two orthogonal branches: one along the same direction as the incident wave, and exit at port 1, and the other portion along a direction orthogonal to the incident wave, and exit at port 2. The splitting percentage is proportional to the radius and number of layers of the splitting structure. In this



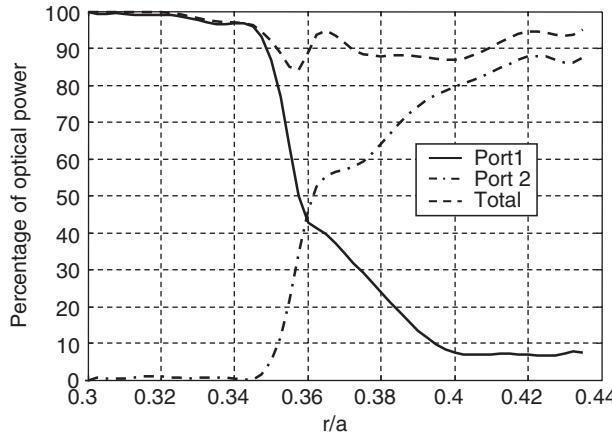
**Figure 3.31** Dispersion based tunable beam splitter utilizing a hybrid PhC structure

section, we study the effect of varying the radii of the splitting section to achieve arbitrary power ratio between output ports. If the radius of air holes in the splitting region is the same as the one in the dispersion guiding structure, it is clear that the wave will completely output at port 1. This output energy, in fact, is used as reference value for the calculation of splitting percentages. Figure 3.32(c) shows the steady-state result of magnetic field (Hz) for the radius of  $0.42a$  in the splitting structure, most of the energy exits at port 2. At  $r_s = 0.36a$ , one obtains approximately 3 dB splitting, and the corresponding steady state result of magnetic field is shown in Figure 3.32(b).

As the radius continually varies from  $0.3a$  to  $0.435a$ , the percentage of optical power versus the normalized radius is plotted in Figure 3.33. At radius of  $0.36a$ , the output ports have nearly the same power. Within the radius range between  $0.3a$  and  $0.345a$ , there is no stop band. As



**Figure 3.32** Steady-state result of  $H_z$  field with radii of air holes of: (a)  $0.3a$ ; (b)  $0.36a$ ; (c)  $0.42a$  for the splitting structure



**Figure 3.33** Percentages of output optical power of port 1 and port 2 vary with the radius of air holes of splitting structure

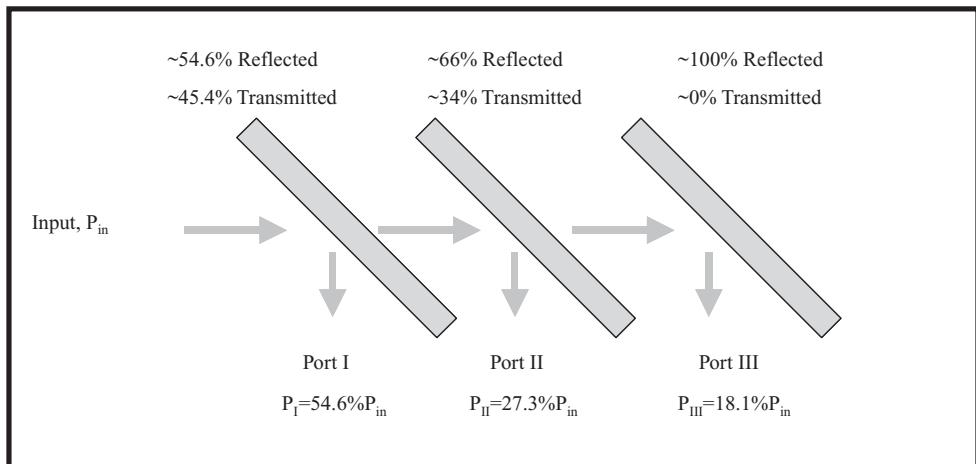
such, the dispersion contour at frequency  $0.26c/a$  is still square-like shape; hence, the light will propagate in the same direction as the incident wave and a small amount light will propagate to port 2 due to the interface of two slightly different types of PhC structures. As the radius of air holes continues to increase, the frequency  $0.26c/a$  falls into the bandgap. However, as we mentioned above, since the frequency is close to edge of the bandgap, a few PhC layers are not sufficient to completely reflect the light, and a portion of the signal will output to port 1.

The beam splitting device shown in Figure 3.31 was experimentally validated by fabricating the respective structure in a 260-nm-thick device layer of a silicon-on-insulator substrate. The fabricated device consisted of air holes with a lattice constant  $a = 442$  nm and radii  $r_g = 0.26a$  and  $r_s = 0.35a$  for the guiding and splitting regions respectively [166, 167].

### 3.4.5 Photonic Crystal Analog-to-digital Converter

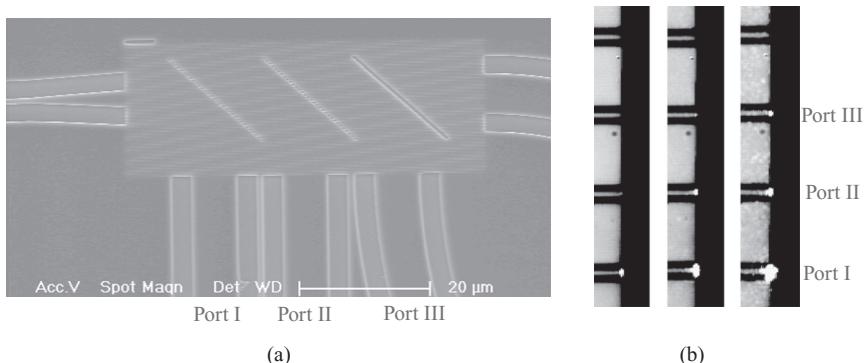
Recently, we demonstrated a 2-bit analog-to-digital (A/D) converter, as shown in Figure 3.34. Four distinguishable states were achieved using three serial photonic crystal splitters with splitting ratios of 54.5:45.4, 66:34, and 100: 0, respectively. The concept of the 2-bit optical A/D converter is illustrated in Figure 3.34.

The input analog signal is assumed to be a sinusoidal wave. When the power of the incident wave is lower than twice the threshold value, namely  $P_{in} < 2P_{th}$ , the output powers at three output ports are smaller than the threshold level and the quantized states of three output ports is hence defined as ‘000’. When the input power is increased to  $2P_{th}$ , the output power at the port 1 reaches threshold. Therefore, the states at the three output ports in this case can be similarly recognized as ‘100’. Accordingly, the output power at port II and III reaches threshold value and their states change from ‘0’ to ‘1’ subsequently when the power of the input signal continues to increase. Consequently, we have four different states, ‘000’, ‘100’, ‘110’, and ‘111’ at different levels of incident power. These four states can be coded into the desired four states ‘00’, ‘01’, ‘10’, and ‘11’ for a 2-bit optical A/D converter.



**Figure 3.34** Schematic of the two-bit A/D converter

To experimentally validate the design concepts of the 2-bit optical A/D converter, we fabricated the device on a silicon-on-insulation (SOI) wafer, which has 260-nm-thick silicon device layer on a 1-μm-thick  $\text{SiO}_2$  insulating layer. E-beam lithography and inductively coupled plasma (ICP) dry etching was employed to pattern and transfer the structure to the silicon device layer. The underlying  $\text{SiO}_2$  layer was removed using buffered oxide etching (BOE). Figure 3.35(a) shows the SEM micrograph of the fabricated device. Figure 3.35(b) shows the captured IR images of output ports at  $\lambda = 1560 \text{ nm}$  at different incident powers. From the figure, one can see that the output powers at port I, II, and III reach threshold level at different incident powers and input signals at different levels are thus be converted into four distinct states of the 2-bit A/D converter.



**Figure 3.35** Fabrication and characterization of a 2-bit A/D converter: (a) the SEM picture of the device; (b) the captured IR images of output ports

### 3.5 Conclusions

It is clear that photonic crystal technology is just beginning to come of age and many applications yet to be realized lie in wait. Optical components that can permit the miniaturization of an application-specific optical integrated circuit (ASIC) to a scale comparable to the wavelength of light will be a good candidate for next-generation high-density optical interconnects and integration. In recent years, there has been a growing interest in the realization of photonic crystals (PhCs) or photonic bandgap structures as optical components and circuits. In this chapter we applied available computational electromagnetics modeling and simulation techniques to develop and optimize application-specific photonic integrated circuits (ASIC) in photonic bandgap structures for near-infrared or telecommunication applications which will be good candidates for next-generation high-density optical computing systems and interconnects. The implication of this work is the ability to incorporate on-chip optical signal processing and routing, on a scale comparable to the wavelength of light. Currently, optical processing devices tend to have a scale much larger than the wavelength of light, which prohibits their use in ‘on-chip applications’.

Patterning photonic crystal thin films into optical circuits would represent the ultimate limit of optoelectronic miniaturization. Integrated circuits that combine conventional electronics and photonics will extend the integrated circuit revolution into the domain of high-bandwidth optical signals.

Results of this study will hopefully be used to realize a new generation of optoelectronic chips to satisfy the growing demand in terms of next-generation optical telecommunication systems.

Photonic crystal telecommunication systems built on a sub-wavelength scale will not only open many exciting opportunities in integrated optics and high-density optical interconnects, but will also provide the basic building blocks for nano-photonics (NPC) of the future. As the development of semiconductor materials has led to the ongoing electronic revolution, high-density optical interconnects in photonic crystals may hold the key for achieving the long-sought goal of large-scale integrated photonic circuits (LSPIC) or optical processing.

The prominent contribution of this work includes the ability to incorporate on-chip optical signal processing and routing, on a scale comparable to the wavelength of light. Currently, optical processing devices tend to have a scale much larger than the wavelength of light, which prohibits their use in ‘on-chip applications’. Optical functions and devices presented include wavelength division multiplexing (channel drop/add filtering) both for WWDM and DWDM.

Towards this end, much remains be done to investigate the properties of individual components incorporated in a photonic crystal, and to identify their breakthrough application to be introduced to the telecommunication market, which will indeed benefit from such miniaturization and integration.

## References

- [1] V. Bykov, ‘Spontaneous Emission in a Periodic Structure’, *Zhurnal Experimentalnoi i Teoreticheskoi Fiziki*, **62**, 1972.
- [2] V. Bykov, ‘Free Oscillations of Elastic Dipole Located in Periodic Structure’, *Zhurnal Tekhnicheskoi Fiziki*, **48**, 1978.
- [3] E. Yablonovitch, ‘Inhibited Spontaneous Emission in Solid-State Physics and Electronics’, *Phys. Rev. Lett.*, **58**, 1987, 2059–2062.

- [4] S. John, ‘Strong Localization of Photons in Certain Disordered Dielectric Superlattices’, *Phys. Rev. Lett.*, **58**, 1987, 2486–2489.
- [5] G. Parker and M. Charlton, ‘Photonic Crystals’, *Physics World*, **13**, 2000.
- [6] M. Bayindir, B. Temelkuran and E. Ozbay, ‘Photonic Crystal Based Beam Splitters’, *Appl. Phys. Lett.*, **77**, 2000, 3902–3904.
- [7] T. Sondergaard and K. H. Dridi, ‘Energy Flow in photonic crystal waveguides’, *Phys. Rev. B*, **61**, 2000, 15688–15696.
- [8] S. Fan, S. G. Johnson and J. D. Joannopoulos, ‘Waveguide branches in photonic crystals’, *J. Opt. Soc. Am. B*, **18**, 2001, 162–165.
- [9] R. W. Ziolkowski, ‘FDTD Analysis of PBG waveguides, power splitters and switches’, *Optical and Quantum Electronics*, **31**, 1999, 843–855.
- [10] *Microcavities and photonic Bandgaps: Physics and Applications*, **324**. Elounda, Crete: Kluwer Academic Publishers, 1995.
- [11] B. D’Urso, O. Painter, J. O’Brien, T. Tombrello, A. Yariv and A. Scherer, ‘Modal reflectivity in finite-depth two-dimensional photonic-crystal microcavities’, *J. Opt. Soc. Am. B*, **15**, 1998, 1155–1159.
- [12] H.-B. Sun, V. Mizeikis, Y. Xu, S. Juodkazis, J.-Y. Ye, S. Matsuo and H. Misawa, ‘Microcavities in polymeric photonic crystals’, *Appl. Phys. Lett.*, **79**, 2001, 1–3.
- [13] S. Fan, P. R. Villeneuve and J. D. Joannopoulos, ‘Microcavities in Photonic Crystals: Mode Symmetry, Tunability, and Coupling Efficiency’, *Physical Review B*, **54**, 1996, 7837–7842.
- [14] S. G. Johnson, F. Shanhui, A. Mekis and J. D. Joannopoulos, ‘Multipole-cancelation mechanism for high-Q cavities in the absence of a complete photonic band gap’, *Appl. Phys. Lett.*, **78**, 2001, 3388–3390.
- [15] R. D. Meade, A. Devenyi, J. D. Joannopoulos, O. L. Alerhand, D. A. Smith and K. Kash, ‘Novel Applications of Photonic Band-Gap Materials – Low-Loss Bends and High Q-Cavities’, *Journal of Applied Physics*, **75**, 1994, 4753–4755.
- [16] A. Scherer, O. Painter, B. D’Urso, R. Lee and A. Yariv, ‘InGaAsP photonic band gap crystal membrane microresonators’, *J. Vac. Sci. Tech. B*, **16**, 1998, 3906–3910.
- [17] A. Sharkawy, S. Shi and D. W. Prather, ‘Multichannel Wavelength Division Multiplexing Using Photonic Crystals’, *Appl. Opt.*, **40**, 2001, 2247–2252.
- [18] R. Stoffer, H. J. W. M. Hoekstra, R. M. D. Ridder, E. V. Groesen and F. P. H. V. Beckum, ‘Numerical Studies of 2D Photonic Crystals: Waveguides, Coupling Between Waveguides and Filters’, *Opt. Quantum Electron.*, **32**, 2000, 947–961.
- [19] R. Fan, R. Villeneuve, J. D. Joannopoulos and H. A. Haus, ‘Channel Drop Tunneling Through Localized States’, *Phys. Rev. Lett.*, **80**, 1998, 960–963.
- [20] H. A. Haus, S. Fan, P. R. Villeneuve and J. D. Joannopoulos, ‘Channel Drop Filters in Photonic Crystals’, *Opt. Express*, **3**, 1998, 4–11.
- [21] S. G. Johnson, P. R. Villeneuve, S. Fan and J. D. Joannopoulos, ‘Linear Waveguides in Photonic-Crystal Slabs’, *Phys. Rev. B*, **62**, 2000, 8212–8222.
- [22] A. Adibi, Y. Xu, R. K. Lee, A. Yariv and A. Scherer, ‘Properties of the slab modes in photonic crystal optical waveguides’, *J. Lightwave Technol.*, **18**, 2000, 1554–1564.
- [23] S. G. Johnson, S. Fan, P. R. Villeneuve and J. D. Joannopoulos, ‘Guided modes in photonic crystal slabs’, *Phys. Rev. B*, **60**, 1999, 5751–5758.
- [24] O. Painter, J. Vuckovic and A. Scherer, ‘Defect modes of a two-dimensional photonic crystal in an optically think dielectric slab’, *J. Opt. Soc. Am. B*, **16**, 1999, 275–285.
- [25] M. D. B. Charlton, M. E. Zoorob, G. J. Parker, M. C. Netti, J. J. Baumberg, S. J. Cox and H. Kemhadjian, ‘Experimental inverstigation of photonic crystal waveguide devices and line-defect waveguide bends’, *Mat. Sci. Eng. B*, **74**, 2000, 17–24.
- [26] M. Gander, R. McBride, J. Jones, D. Mogilevtsev, T. Birks, J. Knight and P. Russell, ‘Experimental measurement of group velocity dispersion in photonic crystal fibre’, *Electron. Lett.*, **35**, 1999, 63–64.
- [27] C. J. M. Smith, R. M. D. L. Rue, M. Rattier, S. Olivier, H. Benisty, C. Weisbuch, P. R. Krauss, R. Houdre and U. Oesterle, ‘Coupled guide and cavity in a two-dimensional photonic crystal’, *Appl. Phys. Lett.*, **78**, 2001, 1487–1489.
- [28] S.-Y. Lin and J. G. Fleming, ‘A three-dimensional optical photonic crystal’, *J. Lightwave Technol.*, **17**, 1999, 1944–1947.
- [29] H. Hirayama, T. Hamano and Y. Aoyagi, ‘Novel surface emitting laser diode using photonic band-gap crystal cavity’, *Appl. Phys. Lett.*, **69**, 1996, 791–793.

- [30] Y. Xia, B. Gates and S. H. Park, ‘Fabrication of three-dimensional photonic crystals for use in the spectral region from ultraviolet to near-infrared’, *J. Lightwave Technol.*, **17**, 1999, 1956–1962.
- [31] S. Rowson, A. Chelnokov, C. Cuisin and J.-M. Lourtioz, ‘Two-Dimensional Photonic Bandgap reflectors for free-propagating beams in the mid-infrared’, *Journal of Optics A-Pure and Applied Optics*, **1**, 1999, 483–489.
- [32] S. Rowson, A. Chelnokov and J. M. Lourtioz, ‘Two-dimensional photonic crystals in macroporous silicon: From mid-infrared (10  $\mu\text{m}$ ) to telecommunication wavelengths (1.3–1.5  $\mu\text{m}$ )’, *Journal of Lightwave Technology*, **17**, 1999, 1989–1995.
- [33] S. Noda, N. Yamamoto, M. Imada, H. Kobayashi and M. Okano, ‘Alignment and stacking of semiconductor photonic bandgaps by wafer-fusion’, *Journal of Lightwave Technology*, **17**, 1999, 1948–1955.
- [34] J. Moosburger, T. Happ, M. Kamp and A. Forchel, ‘Nanofabrication techniques for lasers with two-dimensional photonic crystal mirrors’, *J. Vac. Sci. Tech. B*, **18**, 2000, 3501–3504.
- [35] P. Sabouroux, G. Tayeb and D. Maystre, ‘Experimental and theoretical study of resonant microcavities in two-dimensional photonic crystals’, *Optics Communications*, **160**, 1999, 33–36.
- [36] C. C. Cheng and A. Scherer, ‘Fabrication of photonic band-gap crystals’, *J. Vac. Sci. Tech. B*, **13**, 1995, 2696–2700.
- [37] C. C. Cheng, A. Scherer, R.-C. Tyan, Y. Fainman, G. Witzgall and E. Yablonovitch, ‘New Fabrication Techniques for High Quality Photonic Crystals’, *J. Vac. Sci. Technol. B*, **15**, 1997, 2764–2767.
- [38] A. C. Edrington, A. M. Urbas, P. DeRege, C. X. Chen, T. M. Swager, N. Hadjichristidis, M. Xenidou, L. J. Fetter, J. D. Joannopoulos, Y. Fink and E. L. Thomas, ‘Polymer-Based Photonic Crystals’, *Advanced Materials*, **13**, 2001, 421–425.
- [39] D. J. Norris and Y. Vlasov, ‘Chemical approaches to three-dimensional semiconductor photonic crystals’, *Advanced Materials*, **13**, 2001, 371–376.
- [40] P. Jiang, G. N. Ostojic, R. Narat, D. M. Mittleman and V. L. Colvin, ‘The fabrication and bandgap engineering of photonic multilayers’, *Advanced Materials*, **13**, 2001, 389–393.
- [41] Y. Xia, B. Gates and Z.-Y. Li, ‘Self-Assembly Approaches to Three-Dimensional Photonic Crystals’, *Advanced Materials*, **13**, 2001, 409–413.
- [42] G. Feiertag, W. Ehrfeld, H. Freimuth, H. Kolle, H. Lehr, M. Schmidt, M. M. Sigalas, C. M. Soukoulis, G. Kiriakidis, T. Pedersen, J. Kuhl and W. Koenig, ‘Fabrication of photonic crystals by deep x-ray lithography’, *Appl. Phys. Lett.*, **71**, 1997, 1441–1443.
- [43] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*: Holt, Rinehart and Winston, 1976.
- [44] K. M. Leung and Y. F. Liu, ‘Photon Band Structures: The Plane-Wave Method’, *Phys. Rev. B*, **41**, 1990, 10188–10190.
- [45] L. Liu and J. T. Liu, ‘Photonic Band Structure in the Nearly Plane Wave Approximation’, *European Physical Journal B*, **9**, 1999, 381–388.
- [46] Z. Zhang and S. Satpathy, ‘Electromagnetic Wave Propagation in Periodic Structures: Bloch Wave Solution of Maxwell’s Equations’, *Phys. Rev. Lett.*, **65**, 1990, 2650–2653.
- [47] K. M. Ho, C. T. Chan and C. M. Soukoulis, ‘Existence of a Photonic Gap in Periodic Dielectric Structures’, *Phys. Rev. Lett.*, **65**, 1990, 3152–3155.
- [48] H. Kosaka, T. Kawashima, A. Tomita, M. Notomi, T. Tamamura, T. Sato and S. Kawakami, ‘Superprism Phenomena in Photonic Crystals’, *Phys. Rev. B*, **58**, 1998, R10096–R10099.
- [49] H. Kosaka, T. Kawashima, A. Tomita, M. Notomi, T. Tamamura, T. Sato and S. Kawakami, ‘Photonic Crystals for Micro Lightwave Circuits Using Wavelength-Dependent Angular Beam Steering’, *Appl. Phys. Lett.*, **74**, 1999, 1370–1372.
- [50] M. Tokushima, H. Kosaka, A. Tomita and H. Yamada, ‘Lightwave Propagation Through a 120 Sharply Bent Single-Line-Defect Photonic Crystal Waveguide’, *Appl. Phys. Lett.*, **76**, 2000, 952–954.
- [51] A. Taflove and S. C. Hagness, *Computational Electromagnetics: The Finite-Difference Time-Domain Method*. Boston: Artech House, 2000.
- [52] V. Varadarajan and R. Mittra, ‘Finite-Difference Time-Domain (Fdtd) Analysis Using Distributed Computing’, *4*, 1994, 144–145.
- [53] J. Durban, P. F. Curt, J. R. Humphrey, F. E. Ortiz, D. Prather and M. S. Miroznik, ‘Hardware Implementation of a three dimensional finite difference time domain algorithm’, *IEEE Antennas and Wireless Propagation Letters*, **2**, 2003, 54–57.
- [54] J. Durban, P. F. Curt, J. R. Humphrey, F. E. Ortiz, D. Prather and M. S. Miroznik, ‘Implementation of Three-Dimensional FPGA Based FDTD solvers: An Architectural Overview’, *Proc. 11th Annual IEEE Symposium on Field Programmable Custom Computing Machines (FCCM)*, 2003, pp 269–270.

- [55] J. Durbano, P. F. Curt, J. R. Humphrey, F. E. Ortiz and D. Prather, ‘FPGA based acceleration of the three-dimensional finite difference time domain method for electromagnetic calculations’, *Proc. Global Signal Processing Expo and Conference (GSPx)*, 2004.
- [56] E. Kelmelis, J. Durbano, J. R. Humphrey and F. E. Ortiz, ‘Modeling and simulation of nanoscale devices with a desktop supercomputer’, *Optics and Photonics*, San Diego, 2006.
- [57] Optiwave, ‘OptiFDTD MAX’, 6.0 edn, 2006.
- [58] T. Baba, A. Motegi, T. Iwai, N. Fukaya, Y. Watanabe and A. Sakai, ‘Light propagation characteristics of straight single-line-defect waveguides in photonic crystal slabs fabricated into a silicon-on-insulator substrate’, *IEEE J. Quant. Electron.*, **38**, 2002, 2002.
- [59] A. Scherer, O. Painter, J. Vuckovic, M. Loncar and T. Yoshie, ‘Photonic Crystals for Confining, Guiding, and Emitting Light’, *IEEE Trans. Nanotech.*, **1**, 2002, 4.
- [60] J. P. Berenger, ‘A perfectly matched layer for the absorption of electromagnetic waves’, *J. Comp. Physics*, **114**, 1994, 185–200.
- [61] Z. S. Sacks, D. M. Kingsland, R. Lee and J. Lee, ‘A Perfectly Matched Anisotropic Absorber for Use as an Absorbing Boundary Condition’, *IEEE Trans. Antennas Propagat.*, **AP-43**, 1995, 1460–1463.
- [62] S. Y. Shi, C. H. Chen and D. W. Prather, ‘Plane-wave expansion method for calculating band structure of photonic crystal slabs with perfectly matched layers’, *J. Appl. Phys.*, **21**, 2004, 1769–1775.
- [63] M. Qiu, ‘Effective index method for heterostructure-slab-waveguide-based two-dimensional photonic crystals’, *Appl. Phys. Lett.*, **81**, 2002, 1163–1165.
- [64] C. Chen, S. Shi and D. W. Prather, ‘Stepwise effective index method in modeling of high index contrast planar photonic crystals’, *Proceedings of SPIE – The International Society for Optical Engineering*, Vol. 5360, *Photonic Crystal and Devices II*, 2004, 390–399.
- [65] S. Shi, C. Chen and D. W. Prather, ‘Revised plane wave method for dispersive material and its applications to band structure calculations of photonic crystals’, *Applied Physics Letters*, **86**, 2005, 43104-1–43104-3.
- [66] J. D. Joannopoulos, R. D. Meade and J. N. Winn, *Photonic Crystals: Molding the Flow of Light*. Princeton, N.J.: Princeton University Press, 1995.
- [67] E. Yablonovitch and T. J. Gmitter, ‘Donor and Acceptor Modes in Photonic Band Structures’, *Phys. Rev. Lett.*, **67**, 1991, 3380–3383.
- [68] J. Knight, T. Birks, P. Russell and D. Atkin, ‘All-silica single-mode optical fiber with photonic crystal cladding’, *Optics Letters*, **21**, 1996, 1547–1549.
- [69] T. Sondergaard, J. Arentoft, A. Bjarklev, M. Kristensen, J. Erland, J. Broeng and S. E. Barkou, ‘Designing finite-height photonic crystal waveguides: confinement of light and dispersion relations’, *Optics Communications*, **194**, 2001, 341–351.
- [70] T. Sondergaard, A. Bjarklev, M. Kristensen, J. Erland and J. Broeng, ‘Designing finite-height two-dimensional photonic crystal waveguides’, *Applied Physics Letters*, **77**, 2000, 785–787.
- [71] S. H. Fan, S. G. Johnson, J. D. Joannopoulos, C. Manolatou and H. A. Haus, ‘Waveguide branches in photonic crystals’, *Journal of the Optical Society of America B—Optical Physics*, **18**, 2001, 162–165.
- [72] S. G. Johnson, P. R. Villeneuve, S. H. Fan and J. D. Joannopoulos, ‘Linear waveguides in photonic-crystal slabs’, *Physical Review B*, **62**, 2000, 8212–8222.
- [73] M. D. B. Charlton, M. E. Zoorob, G. J. Parker, M. C. Netti, J. J. Baumberg, S. Cox and H. Kemhadjian, ‘Experimental investigation of photonic crystal waveguide devices and line-defect waveguide bends’, *Materials Science and Engineering B-Solid State Materials for Advanced Technology*, **74**, 2000, 17–24.
- [74] A. Mekis, J. C. Chen, I. Kurland, S. Fan, P. R. Villeneuve and J. D. Joannopoulos, ‘High Transmission through sharp bends in photonic crystal waveguides’, *Phys. Rev. Lett.*, **77**, 1996, 3787–3790.
- [75] Attila Mekis and J. D. Joannopoulos, ‘Tapered Couplers for efficient interfacing between dielectric and photonic crystal waveguides’, *J. Lightwave Technol.*, **19**, 2001, 861–865.
- [76] E. Chow, S. Y. Lin, J. R. Wendt, S. G. Johnson and J. D. Joannopoulos, ‘Quantitative analysis of bending efficiency in photonic-crystal waveguide bends at  $\lambda = 1.55 \mu\text{m}$  wavelengths’, *Optics Letters*, **26**, 2001, 286–288.
- [77] I. EL-Kady, M. M. Sigalas, R. Biswas and K. M. Ho, ‘Dielectric Waveguides in Two-Dimensional Photonic Bandgap Materials’, *J. Lightwave Technol.*, **17**, 1999, 2042–2049.
- [78] M. Loncar, T. Doll, J. Vuckovic and A. Scherer, ‘Design and fabrication of silicon photonic crystal optical waveguides’, *J. Lightwave Technol.*, **18**, 2000, 1402–1411.

- [79] A. R. McGurn, ‘Intrinsic localized modes in nonlinear photonic crystal waveguides: Dispersive modes’, *Physica Letters A*, **260**, 1999, 314–321.
- [80] C. J. M. Smith, H. Benisty, S. Olivier, M. Rattier, C. Weisbuch, T. F. Krauss, R. M. D. L. Rue, R. Houdre and U. Osterle, ‘Low-Loss Channel Waveguides with Two-Dimensional Photonic Crystal Boundaries’, *Appl. Phys. Lett.*, **77**, 2000, 2813–2815.
- [81] V. Berger, O. Gauthier-Lafaye and E. Costard, ‘Photonic band gaps and holography’, *Journal of Applied Physics*, **82**, 1997, 60–64.
- [82] G. J. Schneider, J. Murakowski, S. Venkataraman and D. W. Prather, ‘Combination lithography for photonic-crystal circuits’, *Journal of Vacuum Science and Technology B*, **22**, 2004, 146–151.
- [83] D. W. Prather, J. Murakowski, S. Y. Shi, S. Venkataraman, A. Sharkawy, C. H. Chen and D. Pustai, ‘High-efficiency coupling structure for a single-line-defect photonic-crystal waveguide’, *Optics Letters*, **27**, 2002, 1601–1603.
- [84] D. M. Pustai, A. Sharkawy, S. Y. Shi and D. W. Prather, ‘Tunable photonic crystal microcavities’, *Applied Optics*, **41**, 2002, 5574–5579.
- [85] S. Venkataraman, J. Murakowski, T. N. Adam, J. Kolodzey and D. W. Prather, ‘Fabrication of high-fill-factor photonic crystal devices on silicon-on-insulator substrates’, *Journal of Microlithography Microfabrication and Microsystems*, **2**, 2003, 248–254.
- [86] D. M. Pustai, A. Sharkawy, S. Y. Shi, G. Jin, J. Murakowski and D. W. Prather, ‘Characterization and analysis of photonic crystal coupled waveguides’, *Journal of Microlithography Microfabrication and Microsystems*, **2**, 2003, 292–299.
- [87] D. W. Prather, S. Y. Shi, D. M. Pustai, C. H. Chen, S. Venkataraman, A. Sharkawy, G. J. Schneider and J. Murakowski, ‘Dispersion-based optical routing in photonic crystals’, *Optics Letters*, **29**, 2004, 50–52.
- [88] D. M. Pustai, S. Y. Shi, C. H. Chen, A. Sharkawy and D. W. Prather, ‘Analysis of splitters for self-collimated beams in planar photonic crystals’, *Optics Express*, **12**, 2004, 1823–1831.
- [89] S. Y. Shi, A. Sharkawy, C. H. Chen, D. M. Pustai and D. W. Prather, ‘Dispersion-based beam splitter in photonic crystals’, *Optics Letters*, **29**, 2004, 617–619.
- [90] C. C. Lin, C. H. Chen, A. Sharkawy, G. J. Schneider, S. Venkataraman and D. W. Prather, ‘Efficient terahertz coupling lens based on planar photonic crystals on silicon on insulator’, *Optics Letters*, **30**, 2005, 1330–1332.
- [91] S. Venkataraman, ‘Fabrication of two-dimensional and three-dimensional photonic crystal devices for applications in chip-scale optical interconnects’, In: *Electrical and Computer Engineering*. Newark: University of Delaware, 2005.
- [92] J. Murakowski, D. Pustai and D. Prather, ‘Etchless fabrication of photonic crystals in silicon’, *Journal of Vacuum Science and Technology B*, **20**, 2002, 2013–2016.
- [93] S. G. Johnson and J. D. Joannopoulos, ‘Three-dimensionally periodic dielectric layered structure with omnidirectional photonic band gap’, *Applied Physics Letters*, **77**, 2000, 3490–3492.
- [94] J. G. Fleming and S. Y. Lin, ‘Three-dimensional photonic crystal with a stop band from 1.35 to 1.95  $\mu\text{m}$ ’, *Optics Letters*, **24**, 1999, 49–51.
- [95] Y. A. Vlasov, X. Z. Bo, J. C. Sturm and D. J. Norris, ‘On-chip natural assembly of silicon photonic bandgap crystals’, *Nature*, **414**, 2001, 289–293.
- [96] W. M. Lee, S. A. Pruzinsky and P. V. Braun, ‘Multi-photon polymerization of waveguide structures within three-dimensional photonic crystals’, *Advanced Materials*, **14**, 2002, 271–274.
- [97] M. Campbell, D. N. Sharp, M. T. Harrison, R. G. Denning and A. J. Turberfield, ‘Fabrication of photonic crystals for the visible spectrum by holographic lithography’, *Nature*, **404**, 2000, 53–56.
- [98] H. B. Sun, S. Matsuo and H. Misawa, ‘Three-dimensional photonic crystal structures achieved with two-photon-absorption photopolymerization of resin’, *Applied Physics Letters*, **74**, 1999, 786–788.
- [99] B. H. Cumpston, S. P. Ananthavel, S. Barlow, D. L. Dyer, J. E. Ehrlich, L. L. Erskine, A. A. Heikal, S. M. Kuebler, I. Y. S. Lee, D. McCord-Maughon, J. Q. Qin, H. Rockel, M. Rumi, X. L. Wu, S. R. Marder and J. W. Perry, ‘Two-photon polymerization initiators for three-dimensional optical data storage and microfabrication’, *Nature*, **398**, 1999, 51–54.
- [100] Y. Xu, G. J. Schneider, E. D. Wetzel and D. W. Prather, ‘Centrifugation and spin-coating method for fabrication of three-dimensional opal and inverse-opal structures as photonic crystal devices’, *Journal of Microlithography Microfabrication and Microsystems*, **3**, 2004, 168–173.
- [101] P. Yao, G. J. Schneider, D. W. Prather, E. D. Wetzel and D. J. O’Brien, ‘Fabrication of three-dimensional photonic crystals with multilayer photolithography’, *Optics Express*, **13**, 2005, 2370–2376.

- [102] P. Yao, G. J. Schneider, B. L. Miao, J. Murakowski, D. W. Prather, E. D. Wetzel and D. J. O'Brien, 'Multilayer three-dimensional photolithography with traditional planar method', *Applied Physics Letters*, **85**, 2004, 3920–3922.
- [103] J. Murakowski, G. J. Schneider and D. W. Prather, 'Fabrication of 3D photonic crystals with embedded defects', *Proc. SPIE*, **5347**, 2003, 181–198.
- [104] G. J. Schneider, E. D. Wetzel, J. A. Murakowski and D. W. Prather, 'Fabrication of three-dimensional Yablonovite photonic crystals by multiple-exposure UV interference lithography', *Proc. SPIE*, **5720**, 2005, 9–17.
- [105] P. Yao, G. J. Schneider and D. W. Prather, 'Three-dimensional lithographical fabrication of microchannels', *Journal of Microelectromechanical Systems*, **14**, 2005, 799–805.
- [106] S. Venkataraman, G. J. Schneider, J. Murakowski, S. Y. Shi and D. W. Prather, 'Fabrication of three-dimensional photonic crystals using silicon micromachining', *Applied Physics Letters*, **85**, 2004, 2125–2127.
- [107] D. W. Prather, S. Shi, S. Venkataraman, Z. Lu, J. Murakowski and G. Schneider, 'Self-collimation in 3D photonic crystals', *Proceedings of SPIE – The International Society for Optical Engineering*, Vol. 5733, *Photonic Crystal Materials and Devices III*, 2005, 84–93.
- [108] D. W. Prather and Z. Lu, *article in preparation*, 2005.
- [109] R. D. Meade, A. Devenyi, J. D. Joannopoulos, O. L. Alerhand, D. A. Smith and K. Kash, 'Novel applications of photonic band gap materials: low-loss bends and high Q cavities', *J. Appl. Phys.*, **75**, 1994, 4753.
- [110] S. Y. Lin, E. Chow, V. Hietola, P. R. Villeneuve and J. D. Joannopoulos, 'Experimental Demonstration of Guiding and Bending of Electromagnetic Waves in Photonic Crystal', *Science*, **282**, 1998, 274–276.
- [111] El-Kady, M. M. Sigalas, R. Biswas and K. M. Ho, 'Dielectric waveguides in Two-Dimensional Photonic Bandgap Materials', *J. Lightwave Tech.*, **17**, 1999, 2042.
- [112] S. Y. Lin, E. Chow, S. G. Johnson and J. D. Joannopoulos, 'Demonstration of Highly Efficient Waveguiding in a Photonic Crystal Slab at the 1.5- $\mu\text{m}$  Wavelength', *Optics Letters*, **25**, 2000, 1297–1299.
- [113] M. Qiu, B. Jaskorzynska, M. Swillo and H. Bensity, 'Time-domain 2D modelling of slab-waveguide-based photonic-crystal devices in the presence of radiation losses', *Micro. Opt. Tech. Lett.*, **34**, 2000, 387–393.
- [114] M. Notomi, A. Shinya, K. Yamada, J.-I. Takahashi, C. Takahashi and I. Yokohama, 'Structural tuning of guiding modes of line-defect waveguides of silicon-on-insulator photonic crystal slabs', *IEEE J. Quant. Electron.*, **38**, 2002, 736–742.
- [115] Y. Ohtera, T. Kawashima, Y. Sakai, T. Sato, I. Yokohama, A. Ozawa and S. Kawakami, 'Photonic crystal waveguides utilizing a modulated lattice structure', *Optics Letters*, **27**, 2002, 2158–2160.
- [116] S. J. McNab, N. Moll and Y. A. Vlasov, 'Ultra-low loss photonic integrated circuit with membrane-type photonic crystal waveguides', *Optics Express*, **11**, 2003, 2927–2939.
- [117] M. Loncar, J. Vuckovic and A. Scherer, 'Methods for controlling positions of guided modes of photonic-crystal waveguides', *J. Opt. Soc. Am. B—Optical Physics*, **18**, 2001, 1362–1368.
- [118] H. Benisty, S. Olivier, C. Weisbuch, M. Agio, M. Kafesaki, C. M. Soukoulis, M. Qiu, M. Swillo, A. Karlsson, B. Jaskorzynska, A. Talneau, J. Moosburger, M. Kamp, A. Forchel, R. Ferrini, R. Houdre and U. Oesterle, 'Models and measurements for the transmission of submicron-width waveguide bends defined in two-dimensional photonic crystals', *IEEE Journal of Quantum Electronics*, **38**, 2002, 770–785.
- [119] A. Sharkawy, D. Pustai, S. Y. Shi and D. W. Prather, 'High transmission through waveguide bends by use of polycrystalline photonic-crystal structures', *Optics Letters*, **28**, 2003, 1197–1199.
- [120] A. Chutinan, M. Okano and S. Noda, 'Wider bandwidth with high transmission through waveguide bends in two-dimensional photonic crystal slabs', *Appl. Phys. Lett.*, **80**, 2002, 1698–1700.
- [121] E. M. Purcell, *Phy. Rev.*, **69**, 1946, 681.
- [122] P. L. Gourley, 'Microstructured Semiconductor Lasers for High-Speed Information-Processing', *Nature*, **371**, 1994, 571–577.
- [123] T. Yoshie, J. Vuckovic and A. Scherer, 'High Quality two-dimensional photonic crystal slab cavities', *Appl. Phys. Lett.*, **79**, 2001, 4289–4291.
- [124] K. Srinivasan and O. Painter, 'Momentum space design of high-Q photonic crystal optical cavities', *Optics Express*, **10**, 2002, 670–684.
- [125] P. Lalanne and J. P. Hugonin, 'Bloch-Wave Engineering for High-Q Small-V Microcavities', *IEEE J. Quantum Electronics*, **39**, 2003, 1430–1438.
- [126] P. Lalanne, S. Mias and J. P. Hugonin, 'Two physical mechanisms for boosting the quality factor to cavity volume ratio of photonic crystal microcavities', *Opt. Express*, **12**, 2004, 458–467.

- [127] Y. Akahane, T. Asano, B. Song and S. Noda, ‘High-Q photonic nanocavity in two-dimensional photonic crystal’, *Nature*, **425**, 2003, 944–947.
- [128] J. S. Foresi, P. R. Villeneuve, J. Ferrera, E. R. Thoen, G. Steinmeyer, S. Fan, J. D. Joannopoulos, L. C. Kimerling, H. I. Smith and E. P. Ippen, ‘Photonic-bandgap microcavities in optical waveguides’, *Nature*, **390**, 1997, 143–145.
- [129] R. K. Lee, O. J. Painter, B. D’Urso, A. Scherer and A. Yariv, ‘Measurement of spontaneous emission from a two-dimensional photonic band gap defined microcavity at near-infrared wavelengths.’, *Appl. Phys. Lett.*, **74**, 1999, 1522–1524.
- [130] O. Painter, R. K. Lee, A. Scherer, A. Yariv, J. O’Brien, P. D. Dapkus and I. Kim, ‘Two-Dimensional Photonic Band-Gap Defect Mode Laser’, *Science*, **284**, 1999, 1819–1821.
- [131] O. Painter, A. Husain, A. Scherer, J. O’Brien, I. Kim and P. D. Dapkus, ‘Room Temperature photonic crystal defect lasers at near-infrared wavelengths in InGaAsP’, *J. Lightwave Technol.*, **17**, 1999, 2082–2088.
- [132] H. G. Park, J. K. Hwang, J. Huh, H. Y. Ryu, Y. H. Lee and J. S. Kim, ‘Nondegenerate monopole-mode two-dimensional photonic band gap laser’, *Applied Physics Letters*, **79**, 2001, 3032–3034.
- [133] H. Y. Ryu, S. H. Kwon, Y. J. Lee, Y. H. Lee and J. S. Kim, ‘Very-low-threshold photonic band-edge lasers from free-standing triangular photonic crystal slabs’, *Applied Physics Letters*, **80**, 2002, 3476–3478.
- [134] O. Painter, A. Husain, A. Scherer, P. T. Lee, I. Kim, J. O’Brien and P. D. Dapkus, ‘Lithographic Tuning of a two-dimensional photonic crystal laser array’, *IEEE Photon. Tech. Lett.*, **12**, 2000, 1126–1128.
- [135] J. K. Hwang, H. Y. Ryu, D. S. Song, I. Y. Han, H. W. Song, H. K. Park, Y. H. Lee and D. H. Jang, ‘Room-temperature triangular-lattice two-dimensional photonic band gap lasers operating at 1.54 μm’, *Applied Physics Letters*, **76**, 2000, 2982–2984.
- [136] W. D. Zhou, J. Sabarinathan, B. Kochman, E. Berg, O. Qasaimeh, S. Pang and P. B. Hattacharya, ‘Electrically injected single-defect photonic bandgap surface-emitting laser at room temperature’, *Electronics Letters*, **36**, 2000, 1541–1542.
- [137] M. Qiu, M. Mulot, M. Swillo, S. Anand, B. Jaskorzynska, A. Karlsson, M. Kamp and A. Froehel, ‘Photonic crystal optical filter based on contra-directional waveguide coupling’, *Appl. Phys. Lett.*, **83**, 2003, 5121–5123.
- [138] E. A. Camargo, H. M. H. Chong and R. M. Rue, ‘2D Photonic crystal thermo-optic switch based on AlGaAs/GaAs epitaxial structure’, *Optics Express*, **12**, 2004, 588–592.
- [139] M. F. Yanik, S. Fan, M. Soljacic and J. D. Joannopoulos, ‘All-optical transistor action with bistable switching in a photonic crystal cross-waveguide geometry’, *Opt. Lett.*, **28**, 2003, 2506.
- [140] J. S. Shirk, R. G. S. Pong, S. R. Flom, A. Hiltner and E. Baer, ‘Nonlinear nanolayered polymers: 1D photonic materials with applications to optical limiting’, *Conference on Lasers and Electro-Optics Europe-Technical Digest*, 2001, 83–84.
- [141] H. B. Lin, R. J. Tonucci and A. J. Campillo, ‘Two-dimensional photonic bandgap optical limiter in the visible’, *Optics Letters*, **23**, 1998, 94–96.
- [142] A. M. Zheltikov, N. I. Koroteev, S. A. Magnitsky and A. V. Tarasishin, ‘Self-phase modulation and compression of short laser pulses in a nonlinear photonic crystal’, *Izvestiya Akademii Nauk Seriya Fizicheskaya*, **63**, 1999, 717–724.
- [143] I. Abdulhalim, ‘Reflective phase-only modulation using one-dimensional photonic crystals’, *Journal of Optics A—Pure and Applied Optics*, **2**, 2000, L9–L11.
- [144] N. Susa, ‘Change in transmittance due to free carriers in two-dimensional photonic crystals’, *Japanese Journal of Applied Physics Part 1—Regular Papers Short Notes and Review Papers*, **39**, 2000, 6288–6289.
- [145] S. Hava, J. Ivri and M. Auslender, ‘Wavenumber-modulated patterns of transmission through one- and two-dimensional gratings on a silicon substrate’, *Journal of Optics A—Pure and Applied Optics*, **3**, 2001, S190–S195.
- [146] H. L. Guo, H. Y. Chen, P. G. Ni, Q. Zhang, B. Y. Cheng and D. Z. Zhang, ‘Transmission modulation in the passband of polystyrene photonic crystals’, *Applied Physics Letters*, **82**, 2003, 373–375.
- [147] K. G. Hougaard, J. Broeng and A. Bjarklev, ‘Low pump power photonic crystal fibre amplifiers’, *Electronics Letters*, **39**, 2003, 599–600.
- [148] T. V. Andersen, J. Thogersen, S. R. Keiding and J. J. Larsen, ‘High-power intracavity frequency doubling of a Ti : sapphire femtosecond oscillator’, *Applied Physics B—Lasers and Optics*, **76**, 2003, 639–644.
- [149] S. Lan, K. Kanamoto, T. Yang, S. Nishikawa, Y. Sugimoto, N. Ikeda, H. Nakamura, K. Asakawa and H. Ishikawa, ‘Similar role of waveguide bends in photonic crystal circuits and disordered defects in coupled cavity waveguides: An intrinsic problem in realizing photonic crystal circuits’, *Physical Review B*, **67**, 2003, 115208.

- [150] S. Noda, A. Chutinan and M. Imada, ‘Trapping and emission of photons by a single defect in a photonic bandgap structure’, *Nature*, **407**, 2000, 608–610.
- [151] Y. Ohtera, T. Sato, T. Kawashima, T. Tamamura and S. Kawakami, ‘Photonic crystal polarisation splitters’, *Electronics Letters*, **35**, 1999, 1271–1272.
- [152] S. Y. Lin, E. Chow, J. Bur, S. G. Johnson and J. D. Joannopoulos, ‘Low-loss, wide-angle Y splitter at similar to 1.6- $\mu$ m wavelengths built with a two-dimensional photonic crystal’, *Optics Letters*, **27**, 2002, 1400–1402.
- [153] Y. Sugimoto, N. Ikeda, N. Carlsson, K. Asakawa, N. Kawai and K. Inoue, ‘AlGaAs-based two-dimensional photonic crystal slab with defect waveguides for planar lightwave circuit applications’, *IEEE Journal of Quantum Electronics*, **38**, 2002, 760–769.
- [154] R. Wilson, T. J. Karle, I. Moerman and T. F. Krauss, ‘Efficient photonic crystal Y-junctions’, *J. of Opt. A*, **5**, 2003, S76–S80.
- [155] M. H. ShiH, W. J. Kim, W. Kuang, J. R. Cao, H. Yukawa, S. J. Choi, J. D. O’Brien, P. D. Dapkus and W. K. Marshall, ‘Two-dimensional photonic crystal Mach-Zehnder interferometers’, *Appl. Phys. Lett.*, **84**, 2004, 460–462.
- [156] S.-Y. Lin, V. M. Hietala, L. Wang and E. D. Jones, ‘Highly dispersive photonic band-gap prism’, *Optics Letters*, **21**, 1996, 1771–1773.
- [157] H. Kosaka, T. Kawashima, A. Tomita, M. Notomi, T. Tamamura, T. Sato and S. Kawakami, ‘Self-Collimating Phenomena in Photonic Crystals’, *Appl. Phys. Lett.*, **74**, 1999, 1212–1214.
- [158] J. Witzens, M. Loncar and A. Scherer, ‘Self-collimation in planar photonic crystals’, *IEEE J. Sel. Topics Quantum Electron.*, **8**, 2002, 1246–1257.
- [159] L. Wu, M. Mazilu and T. F. Krauss, ‘Beam steering in Planar-Photonic Crystals: From Superprism to Super-collimator’, *J. Lightwave Tech.*, **21**, 2003, 561–566.
- [160] D. N. Chigrin, S. Enoch, S. C. M. Torres and G. Tayeb, ‘Self-guiding in two-dimensional photonic crystals’, *Optics Express*, **11**, 2003, 1203–1211.
- [161] D. W. Prather, S. Shi, D. Pustai, C. Chen, S. Venkataraman, A. Sharkawy, G. Schneider and J. Murakowski, ‘Dispersion-based optical routing in photonic crystals’, *Opt. Lett.*, **29**, 2004, 50–52.
- [162] D. W. Prather, J. Murakowski, S. Shouyuan, S. Venkataraman, A. Sharkawy, C. Chen and D. Pustai, ‘High Efficiency Coupling Structure for a single Line-Defect Photonic Crystal Waveguide’, *Optics Letters*, **27**, 2002, 1601–1603.
- [163] C. Chen, A. Sharkawy, D. M. Pustai, S. Shi and D. W. Prather, ‘Optimizing bending efficiency of self-collimated beams in non-channel planar photonic crystal waveguides’, *Opt. Express*, **11**, 2003, 3153–3159.
- [164] C. Chen, G. Jin, S. Shouyuan, A. Sharkawy and D. W. Prather, ‘A unidirectional photonic crystal dispersion-based emitter’, *Appl. Phys. Lett.*, **84**, 2004, 3151–3153.
- [165] S. G. Johnson and J. D. Joannopoulos, ‘Block-Iterative Frequency-Domain Methods for Maxwell’s Equations in a Planewave Basis’, *Opt. Express*, **8**, 2001, 173–180.
- [166] D. M. Pustai, S. Shouyuan, C. Chen, A. Sharkawy and D. W. Prather, ‘Analysis of splitters for self-collimated beams in planar photonic crystals’, *Optics Express*, **12**, 2004, 1823–1831.
- [167] S. Shi, A. Sharkawy, C. Chen, D. M. Pustai and D. W. Prather, ‘Dispersion based beam-splitter in photonic crystals’, *Optics Letters*, **29**, 2004, 617.

# 4

## Optical Modulators in Silicon Photonic Circuits

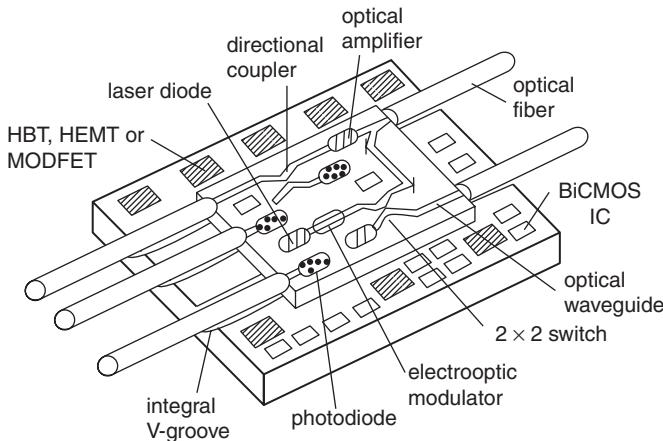
F.Y. Gardes, G.T. Reed, G.Z. Mashanovich and C.E. Png

### 4.1 Introduction

Ever since the earliest research on optical circuits, dating back to the 1970s, there have been visions of an optical superchip, e.g. [1, 2], containing a variety of integrated optical components to carry out light generation, modulation, manipulation, detection and amplification (Figure 4.1). The early work was associated with ferroelectric materials such as lithium niobate ( $\text{LiNbO}_3$ ), and III-V semiconductors such as the gallium arsenide (GaAs) and indium phosphide (InP) based systems.  $\text{LiNbO}_3$  was interesting almost solely due to the fact that it possesses a large electro-optic coefficient [1], enabling optical modulation via the Pockels effect. Alternatively, the III-V compounds were interesting because of the relative ease of laser fabrication, and the prospect of optical and electronic integration.

However, the dominance of silicon as the semiconductor of choice for microelectronics eventually led to the investigation of silicon photonic circuits, primarily because of the potential attraction of integration with electronics in a cost-effective manner. Such research began in the mid 1980s. However, silicon is inherently restricted as an optical material for two primary reasons. Firstly the indirect bandgap of silicon means that light emission from silicon is inefficient, and nontraditional techniques must be investigated if light emission is to be realised. Such work is progressing rapidly and significant progress has been made, e.g. [1, 3].

Room-temperature silicon light emitting diodes (LEDs) are now a reality, e.g. [4] and the rate of progress is such that many researchers are predicting the imminent arrival of the electrically pumped silicon laser. These issues are discussed in more detail in Chapter 5. Secondly, the centrosymmetric crystal structure of silicon means that it does not exhibit a linear electro optic (Pockels) effect. Since this is the traditional means of implementing an optical modulator in a waveguide-based device, modulation of the refractive index of silicon must be carried out in



**Figure 4.1** Silicon-based optoelectronic integrated circuits (OEIC) ‘superchip’. Reproduced from [1–2] by permission of IEEE © (1993)

another way. This chapter discusses the progress of optical modulation in silicon, which has been significant in recent years [5].

It is worth noting that the success of research in these two fields of silicon photonics in recent years has resulted in a resurgence of interest, as demonstrated by the plethora of silicon photonics conferences and symposia now in existence, as well as the wealth of scientific literature that has emerged in the last five years. For the second time in a short research history, optimism is high in the field of silicon photonics, that the functional building blocks of the silicon photonic circuit, and the integration with electronics is a viable technological and commercial option.

## 4.2 Mechanisms for Optical Modulation in Silicon

Optical modulators in photonic circuits are implemented via devices that cause direct changes in optical intensity via absorption, or cause changes in the refractive index of the material (and hence the phase of a propagating wave), which can be converted to an intensity change via an interferometer [6] or a resonant device [7]. Such refractive index changes are typically produced via some applied electrical signal, and whilst the design of the modulator will determine the required change in refractive index, typically a change of at least the order of  $10^{-4}$  is required to make the device viable at reasonable dimensions [6]. The preferred means of implementing modulation is by the application of an electric field, because this implies little or no current flow (and hence low power) and a fast response time. The application of an electric field to a material can result in a change to the real and imaginary refractive indices. A change in real refractive index  $\Delta n$ , is known as electrorefraction, and a change in the imaginary part of refractive index  $\Delta\alpha$  is known as electroabsorption. The primary electric field effects that are useful in semiconductor materials are the Pockels effect, the Kerr effect, and the Franz-Keldysh effect and will be discussed later, but lets first introduce the absorption mechanism in silicon.

### 4.2.1 Absorption

The two main potential sources of absorption loss for semiconductor waveguides are band edge absorption (or interband absorption), and free carrier absorption. Interband absorption occurs when photons with energy greater than the bandgap, are absorbed to excite electrons from the valence band to the conduction band. Therefore to avoid interband absorption, a wavelength must be used that is longer than the absorption edge wavelength of the waveguide material. Silicon is an excellent example material to demonstrate this point. The band edge wavelength of silicon is approximately 1.1 μm, above which silicon is used as a waveguide material. For wavelengths shorter than 1.1 μm, silicon absorbs very strongly, and is one of the most common materials used for photodetectors in the visible wavelength range, and at very short infrared wavelengths.

The band edge absorption of a material does not mark an abrupt transition from strong absorbance to transparency, so care must be taken when selecting a wavelength for a given technology. For example, the absorption of pure silicon at a wavelength of  $\lambda = 1.15 \mu\text{m}$  results in an attenuation of 2.83 dB/cm [8], whereas moving to a wavelength of  $\lambda = 1.52 \mu\text{m}$  reduces the loss to 0.004 dB/cm [8]. Therefore, semiconductor waveguides should suffer negligible band edge absorption, provided a suitable wavelength of operation is chosen.

Free carrier absorption however, may be significant in semiconductor waveguides. The concentration of free carriers will affect both the real and imaginary refractive indices. For devices fabricated in silicon, modulation of the free carrier density is used to deliberately modulate refractive index, an effect that will be discussed in more detail later in this chapter when optical modulators are considered in more detail. Changes in absorption in semiconductors can be described by the well known Drude–Lorenz equation [9]:

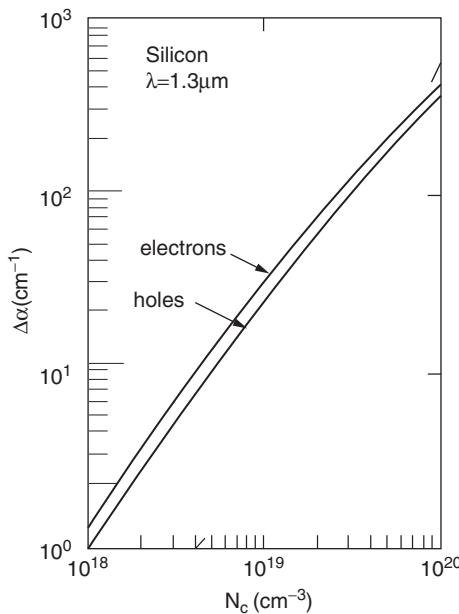
$$\Delta\alpha = \frac{e^3\lambda_0^2}{4\pi^2c^3\varepsilon_0 n} \left( \frac{N_e}{\mu_e(m_{ce}^*)^2} + \frac{N_h}{\mu_h(m_{ch}^*)^2} \right) \quad (4.1)$$

where  $e$  is the electronic charge;  $c$  is the velocity of light in vacuum;  $\mu_e$  is the electron mobility;  $\mu_h$  is the hole mobility;  $m_{ce}^*$  is the effective mass of electrons;  $m_{ch}^*$  is the effective mass of holes;  $N_e$  is the free electron concentration;  $N_h$  is the free hole concentration;  $\varepsilon_0$  is the permittivity of free space; and  $\lambda_0$  is the free space wavelength.

Some of the parameters in Equation (4.1) are interdependent, so care must be taken when evaluating the effect of free carrier absorption. However, in order to demonstrate the significance of free carrier absorption let us consider the additional absorption introduced by free carriers. Soref and Lorenzo evaluated Equation (4.1) for values of  $N$  in the range  $10^{18}\text{--}10^{20} \text{ cm}^{-3}$ , and presented the data graphically, as shown in Figure 4.2 [10]. It can be seen from Figure 4.2, that an injected hole and electron concentration of  $10^{18} \text{ cm}^{-3}$  introduces a total additional loss of approximately  $2.5 \text{ cm}^{-1}$ . This corresponds to a loss of 10.86 dB/cm, indicating the dramatic effect doping of the semiconductor can have on the loss of a waveguide.

### 4.2.2 Optical Modulation Mechanisms in Silicon

One of the requirements of an integrated optical technology, particularly one related to communications, is the ability to perform optical modulation. This implies a change in the optical field due to some applied signal, typically, although not exclusively an electrical signal. The



**Figure 4.2** Additional loss of silicon due to free carriers. Reproduced from [10] by permission of IEEE © (1986)

change in the optical field is usually derived from a change in refractive index of the material involved, with the applied field, although other parametric changes are possible. It is now widely accepted that the most efficient means of implementing optical modulation in silicon via an electrical signal, is to use carrier injection or depletion. This will be discussed later. First however, let us consider other electrical modulation techniques to discover why they are not useful in silicon, even though they are used in other integrated optical technologies. The primary candidates for electrically derived modulation are electric field effects.

#### 4.2.2.1 Electric Field Effects

The application of an electric field to a material can result in a change to the real and imaginary part of the refractive indices. A change in the real part of the refractive index  $\Delta n$  with an applied electric field is known as electrorefraction, and a change in the imaginary part of refractive index  $\Delta\alpha$  with applied electric field is known as electroabsorption. The primary electric field effects that are useful in semiconductor materials to cause either electroabsorption or electrorefraction are the Pockels effect, the Kerr effect, and the Franz–Keldysh effect. Soref and Bennett [11] have examined electric field effects in silicon, and their results will be cited here to demonstrate the relative efficiency of several electric field effects. However, let us first briefly discuss the three electric field effects mentioned above.

*The Pockels effect*, also known as the linear electro-optic effect causes a change in real refractive index  $\Delta n$  which is proportional to the applied field  $E$ . Therefore, if the applied field is uniform, the change in refractive index will be proportional to applied voltage  $V$  for a fixed

modulator geometry. The Pockels effect in general, produces a change in refractive index that is dependent upon the direction of the applied electric field with respect to the crystal axes. Therefore, the effect is also usually polarisation dependent, although simplifications occur due to the symmetry of the crystal structure of a given material, in any given direction. The geometry of the silicon crystal structure is such that the Pockels effect disappears completely, and therefore is not an option for optical modulation in silicon. However, more generally it is usual to align the applied electric field with one of the principal axes of the crystal, to utilise the largest electro-optic coefficient for the material concerned. For example, for lithium niobate ( $\text{LiNbO}_3$ ), the maximum refractive index change occurs if the so-called  $r_{33}$  coefficient is utilised. This results in a refractive index change in  $\text{LiNbO}_3$  given by

$$\Delta n = -r_{33}n_{33} \frac{E_3}{2} \quad (4.2)$$

where  $n_{33}$  is the refractive index in the direction of the applied electric field, and  $E_3$  is the applied electric field. The subscript 3 merely indicates which of the three principal axes of the material is aligned with the applied field. The value of  $r_{33}$  is  $30.8 \times 10^{-12} \text{ m/V}$ .

The *Kerr effect* is a second-order electric field effect in which the change in real refractive index  $\Delta n$  is proportional to the square of the applied electric field. This effect is present in silicon, although it is relatively weak. The change may be expressed as:

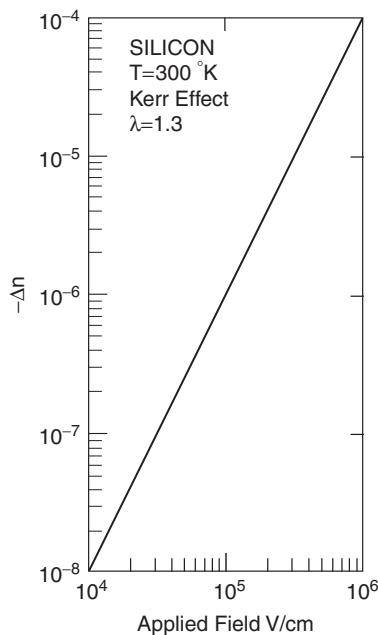
$$\Delta n = s_{33}n_0 \frac{E^2}{2} \quad (4.3)$$

where  $s_{33}$  is the Kerr coefficient,  $n_0$  is the unperturbed refractive index, and  $E$  is the applied field. In this case the sign of the refractive index change is not dependent on the direction within the crystal axis.

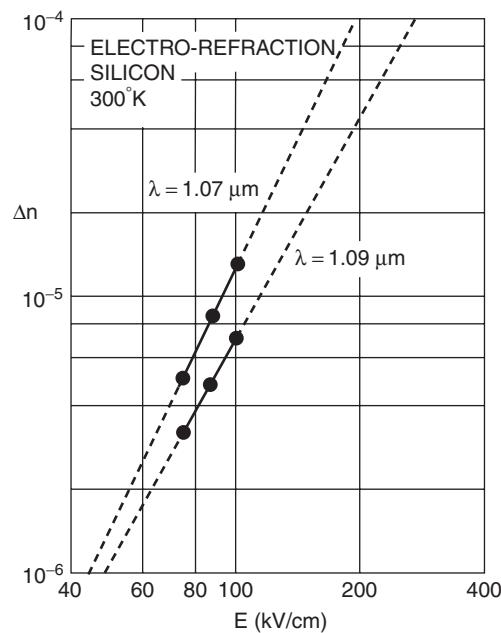
Soref and Bennett theoretically quantified the refractive index change in silicon at a wavelength of  $1.3 \mu\text{m}$ , as a function of the applied electric field. Their results are shown graphically in Figure 4.3. It can be seen from Figure 4.3 that the change in refractive index  $\Delta n$  is predicted to reach  $10^{-4}$  at an applied field of  $10^6 \text{ Vs/cm}$  ( $100 \text{ V}/\mu\text{m}$ ), which is above the breakdown field for lightly doped silicon. It will be seen later that this is a relatively small effect when compared with the plasma dispersion effect (carrier injection or depletion).

Unlike the Pockels effect and the Kerr effect, the *Franz-Keldysh effect* gives rise to both electrorefraction and electroabsorption, although primarily the latter. The effect is due to distortion of the energy bands of the semiconductor upon application of an electric field. In effect, this shifts the bandgap energy, resulting in a change in the absorption properties of the crystal, particularly at wavelengths close to the bandgap, and hence a change in the complex refractive index. Soref and Bennett [11] also quantified the changes in refractive index due to the Franz–Keldysh effect. Whilst all their data was not presented graphically, they plotted the change in refractive index as a function of the applied electric field at a wavelength of  $1.07 \mu\text{m}$ , the wavelength at which the effect is greatest. They also plotted data at  $1.09 \mu\text{m}$  for comparison. It should be noted, however, that the effect diminishes significantly at the telecommunications wavelengths of  $1.31 \mu\text{m}$ , and  $1.55 \mu\text{m}$ . The data are shown in Figure 4.4.

From Figure 4.4, it can be seen that the refractive index change reaches  $10^{-4}$  at an applied field of  $2 \times 10^5 \text{ V/cm}$  ( $20 \text{ V}/\mu\text{m}$ ). Whilst this figure is better than that evaluated for the Kerr effect, it must be remembered that this is not the figure evaluated at the same wavelength. The



**Figure 4.3** The Kerr effect in silicon as a function of applied electric field. Reproduced from [11] by permission of IEEE © (1987)



**Figure 4.4** The Franz–Keldysh effect in silicon. Reproduced from [11] by permission of IEEE © (1987)

Franz–Keldysh effect will fall very significantly at 1.3 μm, the wavelength at which the results above for the Kerr effect are quoted.

*The plasma dispersion effect (carrier injection or depletion):* we observed previously that the concentration of free charges in silicon contributes to the loss via absorption. The imaginary part of the refractive index is determined by the absorption (or loss) coefficient. Therefore it is clear that changing the concentration of free charges can change the refractive index of the material. We saw earlier that the Drude–Lorenz equation relating the concentration of electrons and holes to the absorption (Equation 4.1) was:

$$\Delta\alpha = \frac{e^3\lambda_0^2}{4\pi^2c^3\varepsilon_0n} \left( \frac{N_e}{\mu_e(m_{ce}^*)^2} + \frac{N_h}{\mu_h(m_{ch}^*)^2} \right) \quad (4.4)$$

The corresponding equation relating the carrier concentration  $N$  to the change in refractive index  $\Delta n$  is [9]:

$$\Delta n = \frac{-e^2\lambda_0^2}{8\pi^2c^2\varepsilon_0n} \left( \frac{N_e}{m_{ce}^*} + \frac{N_h}{m_{ch}^*} \right) \quad (4.5)$$

Soref and Bennett [11] studied results in the scientific literature to evaluate the change in refractive index  $\Delta n$  to experimentally produced absorption curves for a wide range of electron and hole densities, over a wide range of wavelengths. In particular they focused on the communications wavelengths of 1.3 and 1.55 μm. Interestingly their results were in good agreement with the classical Drude–Lorenz model, only for electrons. For holes they noted a  $(\Delta N)^{0.8}$  dependence.

The changes in refractive index and absorption coefficient due to injected carriers in silicon can be produced from the Kramers–Kronig relations between the real and the imaginary parts of the refractive index. These relations can be derived using the following approach.

The electric displacement vector in a medium is given by

$$\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P} \quad (4.6)$$

where  $\mathbf{E}$  is electric field and  $\mathbf{P}$  is polarisation. When a medium is isotropic  $\mathbf{E}$ ,  $\mathbf{D}$  and  $\mathbf{P}$  are collinear. For small fields,  $\mathbf{P}$  is proportional to  $\mathbf{D}$ , and therefore, for isotropic media we may write:

$$\begin{aligned} D &= \varepsilon_0 E + P = \varepsilon E \\ P &= D - \varepsilon_0 E = (\varepsilon - \varepsilon_0)E = (n^2 - 1)\varepsilon_0 E \end{aligned} \quad (4.7)$$

where  $n$  is refractive index of the medium. For a homogeneous isotropic medium and electric field with the magnitude of  $E = E_0\exp(j\omega t)$  we can write the following relations for the polarisation  $P$  and displacement  $D$ :

$$P = (n^2 - 1)\varepsilon_0 E = P_0 e^{j\omega t} \quad (4.8)$$

$$D = \varepsilon E = D_0 e^{j\omega t} \quad (4.9)$$

If the medium is dispersive then permittivity  $\varepsilon$  is a function of  $\omega$ . As the medium is homogeneous, permittivity is not a function of coordinates  $x$ ,  $y$  and  $z$ . If we assume that permittivity

does not change with time then:

$$D(t - t_0) = \varepsilon E(t - t_0) \quad (4.10)$$

Although  $\varepsilon$  may be complex, it is not a function of the electric field and the principle of superposition is valid. If the electric field is applied to the medium at  $t = 0$  ( $E(t) = 0$  for  $t < 0$ ), then the polarisation and displacement must be zero for  $t < 0$ :

$$\begin{aligned} P(t) &= 0 & (t < 0) \\ D(t) &= 0 & (t < 0) \end{aligned} \quad (4.11)$$

If the electric field is applied as a very short pulse at  $t = 0$ , that is

$$E = \delta(t) \quad (4.12)$$

where  $\delta(t)$  is delta function, then the electric field can be represented as [12]:

$$E(t) = \sqrt{\frac{2}{\pi}} \Re \int_0^\infty e^{j\omega t} d\omega \quad (4.13)$$

where symbol  $\Re$  represents the real part of the integral. According to Equation (4.9), the displacement is then given by

$$D = \sqrt{\frac{2}{\pi}} \Re \int_0^\infty \varepsilon e^{j\omega t} d\omega \quad (4.14)$$

If the medium is nondispersive, i.e.  $\varepsilon \neq \varepsilon(\omega)$ , then  $D$  is also a delta function (at  $t = 0$ ) and Equation (4.11) is satisfied. If however, the medium is dispersive  $\varepsilon = \varepsilon(\omega)$  this is no longer the case. For a dispersive medium we can express the permittivity as

$$\varepsilon = \varepsilon(\omega) + \varepsilon_c \quad (4.15)$$

where  $\varepsilon_c$  is a constant equal to the value of  $\varepsilon$  when  $\omega \rightarrow \infty$  so that  $\varepsilon(\omega) = 0$  when  $\omega \rightarrow \infty$ . The displacement can be then expressed in the following form:

$$D = \varepsilon_c E + \sqrt{\frac{2}{\pi}} \Re \int_0^\infty \varepsilon(\omega) e^{j\omega t} d\omega \quad (4.16)$$

According to Equation (4.11), for  $t < 0$  the electric field and displacement are zero,  $E(t) = 0$  and  $D(t) = 0$ , and therefore Equation (4.16) can be rewritten as

$$\sqrt{\frac{2}{\pi}} \Re \int_0^\infty \varepsilon(\omega) e^{j\omega t} d\omega = 0 \quad (4.17)$$

As the permittivity  $\varepsilon(\omega)$  is generally a complex number, we can express it as

$$\varepsilon(\omega) = \varepsilon_r(\omega) - j\varepsilon_i(\omega) \quad (4.18)$$

where  $\varepsilon_r$  and  $\varepsilon_i$  are real and imaginary part of the permittivity respectively. By substituting Equation (4.18) into (4.17), the former equation becomes:

$$\sqrt{\frac{2}{\pi}} \int_0^\infty \varepsilon_r(\omega) \cos(\omega t) d\omega + \sqrt{\frac{2}{\pi}} \int_0^\infty \varepsilon_i(\omega) \sin(\omega t) d\omega = 0 \quad (4.19)$$

The previous equation is valid for  $t < 0$ . If we designate  $t$  as  $-t_1$ , and then replace  $t_1$  with  $t$ , Equation (4.19) can be written as

$$\int_0^\infty \varepsilon_r(\omega) \cos(\omega t) d\omega = \int_0^\infty \varepsilon_i(\omega) \sin(\omega t) d\omega \quad (4.20)$$

If we define

$$p(t) = \sqrt{\frac{2}{\pi}} \int_0^\infty \varepsilon_r(\omega) \cos(\omega t) d\omega \quad (4.21)$$

it is clear that  $\varepsilon_r(\omega)$  is Fourier cosine transformation [13] of  $p(t)$

$$\varepsilon_r(\omega) = \sqrt{\frac{2}{\pi}} \int_0^\infty p(t) \cos(\omega t) dt \quad (4.22)$$

As  $p(t)$  is equal to the left-hand side of Equation (4.20) it must be also equal to the right-hand side of the same equation:

$$p(t) = \sqrt{\frac{2}{\pi}} \int_0^\infty \varepsilon_i(\omega) \sin(\omega t) d\omega \quad (4.23)$$

Substituting Equation (4.23) into (4.22) the real part of the permittivity  $\varepsilon_r(\omega)$  becomes:

$$\varepsilon_r(\omega) = \frac{2}{\pi} \int_0^\infty \int_0^\infty \varepsilon_i(\omega) \sin(\omega_1 t) \cos(\omega t) dt d\omega_1 \quad (4.24)$$

where  $\omega_1$  is used instead of  $\omega$  to indicate which part should be integrated. We need to calculate the integral  $I$  from Equation (4.24)

$$I = \int_0^\infty \sin(\omega_1 t) \cos(\omega t) dt \quad (4.25)$$

This integral can be written as

$$I = \Re \int_0^\infty \sin(\omega_1 t) e^{j\omega t} dt = \Re F\{\sin(\omega_1 t)\} \quad (4.26)$$

and is actually a Fourier transformation of  $\sin(\omega_1 t)$ . We can write the transform of  $\sin(\omega_1 t)$  as

$$F\{\sin(\omega_1 t)\} = \lim_{a \rightarrow 0^+} F\{\sin(\omega_1 t)e^{-at}\} \quad (4.27)$$

Now, it is possible to find the Fourier transform

$$F\{\sin(\omega_1 t)\} = \lim_{a \rightarrow 0^+} \int_0^\infty \sin(\omega_1 t) e^{-at} e^{-j\omega t} dt = \lim_{a \rightarrow 0^+} \int_0^\infty \sin(\omega_1 t) e^{-st} dt \quad (4.28)$$

$$F\{\sin(\omega_1 t)\} = \lim_{a \rightarrow 0^+} \int_0^\infty \frac{e^{j\omega_1 t} - e^{-j\omega_1 t}}{2j} e^{-st} dt = \lim_{a \rightarrow 0^+} \int_0^\infty \frac{e^{-(s-j\omega_1)t} - e^{-(s+j\omega_1)t}}{2j} dt \quad (4.29)$$

$$F\{\sin(\omega_1 t)\} = \lim_{a \rightarrow 0^+} \frac{1}{2j} \left[ \frac{1}{s - j\omega_1} - \frac{1}{s + j\omega_1} \right] = \lim_{a \rightarrow 0^+} \frac{\omega_1}{s^2 + \omega_1^2} = \lim_{a \rightarrow 0^+} \frac{\omega_1}{(a + j\omega)^2 + \omega_1^2} \quad (4.30)$$

Finally

$$I = \frac{\omega_1}{\omega_1^2 - \omega^2} \quad (4.31)$$

and  $\varepsilon_r(\omega)$  from Equation (4.24) becomes:

$$\varepsilon_r(\omega) = \frac{2}{\pi} P \int_0^\infty \frac{\omega_1 \varepsilon_i(\omega_1)}{\omega_1^2 - \omega^2} d\omega_1 \quad (4.32)$$

$P$  in Equation (4.32) indicates that the principal part must be taken, as the integral has a singularity at  $\omega_1 = \omega$ . To evaluate this integral it is necessary to use contour integration [14].

Similarly, it can be shown that

$$\varepsilon_i(\omega) = \frac{2}{\pi} \omega P \int_0^\infty \frac{\omega_1 \varepsilon_r(\omega_1)}{\omega^2 - \omega_1^2} d\omega_1 \quad (4.33)$$

The last two equations express a relation between the real and imaginary parts of  $\varepsilon - \varepsilon_c$ . Sometimes it is desirable to have a relation between real and imaginary parts of the refractive index  $\mathbf{n} = n_r - jk$ . Such a relation can be derived from Equation (4.14), and it can be found that:

$$n_r(\omega) - 1 = \frac{c}{\pi} P \int_0^\infty \frac{\alpha(\omega_1)}{\omega_1^2 - \omega^2} d\omega_1 \quad (4.34)$$

where  $\alpha = k\omega/c$  is the absorption coefficient. As we can see from Equation (4.34), the real and imaginary parts of the refractive index are coupled. Therefore, we can say that the index perturbation  $\Delta n_r$  arises physically from a change in the optical absorption spectrum  $\Delta\alpha$ . The differential Kramers–Kronig dispersion relation [15] can be derived from

Equation (4.34):

$$\Delta n_r(\omega) = \frac{c}{\pi} P \int_0^{\infty} \frac{\Delta\alpha(\omega_1)}{\omega_1^2 - \omega^2} d\omega_1 \quad (4.35)$$

where  $\Delta\alpha(\omega) = \alpha(\omega, \Delta N) - \alpha(\omega, 0)$  and  $\Delta N$  is the change in free-carrier concentration within crystalline silicon.

Soref and Bennett used experimental data for the absorption spectrum available in the literature and Equation (4.34) to compute the integral and consequently the dependence of the change in the refractive index with carrier concentration. They produced expressions relating the refractive index and absorption coefficients changes in silicon due to injection or depletion of carriers in silicon [11] at both 1.3 and 1.55  $\mu\text{m}$  wavelengths, which are now almost universally used for evaluation of optical modulation in silicon.

At  $\lambda_0 = 1.55 \mu\text{m}$  these expressions are:

$$\Delta n = \Delta n_e + \Delta n_h = -8.8 \times 10^{-22} \Delta N_e - 8.5 \times 10^{-18} (\Delta N_h)^{0.8} \quad (4.36)$$

$$\Delta\alpha = \Delta\alpha_e + \Delta\alpha_h = 8.5 \times 10^{-18} \Delta N_e + 6.0 \times 10^{-18} \Delta N_h \quad (4.37)$$

where  $\Delta n_e$  = change in refractive index resulting from the change in free electron carrier concentrations;  $\Delta n_h$  = change in refractive index resulting from the change in free hole carrier concentrations;  $\Delta\alpha_e$  = change in absorption resulting from the change in free electron carrier concentrations;  $\Delta\alpha_h$  = change in absorption resulting from the change in free hole carrier concentrations.

Similarly at  $\lambda_0 = 1.3 \mu\text{m}$ :

$$\Delta n = \Delta n_e + \Delta n_h = -6.2 \times 10^{-22} \Delta N_e - 6.0 \times 10^{-18} (\Delta N_h)^{0.8} \quad (4.38)$$

$$\Delta\alpha = \Delta\alpha_e + \Delta\alpha_h = 6.0 \times 10^{-18} \Delta N_e + 4.0 \times 10^{-18} \Delta N_h. \quad (4.39)$$

We will see later that a carrier injection level of the order of  $5 \times 10^{17}$  is readily achievable for both electrons and holes. Therefore, if we evaluate the change in refractive index  $\Delta n$  at a wavelength of 1.3  $\mu\text{m}$  (Equation 4.38), for comparison with the field effects evaluated earlier we obtain:

$$\Delta n = -6.2 \times 10^{-22} (5 \times 10^{17}) - 6.0 \times 10^{-18} (5 \times 10^{17})^{0.8} = -1.17 \times 10^{-3} \quad (4.40)$$

This is more than an order of magnitude larger than the changes due to the electric field effects described earlier. Furthermore, by appropriate high doping of contacts to the silicon layer, it is reasonable to expect higher injection levels, and therefore higher refractive index changes.

*The thermo-optic effect:* in addition to the electric field effects and change in concentration of free carriers in silicon, one other modulation technique has proved viable for optical modulation devices in silicon. It is the thermo-optic effect, in which the refractive index of silicon is varied by applying heat to the material. The thermo-optic coefficient in silicon is [16]:

$$\frac{dn}{dT} = 1.86 \times 10^{-4} / \text{K} \quad (4.41)$$

Therefore if the waveguide material can be raised in temperature by approximately  $6^\circ\text{C}$  in a controllable manner, a refractive index change of  $1.1 \times 10^{-3}$  results. There are of course issues about controlling the temperature rise to the locality of the waveguide, and of efficiency

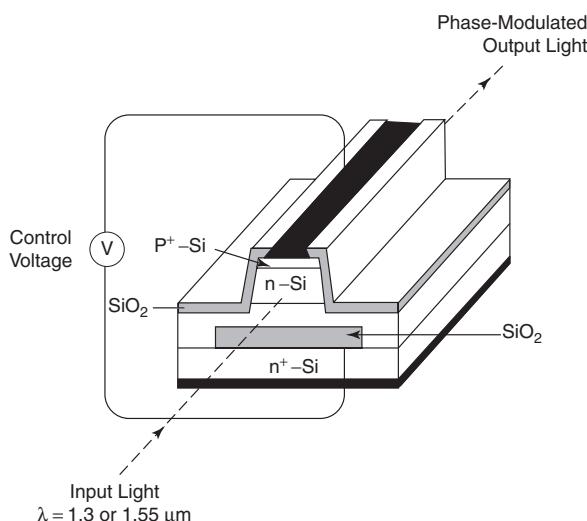
of the mechanism used to deliver the thermal energy. However, experimental results of [16] suggest that a 500  $\mu\text{m}$  device length can deliver a phase shift of  $\pi$  radians for an applied power of 10 mW, if the waveguide is thermally isolated from the substrate. This corresponds to a thermal change of approximately 7°C, and hence a refractive index change of approximately  $1.3 \times 10^{-3}$  over the length of the device.

It is worth noting that the refractive index change is positive with applied thermal energy, whereas the injection of free carriers results in a decrease in refractive index. Therefore the two effects could compete in a poorly designed modulator.

### 4.3 A Brief History of Silicon Modulator Designs

Silicon modulators are as old as the field of silicon photonics, being one of the first devices to be studied, after initial work on the basic waveguide technology showed that devices were viable [17–20]. This section reviews some of the devices that have been studied since the mid 1980s. Whilst it is impossible to provide comprehensive coverage of modulators throughout this period in a short chapter, a selection of interesting and/or important devices is discussed. The important parameters for an optical modulator are the modulation bandwidth, the modulation depth, and the power consumption, and where possible these parameters are included in the discussion for the devices considered.

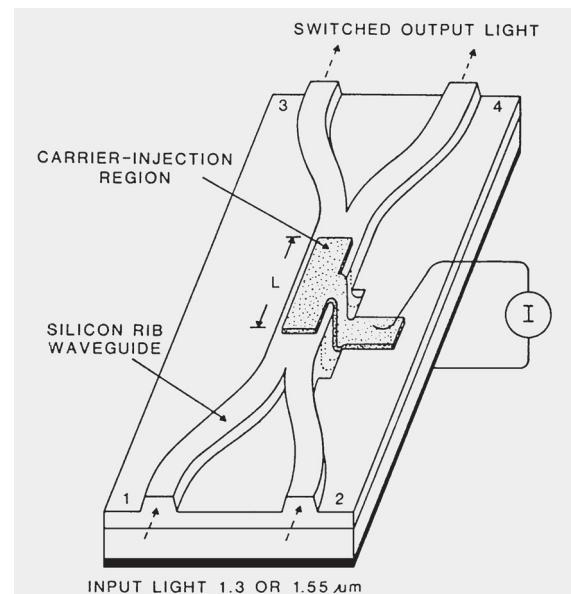
The earliest designs were based upon silicon guiding layers fabricated on doped silicon substrates (to form the lower waveguide boundary), the latter having a reduced refractive index via the plasma dispersion effect. Later, silicon-on-insulator (SOI) waveguiding structures became more popular due to the possibility of much stronger optical confinement. The first plasma dispersion modulator in silicon was proposed by Soref and Bennett [15]. This  $p^+-n^-n^+$  modulator, shown in Figure 4.5, was based on a single-mode silicon rib waveguide. It was



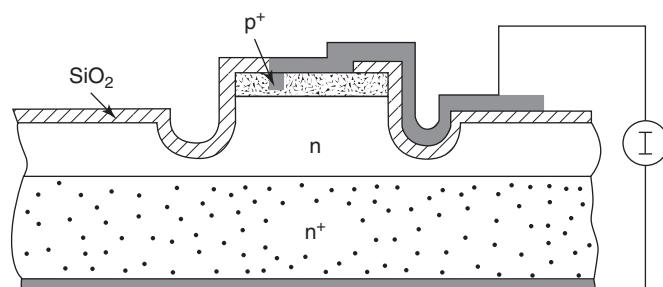
**Figure 4.5** Proposed SOI  $p^+-n^-n^+$  channel-waveguide electro-optical phase modulator. Reproduced from [15] by permission of SPIE

found by modelling that the interaction length of the modulator required for a  $\pi$  radian phase shift was less than 1 mm. The corresponding loss was less than 1 dB at  $\lambda = 1.3 \mu\text{m}$  for both orthogonal polarisation modes. The authors noted that to first approximation, the modulator was polarisation independent. It is interesting to note the buried ‘block’ of  $\text{SiO}_2$  below the waveguide, acting as the lower waveguide boundary. This is similar to the silicon-on-insulator structure used almost universally today in silicon photonics, but the retained contact to the buried  $n^+$  substrate allowed vertical current flow.

In 1987, Lorenzo *et al.* [21] reported the first  $2 \times 2$  electro-optical switch in silicon operating at a wavelength of  $1.3 \mu\text{m}$ , shown in Figure 4.6. The switch was fabricated as a vertical  $p^+ - n$  diode. The injection of holes into the intersection of the crossing waveguides caused a change in refractive index, resulting in some limited switching. For an input current density of

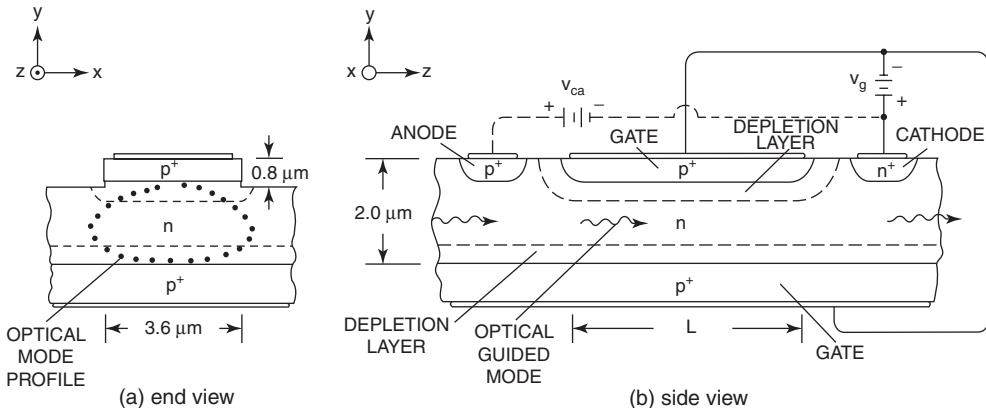


(a) Perspective view



(b) Active midsection

**Figure 4.6**  $2 \times 2$  guided-wave Si optical switch by injection of free carriers. Reproduced from [21] by permission of American Institute of Physics

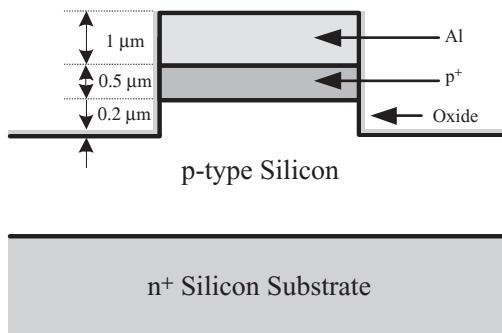


**Figure 4.7** Cross-section views of the proposed dual injection field effect transistors (DIFET). Reproduced from [22] by permission of American Institute of Physics

$J = 1.26 \text{ kA/cm}^2$  applied at the  $p^+$ – $n$  junction, the device experienced an ‘on’ state which switched 50% of the optical power from the output straight-through channel port 3 to the output cross channel port 4. Although this device was not optimised, it reaffirmed the feasibility of the plasma dispersion effect.

In the late 1980s, Friedman *et al.* [22, 23] proposed and theoretically analysed phase modulators in a series of transistor structures integrated into rib waveguides, as shown in Figure 4.7. The devices were based around MOSFETs, utilising injection of single carrier types (holes or electrons), or dual injection field effect transistors (DIFET). The injected charge in the DIFET was to be controlled by the junction field effect in which a voltage variable depletion width controlled the effective cross-sectional area of the conducting channel. The authors predicted effective refractive index changes of the order of  $1 \times 10^{-3}$ , for applied gate voltages of 10–20 V. They also proposed dual gate devices which offered better overlap of the modal field with injected carriers, although fabrication of two gates is impractical for potential SOI structures. They concluded that because the refractive index changes occurred over thin layers, the devices would be optimal in small waveguides, of the order of 0.1  $\mu\text{m}$  in height. At the time this was regarded as unreasonably small, but as the recent trend to smaller cross sectional dimensions continues, and nanophotonic devices are becoming more seriously considered, some of these early devices are being reconsidered. Some of the devices studied also share similarities with a recent Intel device (discussed later), the first modulator to demonstrate experimental operation with a bandwidth in excess of 1 GHz [24].

Treyz *et al.* [25] reported silicon waveguide intensity modulators operating in the range of 1.3 to 1.55  $\mu\text{m}$ . A schematic diagram is shown in Figure 4.8. The intrinsic region was grown to a thickness of 7.7  $\mu\text{m}$  and had a  $p$ -type doping concentration of less than  $5 \times 10^{15} \text{ cm}^{-3}$ . The  $p^+$  layer had a doping concentration greater than  $5 \times 10^{19} \text{ cm}^{-3}$  and a thickness of 0.5  $\mu\text{m}$ . By applying a forward bias to the device, free carriers were injected into the  $p$ -type silicon (guiding) region of the device. The maximum modulation depth achieved was –6.2 dB (76%) for a high current density of  $3.4 \times 10^3 \text{ A/cm}^2$ . The authors measured the 90–10% response time to be less than 50 ns. However, the devices were multimode. This device was further developed

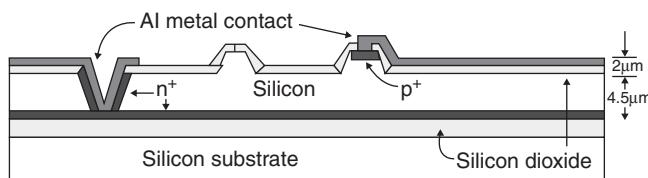


**Figure 4.8** Schematic diagram of the active region. Rib height was 4.6  $\mu\text{m}$  with widths of 13 and 24  $\mu\text{m}$ . Diode lengths were 500 and 1000  $\mu\text{m}$ . Reproduced from [25] by permission of IEEE © (1991)

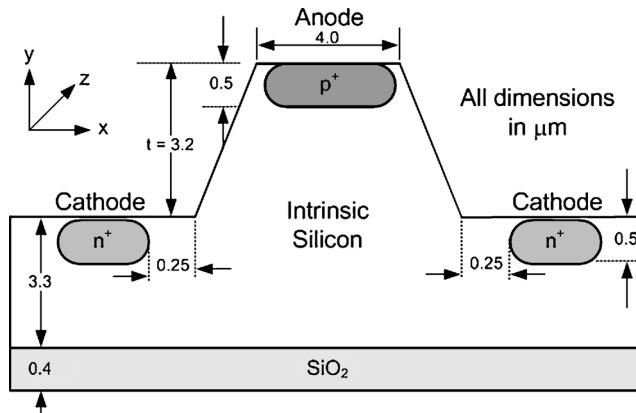
in another paper where the same authors demonstrated a similar device configured as a phase modulator [26]. The device was incorporated into one arm of a Mach–Zehnder Interferometer (MZI) to convert phase modulation into amplitude modulation. The devices also had response times of 50 ns with active region losses estimated to be 0.75–2.0 dB, compared with 0.4–1.0 dB [25] above. In common with the modulator of Lorenzo and Soret [21], this waveguide experienced a high loss due to a relatively weak optical confinement in the vertical direction and hence absorption of guided modes via the doped silicon substrate. The trend to fabrication of SOI-based structures eliminated this problem, but also meant that vertical structures were much more difficult to fabricate, as substrate contacts could not be included. However, Jackson *et al.* [27, 28] later produced a similar ‘vertical’ modulator based on SOI, utilizing a buried  $n^+$  contact at the bottom of the waveguide as shown in Figure 4.9. It is also interesting to note that the cross-sectional dimensions of this device are large, of the order of several micrometres. This was typical of modulation devices throughout the early 1990s, largely to ensure good coupling from telecommunications optical fibres.

In 1993, Fischer *et al.* [29] produced an MZI modulator with a  $p-i-n$  device by using a single-mode SiGe rib waveguide. At a wavelength of 1.3  $\mu\text{m}$ , the measured modulation depth was –10 dB with a current of 150 mA which was the lowest value reported at that time. However, the response time of their device was not reported.

Work carried out at the University of Surrey by Tang *et al.* [30] in 1994, in support of an earlier simulation paper in 1993 [31] showed that it was possible to obtain a 30% increase in the concentration of injected carriers into the waveguiding region of a phase modulator [30, 32] by changing the side wall angle of the rib from vertical to 54.7° (Figure 4.10).



**Figure 4.9** ‘Vertical’ modulator demonstrated experimentally by Jackson *et al.* Reproduced from [27, 28] by permission of IEEE © (1998)



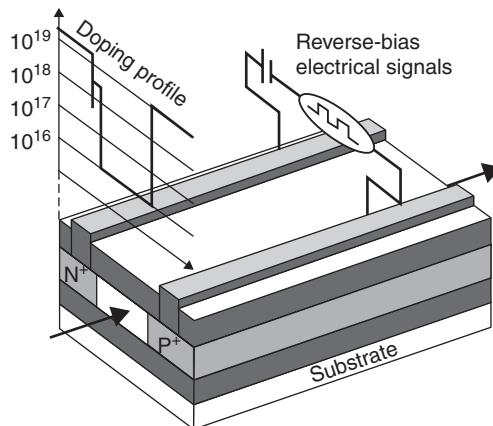
**Figure 4.10** Three-terminal phase modulator with angled rib walls. Reproduced from [30–32] by permission of IEEE © (1993)

The separation distance of the  $n^+$  injecting regions remained constant for both the vertical and angled rib waveguides. The device was designed at the University of Surrey [30] and fabricated in a University of Surrey/University of Southampton collaboration [32]. It was modelled using the two-dimensional (2D) semiconductor device simulator MEDICI. Table 4.1 shows the comparison between the predicted theoretical and experimental performance of two modulators with different cross-sectional areas [30, 32]. From Table 4.1, it can be seen that, despite the differences between the two modulators, the figure of merit is in reasonable agreement, and is a direct consequence of optimisation via the modelling exercise. At the time when these devices were fabricated, the typical current densities were of the order of  $\text{kA}/\text{cm}^2$  [21]. Therefore, with an experimental drive current of 7 mA, and current density of  $175 \text{ A}/\text{cm}^2$ , this device represented an improvement in the current density of approximately an order of magnitude. Modulation bandwidths were in the range 5–20 MHz for different device variants.

Huang *et al.* [33] proposed and modelled a novel guided-wave modulator in 1993 which utilised an IMPATT type diode, to generate carriers by avalanche multiplication (Figure 4.11). The authors argued that the modulation speed of the injection-type of device was limited by the minority-carrier lifetime, ( $\sim \mu\text{s}$ ), and a depletion-type device could only provide modulation over a sub-micrometer device size. Hence they based their design on avalanche generation. Their device utilised two regions. The first was a small and high-field avalanche region where

**Table 4.1** A comparison of theoretical and experimental performance [30, 32]

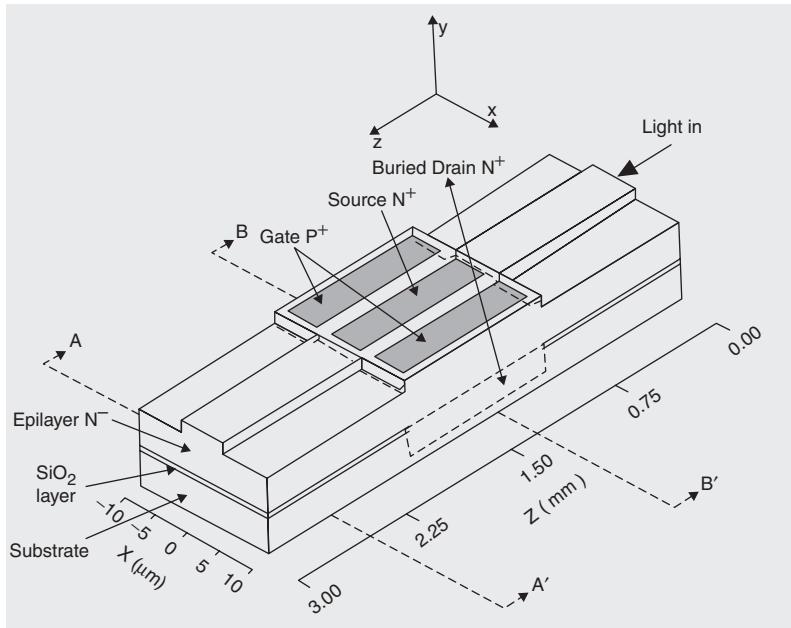
Parameter	Modelled device	Experimental device
Current density $J$ ( $\text{A}/\text{cm}^2$ )	120	175
Active optical loss (dB)	1.37	1.3
Current required for a $\pi$ -radian phase shift $I_\pi$ (mA)	4	7
Refractive index change $\Delta n$	$-9.24 \times 10^{-4}$	$-1.55 \times 10^{-3}$
Peak carrier concentration $N$ ( $\text{cm}^{-3}$ )	$2.5 \times 10^{17}$	$4.6 \times 10^{17}$
Figure of merit $\chi$ ( $^\circ \text{ cm}/\text{A}$ )	17.9	20.6



**Figure 4.11** Schematic diagram and doping profile of the SOI  $n^+-p^+-p^-p^+$  modulator. Reproduced from [33] by permission of American Institute of Physics

impact ionization occurs and carriers were generated. The second was a long and low-field drift region where generated carriers cause the refractive index to be modulated. Modulation speed is improved as the drifting carriers are majority carriers. The proposed modulator structure of  $n^+-p^+-p^-p^+$  based on SOI is reproduced above [33]. The simulated doping concentrations for the  $n^+$  and  $p^+$  claddings were  $1.5 \times 10^{19}$  and  $10^{19}$  cm<sup>-3</sup>, respectively; while the avalanche  $p^+$  region was  $3 \times 10^{18}$  cm<sup>-3</sup>. The authors modelled their devices using the SUPREM-IV process simulator [34] and found the switching time of their proposed device to be less than 1 ns, which implies a modulation bandwidth in the gigahertz regime. However, the predicted current density was two orders of magnitude higher than an injection device ( $\sim 110$  kA/cm<sup>2</sup>), and very lossy as lateral confinement was achieved via doped regions, although presumably an etched rib is also viable.

By the mid 1990s, silicon modulator work was almost exclusively being carried out on SOI material. However, in 1997, Cutolo *et al.* [35] proposed a silicon modulator based on a three-terminal device integrated in a low-loss single-mode SOI waveguide, but with the buried oxide layer removed in the active region of the modulator (Figure 4.12). The device was analysed theoretically and according to the authors, the three terminal electronic structure was significantly more efficient than two terminal  $p-i-n$  diode based modulators, a similar conclusion to that drawn previously by Tang *et al.* [30, 32]. Simulation results were obtained using the MEDICI 2D semiconductor device simulator [34]. An amplitude modulation of 20% was reported, together with an injection power of about 126 mW, and a switching time of 5.6 ns. Furthermore, when the device was operated as a phase modulator, it exhibited a very high figure of merit. The predicted performance was approximately 215°/V/mm, for a driving power of 43 mW and a switching time of less than 3.5 ns (143 MHz). At that time this was a very promising prediction for a modulator based upon a large rib waveguide. The authors attributed the switching performance to the more complex electrical structure, in comparison with classical two-terminal  $p-i-n$  diode modulators. The proposed modulator had dimensions of 6, 6, and 1.5 μm for the rib width, Si layer thickness, and etch depth respectively. The plasma was injected into the optical channel by applying a forward voltage to the source and drain. The free carriers in the active region were then controlled by means of an applied voltage to the

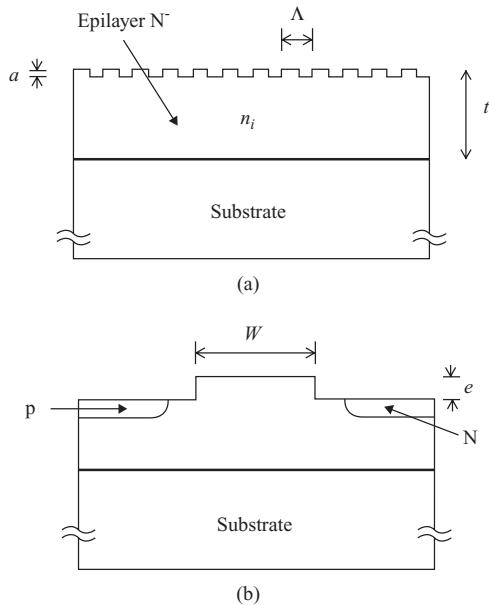


**Figure 4.12** Three-terminal modulator proposed by Cutolo *et al.* Reproduced from [35] by permission of IEEE © (1997)

control terminal. Hence the carriers can be displaced away from the optical mode rather than be depleted from the active region. In this way, the total charge present in the channel is fixed, so achieving shorter switching time as free carrier drift is faster than the injection. The active region of the waveguide would, if fabricated, exhibit significantly larger propagation losses than a purely SOI waveguide, due to additional absorption in the buried Drain. For the TE<sub>0</sub> mode at  $\lambda = 1.55 \mu\text{m}$ , overall active attenuation was predicted to be 11.8 dB/cm (12.2 dB/cm for TM<sub>0</sub>). Despite the high loss, this was one example that demonstrated that designs were emerging to increase device speed.

Cutolo *et al.* [36, 37] also proposed and analysed a lateral *p-i-n* diode combined with a Bragg reflector (Figure 4.13). Numerical simulations were made using MEDICI and showed a 50% modulation depth for a power of 4 mW. The response time was 12 ns ( $\sim 30 \text{ MHz}$ ) and the device exhibited an insertion loss of 1 dB. The device worked by varying the refractive index of the guiding layer, which in turn modulates the transmission characteristic of the Bragg mirror and hence modulates the intensity of light passing through the device. The Bragg grating had a period of  $\Lambda = 227 \text{ nm}$ , a depth  $a = 45 \text{ nm}$  and a width  $W = 3 \mu\text{m}$ . The modulator has an interaction length of 3200  $\mu\text{m}$  with  $n^+$  and  $p^+$  dopant depths of 0.85  $\mu\text{m}$ . Performance of a similar Mach-Zehnder interferometer (MZI) based device was similar in terms of driving signals and response times, but the MZI required about three times more power to achieve the same modulation depth. This is an example of how the carrier injection technique can be applied to resonant structures, as well as causing a direct attenuation or a phase shift.

Breglio *et al.* [38] measured a 20% modulation depth with a power of 500 mW for the FET device proposed by Cutolo *et al.* [36] in Figure 4.13. This is nearly four times higher

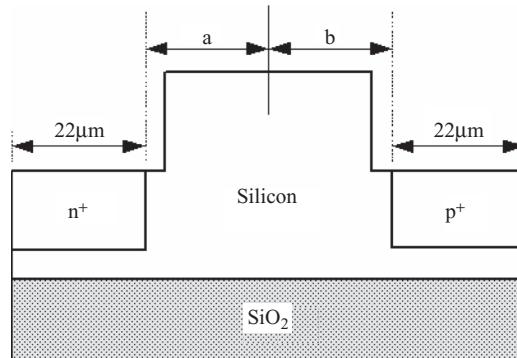


**Figure 4.13** (a) Longitudinal; (b) transverse sections of proposed modulator. Reproduced from [36] by permission of American Institute of Physics

than predicted (126 mW) and is attributed by the authors to a possible contact resistance in the fabricated devices. Scuito *et al.* [39] also reported findings for a similar device, albeit with larger dimensions of 10, 10, 9.25  $\mu\text{m}$  for the rib width, Si layer thickness, and etch depth, respectively. They demonstrated a modulation depth of 75% using a gate current of 10 mA (160 mW) at a wavelength of 1.48  $\mu\text{m}$ . The ‘off’ and ‘on’ drain voltages were 0 V and 16 V, respectively. However, neither Breglio *et al.* [38] nor Scuito *et al.* [39] reported transient analysis. In 1999, Hewitt *et al.* [40] also proposed through simulation, a multi-micrometre three-terminal optical *p-i-n* modulator based on SOI which was also similar to the one proposed by Cutolo *et al.* [35] and earlier by Tang *et al.* [32]. The modulators were based upon a transverse *p-i-n* structures and utilised the plasma dispersion effect to produce the desired refractive index change in an optical rib waveguide. The authors reported a three terminal device based on a doping contact topology of  $n^+p^+p^+$  requiring a drive current of 2.8 mA and a corresponding current density of 112 A/cm<sup>2</sup> to achieve a  $\pi$ -radian phase shift, the lowest predicted at that time for a large rib structure ( $\sim 5 \mu\text{m}$  in cross-sectional dimensions). The rise time was predicted to be 29 ns (12 MHz). However, the devices were not fabricated, as multiple fabrication steps would be required to achieve the  $n^+p^+p^+$  contact topology.

In 2000, Hewitt *et al.* [41] used computer simulation to reconsider a simple two terminal *p-i-n* modulator based on a 5.5  $\mu\text{m}$  SOI rib waveguide. It was predicted that even for a two terminal device, significant optimisation is possible. For example, an increase in the doping concentrations of the  $p^+$  and  $n^+$  regions, from  $10^{19}$  to  $10^{20} \text{ cm}^{-3}$ , results in a drive current decrease from 63 to 8 mA while the transient rise time also decreases from 110 to 105 ns.

At the same time, the placement of the doping windows (Figure 4.14) was also found to improve/degrade device transient characteristics. Rise time was reduced from 184 ns when



**Figure 4.14** Varying the placement of the contact windows. Reproduced from [41] by permission of IEEE © (2000)

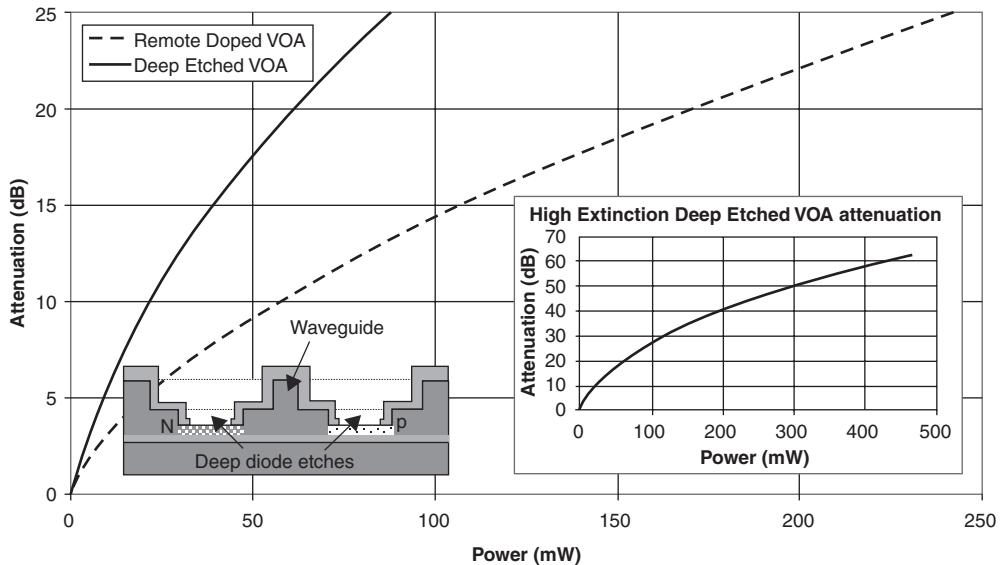
$a = b = 7\text{ }\mu\text{m}$  (see Figure 4.14), to 39 ns when  $a = b = 3\text{ }\mu\text{m}$  [41]. Although the study was conducted on a specific device geometry, it was argued by the authors that the trends observed in the results, would have applicability in other device configurations.

From the work of Hewitt and others [40, 41], it is clear that three-terminal devices require less drive current (2.8 vs 8 mA) and are faster than two-terminal devices(29 vs 39 ns), for an equivalent injection concentration. This is because, three-terminal devices offer more efficient carrier injection. The potential drawback of these devices (three-terminal) is that additional optical attenuation may occur due to the doping contact at the rib top.

For a silicon  $p-i-n$  phase modulator utilising free carrier injection, carrier confinement in the active region is important. Trench isolation was proposed as a means of achieving carrier confinement, by Hewitt *et al.* [42], who predicted that employing lateral trench isolation on either side of the modulator improves both the directcurrent (dc) and transient devices performances by up to 74% and 18% respectively. This is because the lateral carrier movement, which does not contribute to the refractive index change in the active region of the modulator, is ‘blocked’ by the oxide trenches. Another benefit brought by implementing isolation trenches is that the carrier confinement permits dense integration due to electrical isolation between neighbouring devices.

Another method of carrier confinement was demonstrated by Day *et al.* [43, 44]. The authors argued that in order to achieve highly efficient operation, it was necessary to ensure that all injected carriers in the silicon remain in the region between the contact dopants. Rather than locating the doped regions in the surface of the slab region on either side of the waveguide, shown in Figure 4.15, the doped regions are instead placed at the base of the etched regions. This enabled the doped regions to extend to the buried oxide trapping injected carriers in the region of the waveguide, hence increasing device efficiency. This approach is similar in concept to that of Hewitt *et al.* [42], and the conclusions are similar. They reported that the power needed to obtain an optical attenuation of 25 dB was reduced from 240 to 85 mW, and a switching time of  $<300\text{ ns}$  (1.2 MHz). Their waveguide height was approximately  $4\text{ }\mu\text{m}$ .

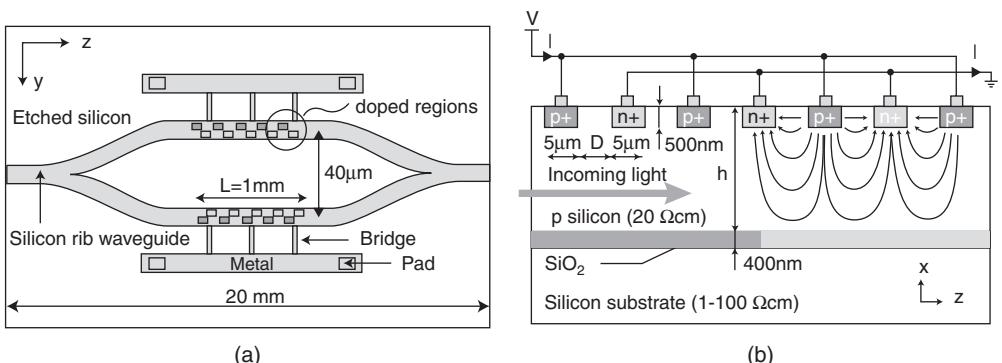
In 2003, Whiteman *et al.* [45] improved the device of Day *et al.* [44] by reducing the waveguide dimensions. They argued that by reducing the waveguide height, the physical extent of the optical mode is reduced, hence allowing the doping contact regions to be closer before an overlap with the optical mode occurs. In other words, with a smaller device, the carriers



**Figure 4.15** Improvement in VOA efficiency with introduction of deep diode etch over standard remote-doped structure. Right inset shows 40 dB attenuation achieved with 200 mW of power. Reproduced from [43, 44] by permission of IEEE © (2003)

take a shorter time to be injected and depleted from the optical region. These devices exhibited 25 dB dynamic attenuation for less than 50 mW with a bandwidth of 10 MHz.

In 2000 Dainesi *et al.* [46] realised a SOI Mach-Zehnder Interferometer (MZI) operating at 10 MHz bandwidth. The MZI device is shown in Figure 4.16(a). The unique feature about this modulator is that the contact regions are arranged in a longitudinal fashion in the  $z$ -direction, as opposed to the lateral approach [47]. The authors suggested that having the diodes, doped regions and contact on the same plane, allows full compatibility with CMOS technology and



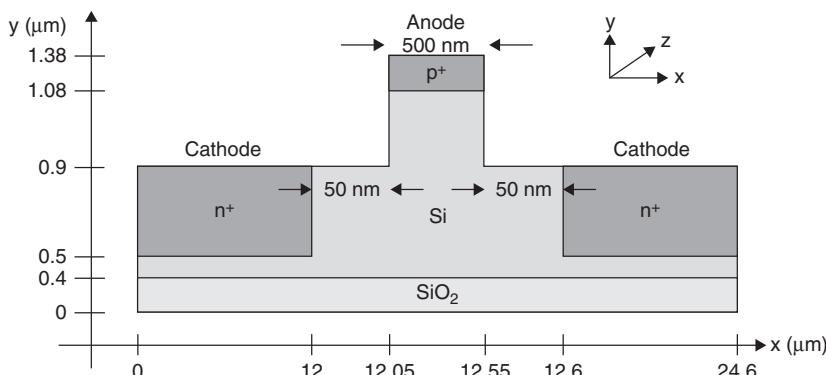
**Figure 4.16** (a) Top view of the MZI modulator; (b) cross-sectional view along the phase modulator of Dainesi *et al.* Reproduced from [46] by permission of IEEE © (2000)

has more uniform charge injection due to the longitudinal layout of the doped regions. The drawback of such a configuration is that parallel diode pairs are created instead of a single lateral diode all along the waveguide structure, hence giving rise to a bigger drive current (136 mA).

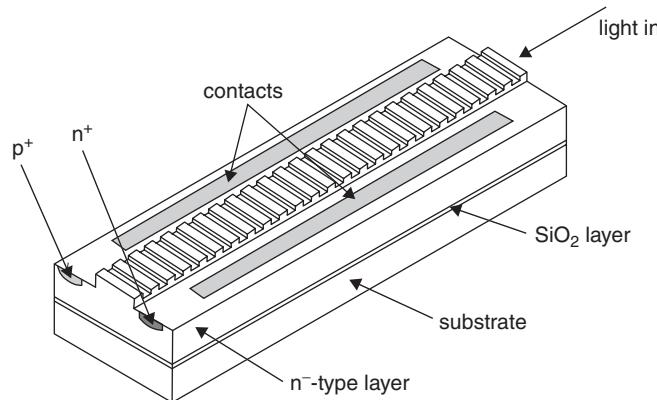
The reduction in device dimensions was a popular means of improving device performance. The work of Ang *et al.* [47] proved to be significant when in 1999, *p-i-n* modulator devices were proposed and modelled to operate above 500 MHz. These were similar in design to previous devices of Hewitt *et al.* [40, 41] and Tang *et al.* [30, 32], but smaller. The modulators with interaction lengths of 500  $\mu\text{m}$  were based upon transverse *p-i-n* structures in an optical rib waveguide with a silicon thickness of 0.98  $\mu\text{m}$ . In that paper, the authors reported a theoretical *npp* device, with doping concentrations of  $10^{20} \text{ cm}^{-3}$ , which required a drive current of just 0.29 mA to achieve a  $\pi$ -radian phase shift, the lowest reported then. The corresponding current density was 112 A/cm<sup>2</sup>. The response time was modelled to be 0.81 ns ( $\sim$ 430 MHz). Those findings represented a significant improvement in device performance from those previously reported in the existing literature. Indeed, at the time of this work [48], more and more researchers were recognising the benefits of reducing the silicon overlayer thickness, not only in terms of modulator performance, but optical circuit performance in general.

Png *et al.* [49], later improved upon the work of Ang *et al.* [48] by modelling devices of similar geometry (see Figure 4.17), but with improved performance [50–52]. In particular, a series of devices were modelled with bandwidths ranging from 70 MHz to in excess of 1 GHz. The devices were based around a rib waveguide, approximately 1  $\mu\text{m}$  in height and between 0.5 and 0.75  $\mu\text{m}$  wide. A feature of these devices was the optimised doping profile in the *n<sup>+</sup>* regions to optimise injection efficiency. Png *et al.* [49] also reported the technique of pre-emphasis on critical device rise and fall times to increase device speed, improving a device based on Figure 4.17 from 95 MHz to 5.8 GHz. Using such a scheme, a class of devices with nominal operating speeds of 1 GHz could theoretically be switched in excess of 40 GHz [53].

In (2003), Irace *et al.* [37] modelled a 1.4 GHz operating bandwidth for a two-terminal Bragg reflector rib waveguide (Figure 4.18). This is essentially a follow up work of Cutolo *et al.* [36], in a smaller waveguide. The device had a waveguide height and rib width of 1  $\mu\text{m}$ , etch depth of 100 nm and an interaction length of 3000  $\mu\text{m}$ . These dimensions are a reduction



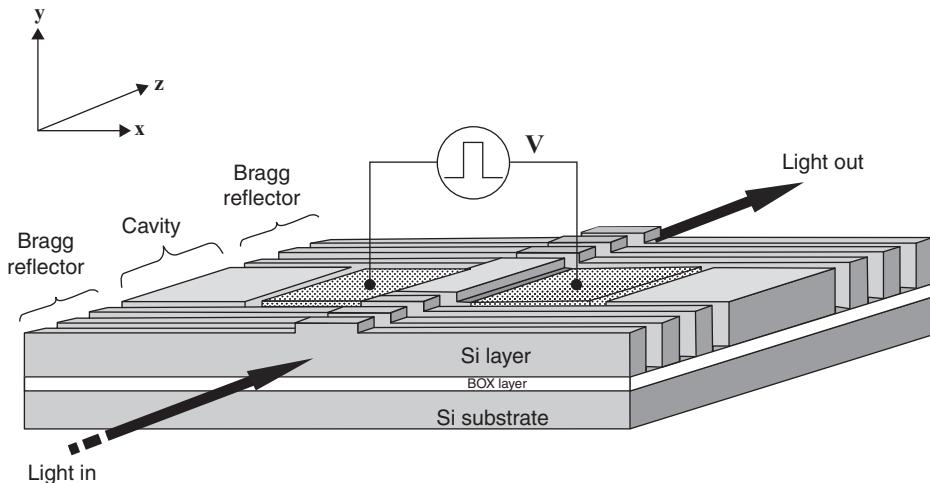
**Figure 4.17** Proposed three-terminal rib waveguide device based on SOI. Reproduced from [49] by permission of IEEE © (2004)



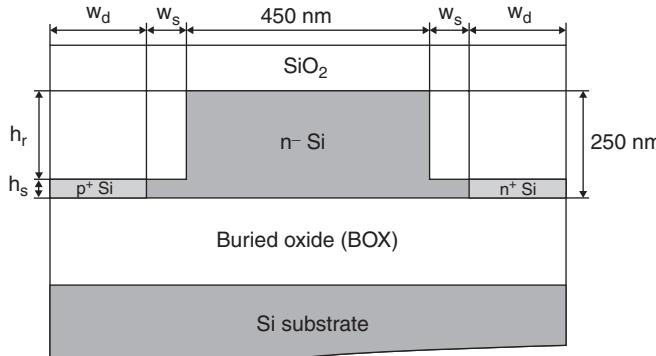
**Figure 4.18** Schematic of the proposed 1 GHz Bragg modulator. Reproduced from [37] by permission of IEEE © (2003)

from the original waveguide height and rib width of 3  $\mu\text{m}$ , etch depth of 45 nm, whilst the interaction length remains comparable (3200  $\mu\text{m}$ ). Thus it can be seen that by reducing the silicon overlayer thickness and rib width, the device bandwidth increased from MHz operation to the GHz regime. Irace *et al.* [37] attributed this improvement in bandwidth to two factors: decreasing and optimising the device dimensions, and applying a pre-bias to the modulator to an ‘off’ level (0.6 V) just below the turn-on voltage (0.8 V), which allowed faster movement of the injected carriers.

Longer devices hinder high integration levels and hence increase cost. Consequently, Barrios *et al.* [54] modelled a modulator based on a Fabry–Perot microcavity with Bragg reflectors as shown in Figure 4.19. The microcavity facilitates confinement of the optical field in a small



**Figure 4.19** Fabry–Perot microcavity with high-reflectivity Bragg reflectors in SOI waveguide. Reproduced from [54] by permission of IEEE © (2003)



**Figure 4.20** Proposed rib waveguide with an integrated lateral  $p-i-n$  diode. Reproduced from [54] by permission of IEEE © (2003)

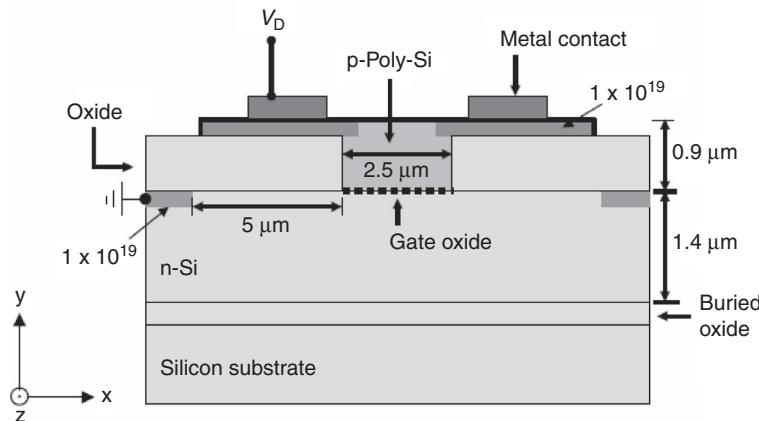
region, and the transmission of the device near its resonance is highly sensitive to small index changes in the cavity. Thus the device required a low concentration of injected carriers to switch to a nonresonant position. The rib width and silicon thickness were 1.5  $\mu\text{m}$ , and the etch depth was 0.45  $\mu\text{m}$ . The 20- $\mu\text{m}$ -long device was predicted to require a dc power of the order of 25  $\mu\text{W}$  at an operating wavelength of 1.55  $\mu\text{m}$ , to achieve 31 MHz operating bandwidth with transmittance of 86% and a modulation depth of 80%. The authors also noted the merits of using trench isolation reported by Hewitt *et al.* [42] and implemented this feature in their proposed modulator device [54].

The same authors [54] proposed another modulator that was only 250 nm in height. The device, shown in Figure 4.20, consisted of two uniformly doped contacts,  $p^+$  and  $n^+$  at a level of  $10^{19} \text{ cm}^{-3}$  and a distance of 200 nm from the rib edge. The device slab thickness was 50 nm. The authors predicted that in order to induce a refractive index change of  $-10^{-3}$ , a dc power of 1.53  $\mu\text{W}/\mu\text{m}$  was required. The switching time for this device was estimated to be 1.29 ns. The dc power is one of the lowest predicted to date, and this device reinforces the merit of reducing device dimensions.

In 2004, researchers from the Intel Corporation experimentally demonstrated a silicon-based optical modulator with a bandwidth that exceeds 1 GHz [55] for the first time. This was a major milestone in silicon photonics and attracted huge media attention.

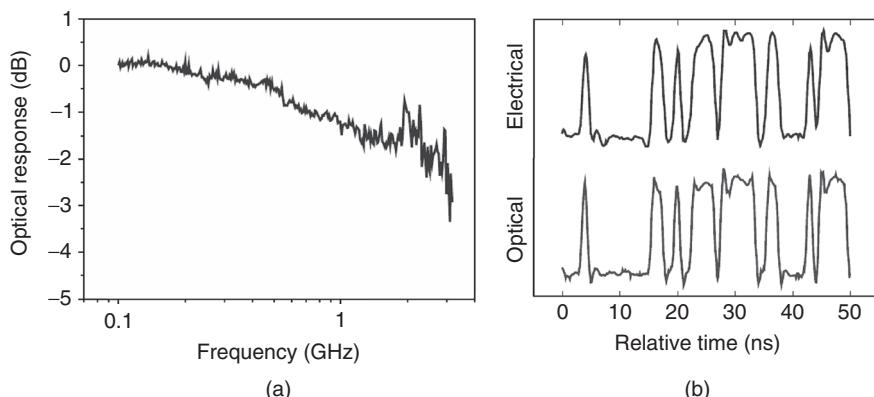
Figure 4.21 shows a schematic of the reported device. This device operates by the free carrier effect and bears a close resemblance to a complimentary metal–oxide–semiconductor (CMOS) transistor. The device structure consists of  $n$ -type crystalline silicon with an upper ‘rib’ of  $p$ -type polysilicon. The  $n$ -type and  $p$ -type regions are separated by a thin insulating oxide layer. Upon application of a positive voltage to the  $p$ -type polysilicon, charge carriers accumulate at the oxide interface, changing the refractive index distribution in the device. This in turn induces a phase shift in the optical wave propagating through the device.

The bandwidth of the device (a single 2.5-mm-long phase modulator) was characterised in two ways in an integrated asymmetric Mach–Zehnder Interferometer (MZI). The first technique was to drive the device with a 0.18 V<sub>rms</sub> sinusoidal source at the wavelength of 1.558  $\mu\text{m}$ , using lensed fibres for coupling into and out of the device. Figure 4.22(a) shows the normalised optical response of the MZI as a function of frequency (photoreceiver output voltage divided

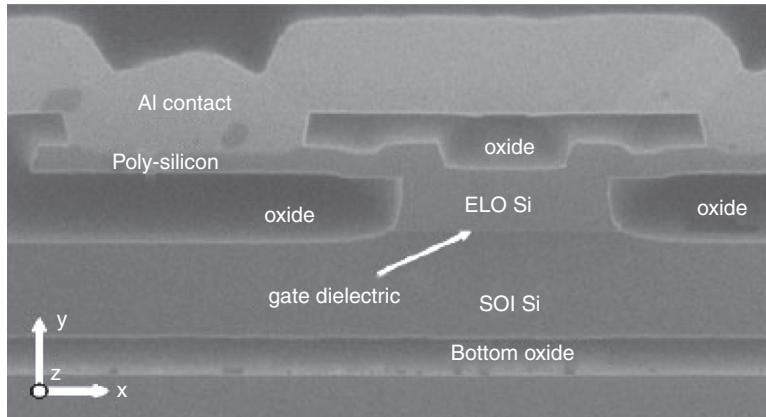


**Figure 4.21** Schematic diagram of the silicon-based optical modulator demonstrated experimentally to exceed 1 GHz bandwidth fabricated using standard CMOS processing techniques [55]

by on-chip drive voltage). Clearly, the 3 dB bandwidth exceeds 1 GHz. The second test was the application of a 3.5 V digital pulse pattern with a dc bias of 3 V. A 1 Gbit/s pseudorandom bit sequence was applied to the device and a high-bandwidth photoreceiver was used for detecting the transmitted optical signal. Figure 4.22(b) shows the optical signal faithfully reproducing the 1 Gbit/s electrical data stream. However, the on-chip loss for this device was rather high, at  $\sim 6.7$  dB, and the device was also highly polarisation dependent due to the horizontal gate oxide. Phase modulation efficiency for TE polarisation was larger than TM polarisation by a factor of 7.



**Figure 4.22** (a) Optical response of the MZI modulator (Figure 4.21) as a function of frequency; (b) pseudorandom bit sequence of a silicon MZI containing a single 2.5-mm-long MOS capacitor device (Figure 4.21) in one arm at a data bit rate of 1 Gb/s. Reproduced from [55] by permission of Macmillan Publishers Ltd.

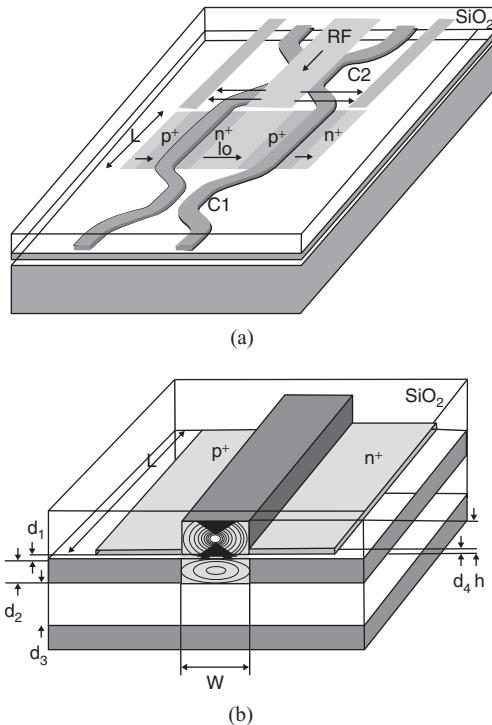


**Figure 4.23** SEM cross-section of the MOS modulator [62]

The authors also suggested several methods that may improve the device performance even further. The first is replacing the *p*-type polysilicon with single crystal silicon, where the latter was expected to reduce on-chip loss by  $\sim 5$  dB. Another suggestion is the reduction of the device dimensions as the capacitance is reduced through such a shrinkage. They also suggested using a graded doping profile in the vertical direction such that higher doping densities exist in the areas close to the gate oxide and lower doping concentrations in the rest of the waveguide. Furthermore, the authors also reported that their modelling predicts that the device can be scaled to operate at 10 GHz, offering very significant improvement of silicon photonic devices.

A year later, Liao *et al.* [62] reported an improved version of the MOS optical modulator. Figure 4.23 shows a scanning electron microscope (SEM) cross-sectional image of the phase shifter. This modulator is smaller than the previous one, and comprises a  $1.0\ \mu\text{m}$  *n*-type doped crystalline Si (the Si layer of the SOI wafer) on the bottom and a  $0.55\ \mu\text{m}$  *p*-type doped crystalline Si on the top, with a  $10.5\ \text{nm}$  gate dielectric, a multi-layer stack of silicon dioxide and nitride, sandwiched between them. In the first version of the device, the waveguide cross-section was  $2.5 \times 2.3\ \mu\text{m}$  and the top Si layer was poly-silicon (poly-Si), which is significantly more lossy than crystalline Si due to defects and grain boundaries. In the improved device the poly-silicon was replaced by crystalline silicon via epitaxial lateral overgrowth (ELO), and the doping concentration was higher. In this smaller version of the phase shifter the mode–charge interaction is much stronger, which according to the authors, improved the  $V_p L_p$  coefficient by 50%. The authors reported data at 10 Gb/s for the modulator with an extinction ratio (ER) of 3.8 dB. Data transmission measurements suggested bandwidth ranging from 6 GHz (ER of 4.5 dB) to 10 GHz. The authors explained that the 6 GHz limitation is due to the driver design and wire bonding, which decreased the cut off frequency.

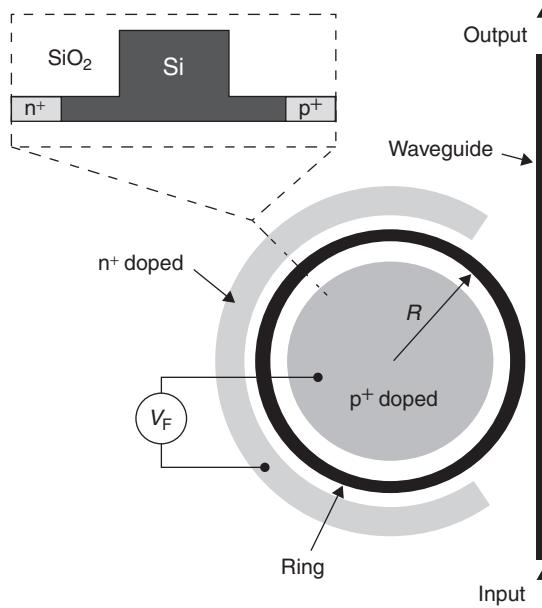
Taking the same interferometer principle as used for Intel's device, Gan and Kartner [56] proposed a Mach–Zehnder interferometer based on a Si–SiO<sub>2</sub> high-index contrast waveguide modulator. This device is based on a split-ridge waveguide that operates under forward-biased conditions and shows corner frequencies of up to 24 GHz. Figure 4.24(a) shows the RF electrodes of the device and Figure 4.24(b) shows the actual device structure where, to obtain single mode operation, the dimensions are  $d_1 = 100\ \text{nm}$ ,  $d_2 = 350\ \text{nm}$ ,  $d_3 = 2\ \mu\text{m}$ ,



**Figure 4.24** (a) Externally biased MZ modulator with coplanar RF feeder; (b) split-ridge waveguide pin-phase modulator with single-mode intensity profile. Reproduced from [56] by permission of IEEE

$d_4 = 100$  nm,  $h = 550$  nm,  $w = 1 \mu\text{m}$ . The pin junction is forward-biased with a dc voltage of 2 V and modulated with a voltage swing of 1 V. By varying the recombination lifetime inside the waveguide, simulations presented in this paper show that, with a carrier lifetime in the intrinsic region, of the order of ps, which could be set by lifetime doping or ion bombardment and operation in the saturation region where carriers travel at saturation velocity close to  $v_s \sim 10^7 \text{ cm/s}$ , a phase shifter with a modulation frequency up to 24 GHz should be possible.

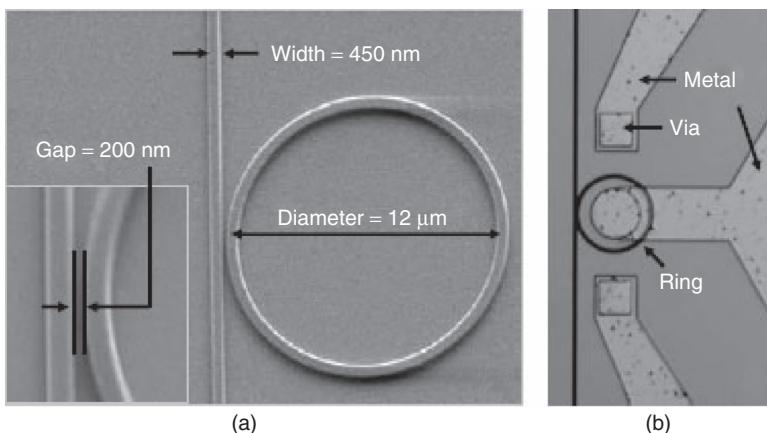
As the trend towards minimizing the real estate of devices continued, Xu *et al.* [57] reported a ring resonator using the waveguide structure proposed by Barrios *et al.* [54]. The diameter of the device was 12  $\mu\text{m}$  which, according to the authors, was at this time three orders of magnitude smaller than previously demonstrated. To use the ring resonator as a modulator it is necessary to operate at a single wavelength. Typically, ring resonator modulators are operated between regions of high throughput and low throughput in the resonator response. One way to achieve such switching is via changes in effective index in the ring, either changed through the plasma dispersion effect or thermally. Figure 4.25 shows a schematic of the device, where the waveguide structure and electrical structure is the one proposed by Barrios *et al.* [54] and is based on carrier injection. In [57] the authors reported a drive voltage of 0.3 V for dc modulation. During modulation, the ring resonator modulator is operated with a peak-to-peak voltage of 3.3 V and shows a modulation depth at 0.4 Gb/s as the extinction ratio at resonant



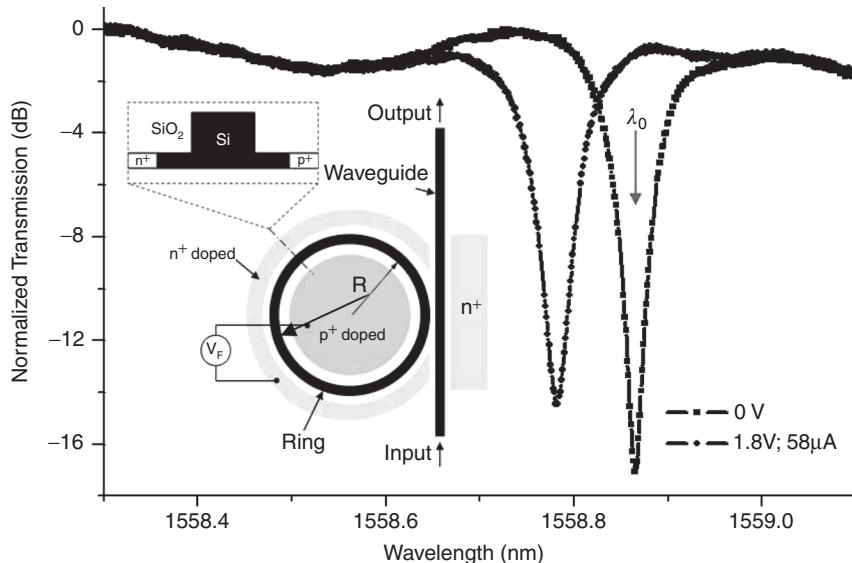
**Figure 4.25** Schematic of the ring resonator structure. Reproduced from [57] by permission of MacMillan Publishers Ltd.

frequency is above 15 dB. The authors also reported a data rate of 1.5 Gb/s when the ring resonator is operated with a peak to peak voltage of 6.9 V. Figure 4.26(a) shows an SEM image of the ring resonator. The waveguide is a rib structure, where the waveguide width is 450 nm and the separation in the coupling region is 200 nm.

In early 2007, Xu *et al.* [58] demonstrated an improvement of the previous device. For this device, shown on Figure 4.27, the 5  $\mu\text{m}$  ring is formed by silicon near-strip waveguides with

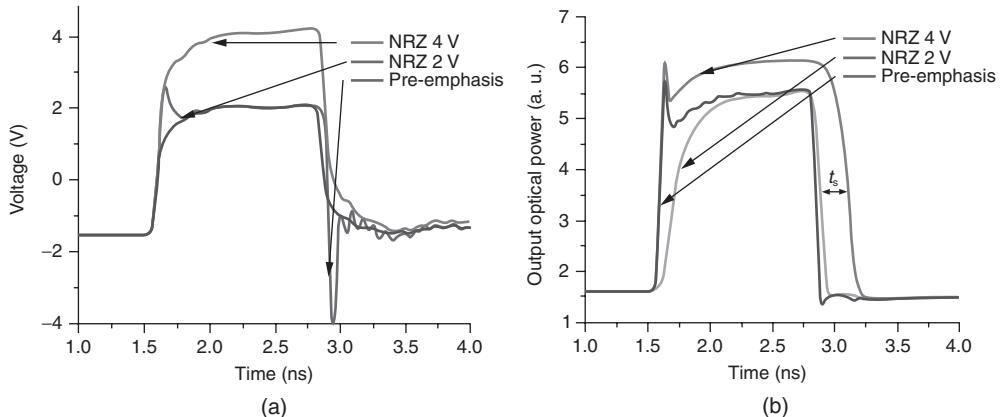


**Figure 4.26** (a) SEM of the ring resonator; (b) top view microscope image of the ring resonator. Reproduced from [57] by permission of MacMillan Publishers Ltd.

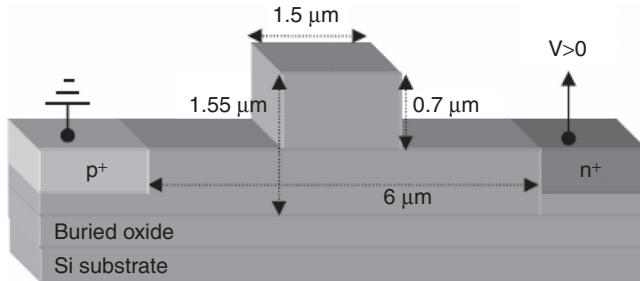


**Figure 4.27** Schematic of the ring resonator structure and normalised transmission spectra of the modulator proposed in. Reproduced from [58] by permission of the Optical Society of America

a height of 200 nm and the width of 450 nm on top of a 50-nm-thick slab layer. The distance between the ring and the straight waveguide is around 200 nm. Furthermore, compared with the previous an additional  $n^+$  doped region is added outside of the straight waveguide to form a nearly closed loop  $p-i-n$  junction, and the distance between the doped regions and the edge of the ring resonators and straight waveguides is reduced to  $\sim 300$  nm. Using a pre-emphasis (shown in Figure 4.28a,b) NRZ signal, they demonstrated the possibility of



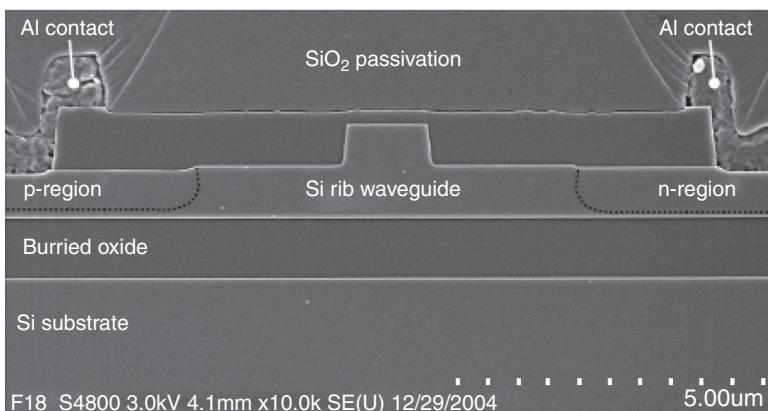
**Figure 4.28** (a) Square-wave driving signals with pre-emphasis and without (NRZ 2-4V) the pre-emphasis; (b) the output optical power when the modulator is driven by voltage signals shown in (a). Reproduced from [58] by permission of the Optical Society of America



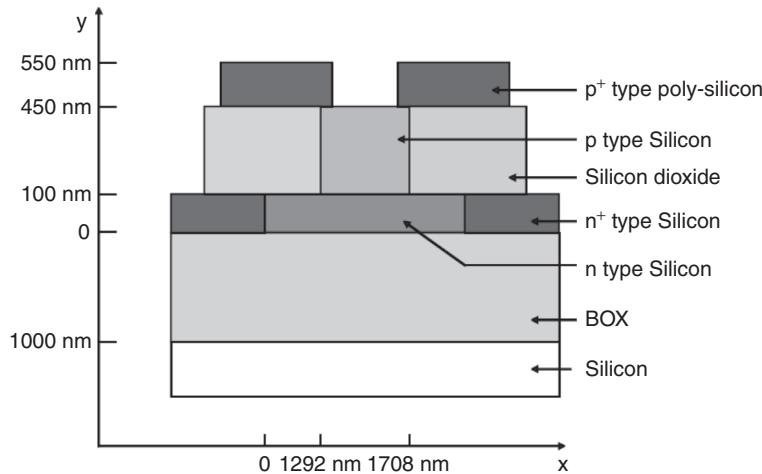
**Figure 4.29** Schematic of the Raman device, showing the electrical structure and dimensions of the waveguide [60]

decreasing further the rise time and fall time of the ring modulator, hence increasing the bandwidth of the modulator to 12.5 Gb/s with an extinction ratio of the signal of around 9 dB. This follows the same principle proposed previously by Png *et al.* [53].

Following a new approach and based on the Intel's silicon Raman laser [59], Jones *et al.* [60] reported a modulator using the Raman effect, in which they demonstrated lossless optical modulation in a silicon waveguide. The device is shown in Figure 4.29, as a schematic and in Figure 4.30, as a front-view SEM picture. To achieve net Raman gain in a silicon waveguide, high pump intensities are required and two-photon-initiated free carriers have to be removed from the intrinsic region of the waveguide. To achieve continuous wave (CW) Raman gain a reverse biased  $p-i-n$  device is embedded in a silicon waveguide to sweep out the two-photon induced free carriers. The electrical structure of the proposed device illustrated Figure 4.29, which shows the position of the  $p-i-n$  diode inside the waveguide. The doping concentration for the  $p$ -type and  $n$ -type region is  $1 \times 10^{20} \text{ cm}^{-3}$  and to increase the interaction length, hence achieving a larger total Raman gain, the waveguide formed an S-shaped curve. The total length of the waveguide was then 4.8 cm with a bend radius of 400  $\mu\text{m}$ . In order to characterise the DC



**Figure 4.30** SEM image of the waveguide and the  $p-i-n$  electrical structure used in the experiment [60]

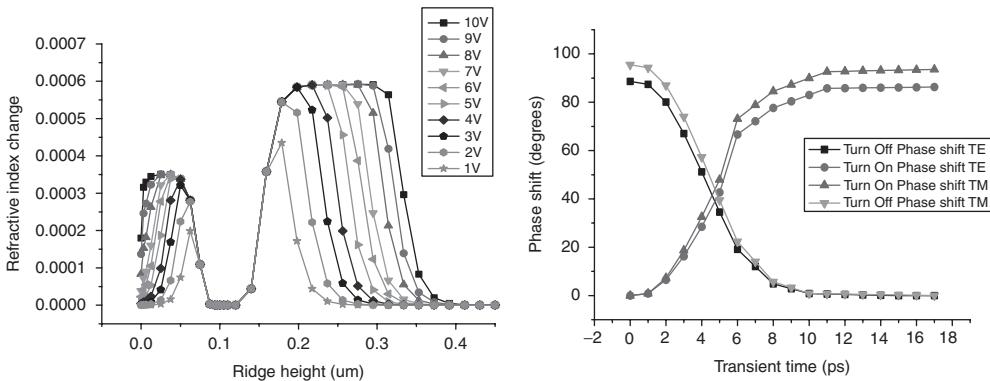


**Figure 4.31** Schematic of a four-terminal depletion-type modulator. Reproduced from [61] by permission of the Optical Society of America

gain measurements, a CW pump laser emitting at around 1548.3 nm was amplified using two EDFA s to a maximum output power of 3 W. The probe laser situated at the Stokes wavelength of 1684 nm was a 2 mW external cavity tunable diode laser. For this device the authors reported a net gain of 3.5 dB for a power pump of 945 mW and showed that a modulation of 5 dB of the net gain should be achievable with a 0–8 V reverse-bias voltage swing. A maximum bandwidth of 80 MHz was reported for a reverse-bias of 0.5–3.5 V swing, with a pump power of 283 mW where the modulation depth was 0.3 dB. The authors stated that the modulation depth could be improved by using the index change due to the free carriers in an interferometer such as a MZI and offered the possibility of a lossless planar waveguide switch or modulator based in silicon.

In order to increase the bandwidth further, a sub-micrometre modulator based on the depletion of a *p*–*n* junction was proposed in 2005 by Gardes *et al.* [61]. In common with the MOS capacitor [24], the depletion-type phase shifter is not limited by the minority carrier recombination lifetime and is based on the principle of removing carriers from the junction area when applying a reverse bias. Figure 4.31 shows a four-terminal asymmetric *pn* structure, where the concentration of *n*-type doping is much higher than the concentration of *p*-type doping. The reason for such a structure is first to minimise the optical losses induced by the *n*-type doping and secondly to enhance the depletion overlap between the optical mode and the *p*-type region, in order to induce a better phase-shift-to-length ratio.

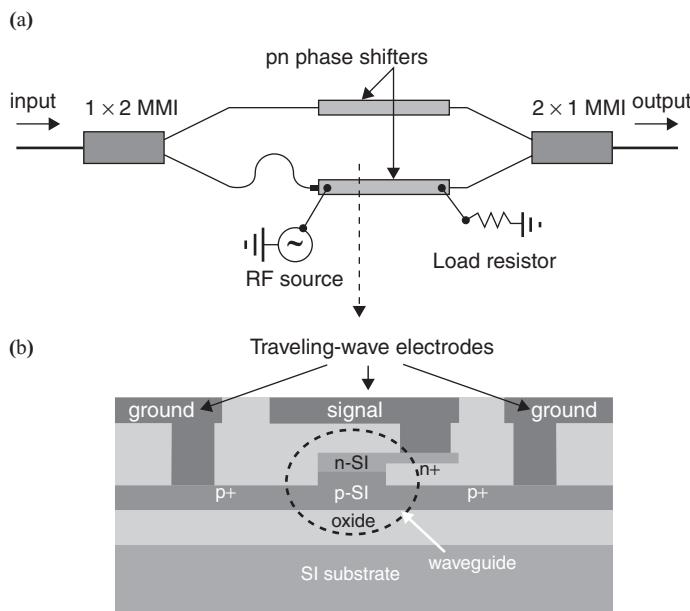
The carrier concentration variation in this kind of device is not uniform, as can be seen in the predictions of the refractive index change in the waveguide shown in Figure 4.32, and arises on both sides of the junction over a width of around 200 nm. One way to optimise the device is by increasing the overlap between the optical mode and the *p*-type depleted region. The main advantage of using depletion is obviously the very fast response time, simulated to be 7 ps for this modulator. This corresponds to an intrinsic bandwidth of approximately 50 GHz. The device proposed in [61] is 2.5 mm long and operates with a reverse bias swing of 5 V in a push–pull configuration as part of a Mach–Zehnder interferometer (MZI).



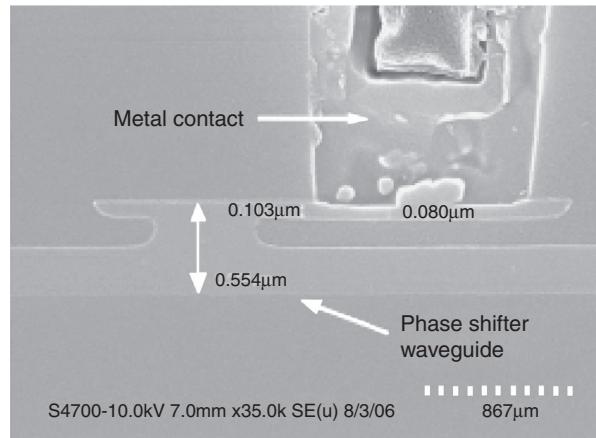
**Figure 4.32** Left: variation of the refractive index in the waveguide. Right: rise and fall time for TE and TM. Reproduced from [61] by permission of the Optical Society of America

In 2007, following the same principle Liu *et al.* [62] demonstrated in a *pn*-junction-based silicon optical modulator. Figures 4.33 and 4.34 show the schematic of the modulator as well as an SEM picture of the modulator cross-section.

The modulator shown in Figure 4.33 comprises a *p*-type doped crystalline silicon rib waveguide having a rib width of  $\sim 0.6 \mu\text{m}$  and a rib height of  $\sim 0.5 \mu\text{m}$  with an *n*-type doped silicon cap layer  $\sim 1.8 \mu\text{m}$  wide. This cap layer of around  $0.1 \mu\text{m}$  thick is formed using a nonselective



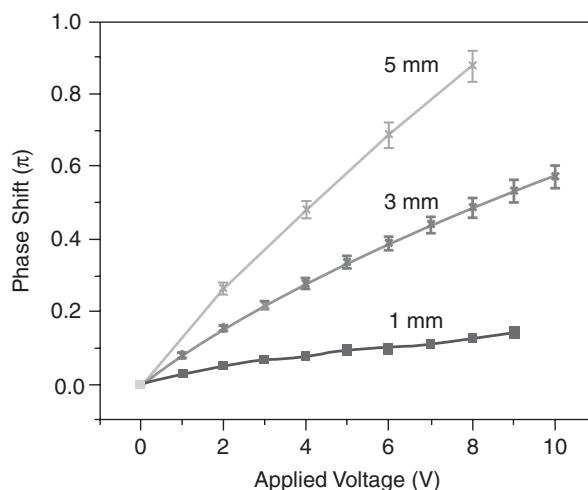
**Figure 4.33** (a) Top view of the proposed MZI modulator; (b) Cross-sectional view of the modulator [62]



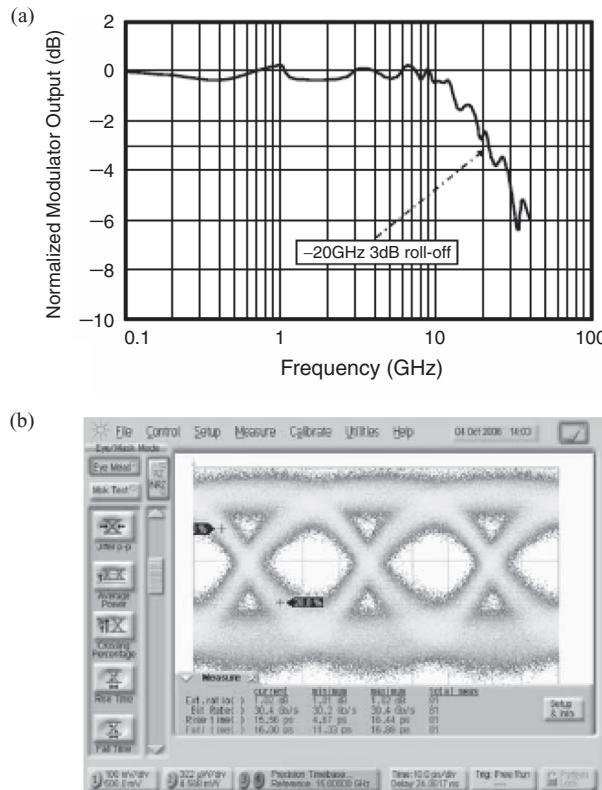
**Figure 4.34** SEM of the cross-section of the optical modulator in. Reproduced from [62] by permission of the Optical Society of America

epitaxial silicon growth process and is used for the formation of the *n*-doped electrical contact. The *p*-doping concentration is about  $1.5 \times 10^{17} \text{ cm}^{-3}$ , and the *n*-doping concentration varies from around  $3 \times 10^{18} \text{ cm}^{-3}$  near the top of the cap layer to  $1.5 \times 10^{17} \text{ cm}^{-3}$  at the *pn* junction. A good ohmic contact is ensured between the silicon and the metal contacts by two slab regions, which are situated 1  $\mu\text{m}$  away from both sides of the rib edge, and a cap layer region at about 0.3  $\mu\text{m}$  away from the rib edge. Those three regions are heavily doped with a dopant concentration of  $1 \times 10^{20} \text{ cm}^{-3}$ .

Figure 4.35 shows the phase shift against voltage of the phase shifter for three different lengths, that results in a modulation efficiency  $V_\pi L$  (where  $V_\pi$  is the bias voltage required for



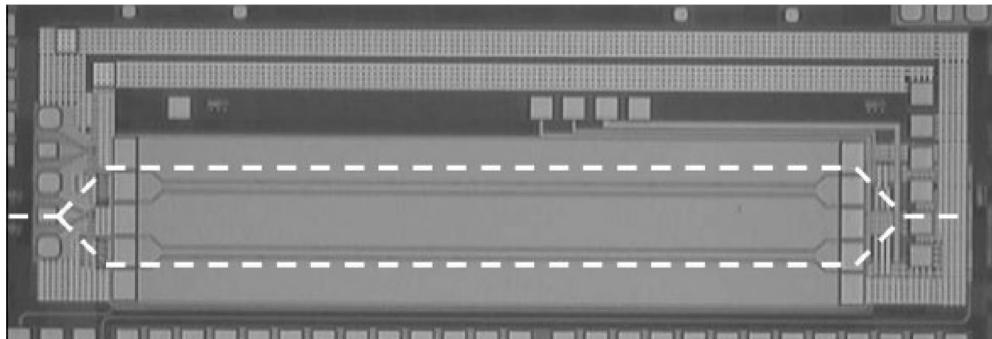
**Figure 4.35** Phase shift against reverse bias for different length of the active area of the phase shifter. Reproduced from [62] by permission of the Optical Society of America



**Figure 4.36** (a) Representation of the response of the modulator as a function of the RF frequency for a 1 mm active area; (b) optical eye diagram of the modulator with a 1-mm-long active area. Reproduced from [62] by permission of the Optical Society of America

$\pi$  phase shift and  $L$  is the device length) of approximately 4 V cm. The modulation bandwidth was also measured on a 1 mm long phase shifter and Figure 4.36 (a) shows that the MZI modulator has a 3 dB roll-off frequency of  $\sim 20$  GHz. Figure 4.36 (b) shows the eye diagram at a bit rate of 30 Gb/s, when using a pseudo random bit sequence with  $2^{31} - 1$  pattern length as the RF source. Whilst this is one of the fastest reported experimental optical modulator in silicon on insulator based on the plasma dispersion effect, although the modulation depth was only of the order of 1dB.

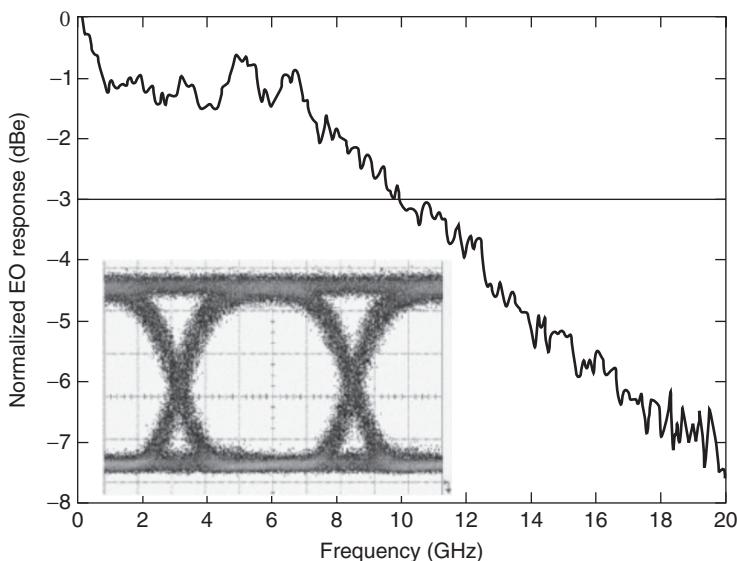
The general trend in silicon photonics is towards devices using sub-micrometre waveguides to achieve good phase and power efficiency, as well as to improve the speed and reduce the real estate of devices. The stage where integration of photonic devices and electronics circuits on the same chip is becoming a major milestone in silicon photonics. In early 2006, Gunn *et al.* [63] demonstrated modulation in both Mach–Zehnder interferometer (MZI) and ring resonator with a data rate up to 10 Gb/s. The authors stated that the modulator drivers were integrated on the chip, but did not provide any details about the electronics or the technology used to change the effective index of the mode in the waveguide. The information provided in [63] indicates that the waveguides are 500 nm wide and have a cross-sectional area of  $0.1 \mu\text{m}^2$ .



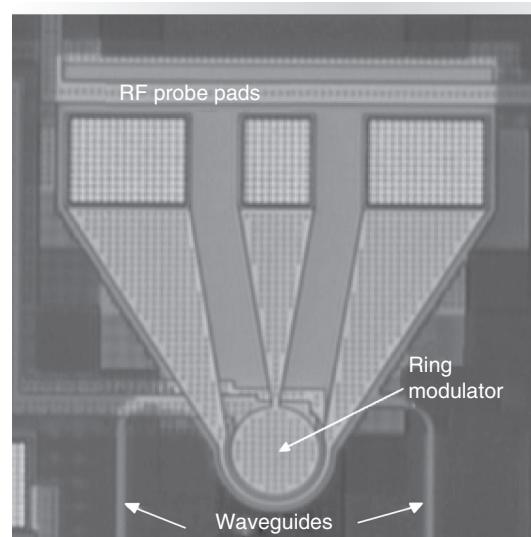
**Figure 4.37** Photograph of the Mach–Zehnder interferometer. Reproduced from [63] by permission of IEEE © (2006)

Figure 4.37 shows a top view from the MZI, where the dashed lines show the optical waveguide structure buried in the silicon and the driver input pads on the left, supply rails across the top and termination pads on the right. Figure 4.38 shows an eye diagram, when the modulator is operated at 10 Gb/s and the frequency response which shows a 3 dB cut-off frequency at 10 GHz. The authors also stated that the typical extinction ratio were 5 dB when the modulator is driven at 2.5 V.

As mentioned above, Gunn *et al.* [63] also demonstrated a ring resonator modulator. The top view of the metal contact is shown in Figure 4.39. The ring was used as a tunable notch filter where the frequency response is shown in Figure 4.41 and centred between 1524 and 1525 nm. The ring radius was 30  $\mu\text{m}$  and was a major improvement in terms of real estate



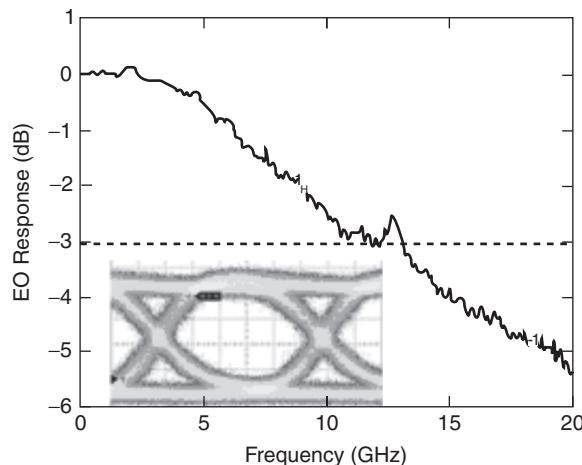
**Figure 4.38** MZI silicon optical modulator 10 Gb/s eye diagram and frequency response. Reproduced from [63] by permission of IEEE © (2006)



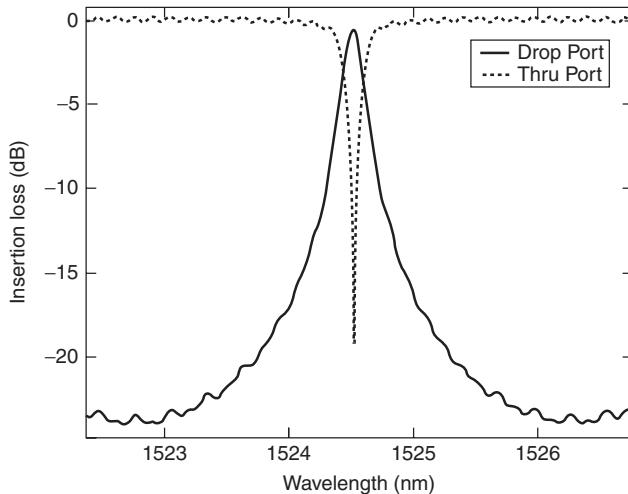
**Figure 4.39** Photograph of a ring resonator modulator with RF probe pads. Reproduced from [63] by permission of IEEE © (2006)

compared to the proposed MZI modulators which occupy approximately  $2\text{ mm}^2$ . Figure 4.40 shows the eye diagram of the ring resonator modulator operated at 10 Gb/s and the frequency response which shows a cut-off frequency of around 10 GHz.

In a more recent publication relating to the device proposed by Gunn *et al.*, Huang *et al.* [64] described the effect used in the optical modulator to be based on the free carrier plasma dispersion. The transducer was a reverse-biased lateral PIN diode, where modulation was obtained when majority carriers are swept in and out of the optical mode by an electric field.



**Figure 4.40** Ring resonator modulator 10 Gb/s eye diagram and frequency response. Reproduced from [63] by permission of IEEE © (2006)

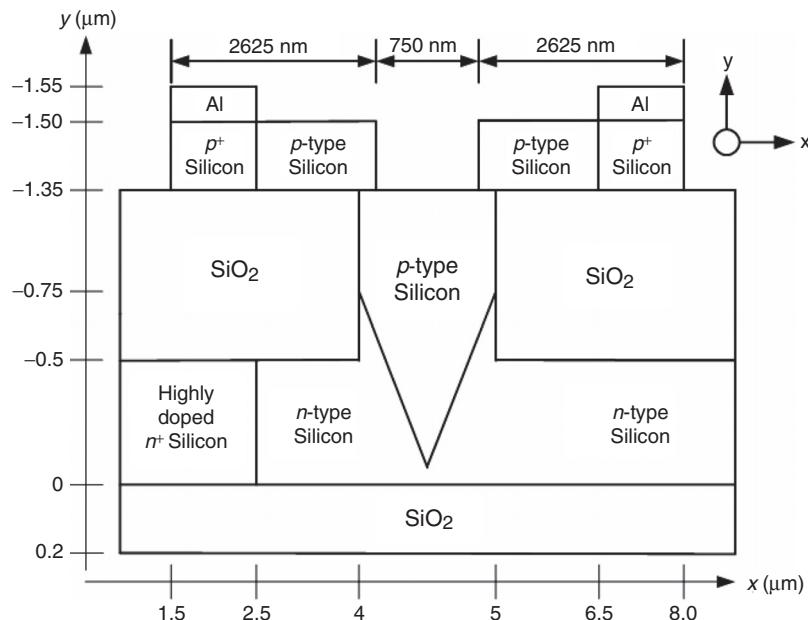


**Figure 4.41** Wavelength response of the drop port and through port of the ring resonator. Reproduced from [63] by permission of IEEE © (2006)

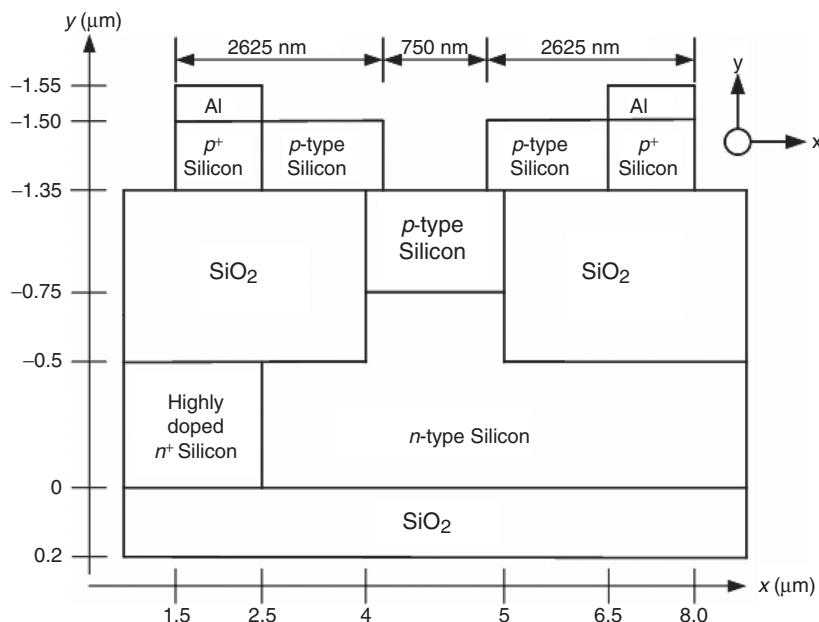
The modulator has a length of 2 mm and a performance of about  $5^\circ/\text{mm}/\text{V}/\text{arm}$ . As high speed modulation was based on majority carriers, the device was entirely limited by RLC parasitics. The diode was driven using a travelling wave electrode design, where the diode junction is designed as part of the microwave transmission line. The geometry of the travelling wave electrode was chosen such that the electrical group velocity and the optical group velocity were approximately matched. The transmission line itself had an impedance of below  $25 \Omega$ , but when loaded with the diode, the total system achieves  $25 \Omega$ .

The trend in silicon photonics, in the last few years has been to reduce waveguide size to obtain maximum gain in the real estate of devices as well as to increase the performance of active devices. Using different methods for the modulation, optical modulators in silicon have seen their bandwidth increased to reach multi-GHz frequencies. In order to simplify fabrication, one requirement for a waveguide, as well as for a modulator, is to retain polarisation independence in any state of operation and to be as small as possible. In 2006, Gardes *et al.* [65] proposed a way to obtain polarisation independence and improve the efficiency of an optical modulator using a V-shaped *pn* junction based on the natural etch angle of silicon,  $54.7^\circ$ . This modulator was compared to a flat junction depletion-type modulator of the same size and doping concentration. The proposed modulator was a *pn* junction formed in a V-shaped structure, as shown in Figure 4.42 [65]. The structure to which this was compared [61] is shown in Figure 4.43 [65].

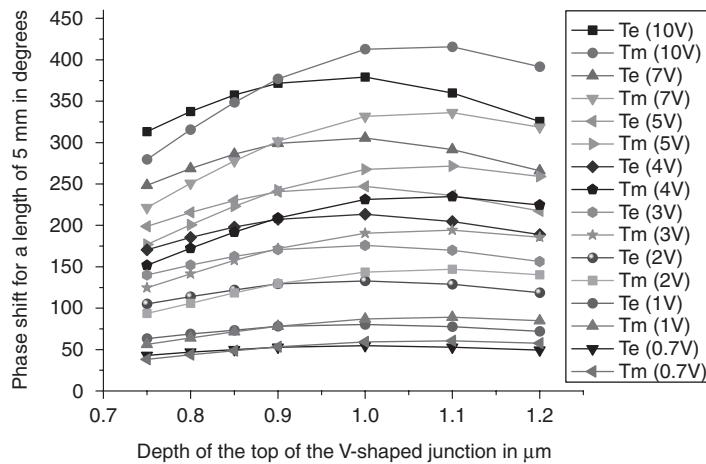
The simulations undertaken were to determine the conditions for polarisation independence of the waveguide during modulation. Figure 4.44 [65] and Figure 4.45 [45] show the phase shift achieved for different junction depths and voltages. Figure 4.44 [65] shows the simulation resulting from the V-shape junction and Figure 4.45 [65] shows similar results for the flat junction. In order to achieve polarisation independence during modulation the optimal positioning of the junction inside the waveguide has to be determined. Figure 4.44 [61] and Figure 4.45 [65] show the junction depth for both devices where polarisation independence is achieved. For the V-shaped junction the polarisation



**Figure 4.42** V-shaped *pn* junction optical modulator in silicon technology. Reproduced from [65] permission of IEEE © (2006)



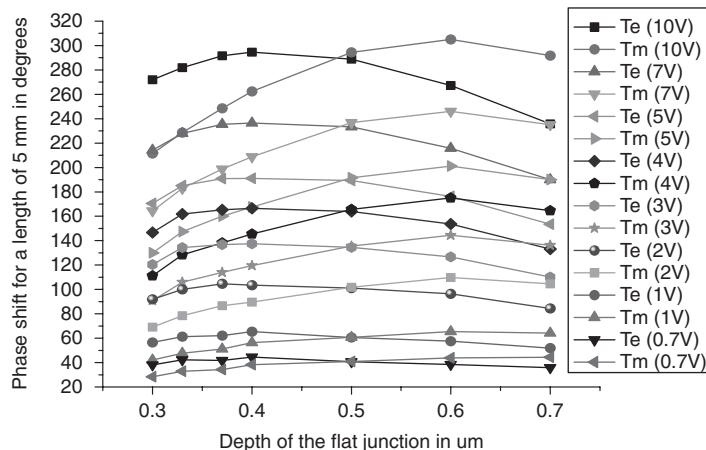
**Figure 4.43** Flat *pn* junction optical modulator in silicon technology. Reproduced from [65] permission of IEEE © (2006)



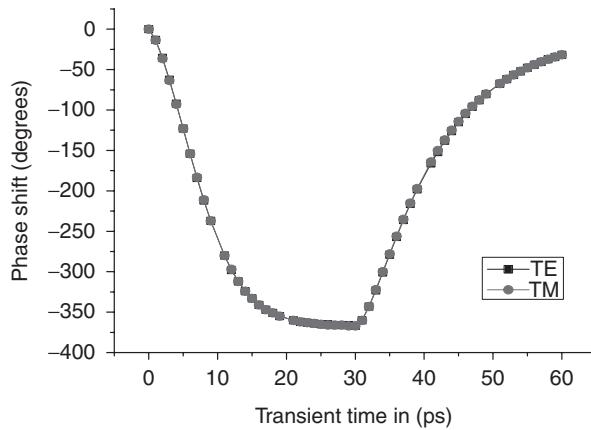
**Figure 4.44** Phase shift modulation (degree) dependence for TE and TM using a V-shaped *pn* junction (depth  $\mu\text{m}$ ). Reproduced from [61] by kind permission of IEEE © (2006)

independence during modulation is achieved when the top of the V-shaped junction is situated at a depth of 0.87  $\mu\text{m}$ . (Figure 4.44) The flat junction modulator achieves polarisation independence with a junction situated at 0.5  $\mu\text{m}$  from the bottom of the waveguide. (Figure 4.44) This is situated at the rib/slab interface which can facilitate fabrication.

The second parameter to extract from those simulations is that the maximum phase shift achieved at polarisation independence for both modulators is different. For a waveguide length of 5 mm the V-shaped junction achieve a phase shift of  $365^\circ$  ( $L_\pi V_\pi = 2.5 \text{ V cm}$ ) whereas for the flat junction the maximum phase shift is 290 degrees ( $L_\pi V_\pi = 3.1 \text{ V cm}$ ). These results show



**Figure 4.45** Phase shift modulation (degree) dependence for TE and TM using a flat *pn* junction (depth  $\mu\text{m}$ ). Reproduced from [61] by kind permission of IEEE © (2006)

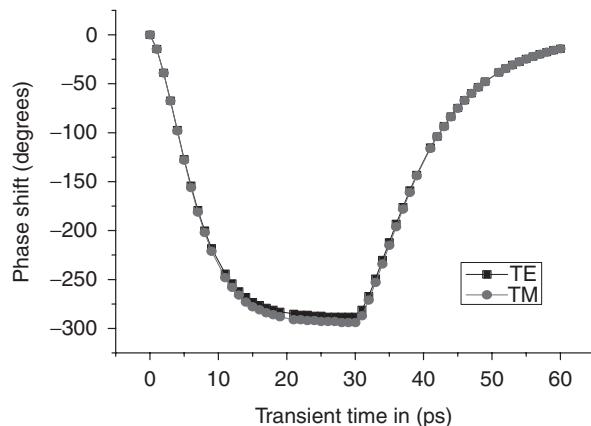


**Figure 4.46** Phase shift against transient time for TE and TM for the V-shaped junction. Reproduced from [65] by kind permission of IEEE © (2006)

that the V-shaped *pn* junction can achieve the same efficiency as the modulator proposed in [61] without the problems involved with submicron waveguides. The next step was to determine the bandwidth of both modulators. This is done by calculating the phase shift variation with transient time and is shown in Figure 4.46 and Figure 4.47 for TE and TM polarisation and for the V-shaped and flat junctions, respectively.

The transients show that the rise and fall time are similar for the both types of junctions. For this type of doping, the V-shaped junction has a rise time of 13 ps and a fall time of 23 ps for both TE and TM. For the flat junction the rise time is 12 ps and the fall time is 21 ps for both TE and TM. In both cases, this corresponds to an intrinsic bandwidth in excess of 15 GHz.

In the previous devices, interferometers and high-quality resonators were typically used to modulate the light, due to relatively weak plasma dispersion effect. Thin quantum well structures made from III-V semiconductors such as GaAs or InP exhibit a much stronger

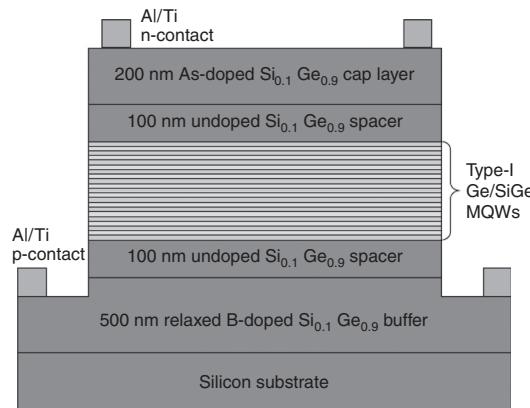


**Figure 4.47** Phase shift against transient time for TE and TM for the flat junction. Reproduced from [65] by kind permission of IEEE © (2006)

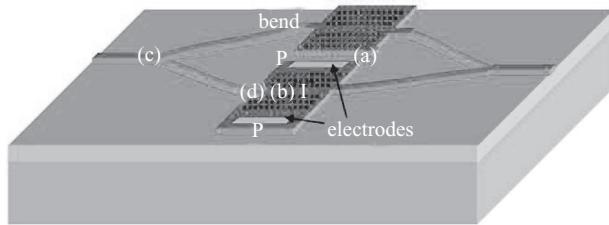
quantum-confined Stark effect (QCSE) mechanism [66], which allows modulator structures with optical path length in micrometres [67, 68]. The QCSE, at room temperature, in thin germanium quantum-well structures grown on silicon (Figure 4.48) has also been reported recently [69]. The QCSE in this structure is comparable to that in III–V materials and the authors claim that the process is CMOS compatible although fabrication complexity is undoubtedly increased. QCSE devices have typical length of the order of 100  $\mu\text{m}$  [70] and do not require carrier injection, as they are typically operated in reverse bias. The resulting low power dissipation, allows large arrays of devices at high data rates, and devices with operation bandwidths larger than 50 GHz have been reported in other materials [70]. This is very encouraging for high-speed, low-power, highly integrated electro-absorption modulators in silicon photonics, although much more development is required.

Pushing the trend in scaling devices even further, a low power and level of high integration has been achieved in silicon by using a photonic crystal (PhC) MZI modulator [71]. Photonic crystals are a class of artificial optical materials with periodic dielectric properties, which result in unusual optical properties (see Chapter 3). A photonic crystal waveguide is formed by introducing a line defect into a two-dimensional slab [72, 73]. The dispersion of photonic crystal optical waveguides offers the possibility to enhance the propagation constant by a factor larger than 100 [71, 73]. Therefore the active length of a rib or strip waveguide can be reduced by the same factor, to achieve the phase shift needed for modulation. The device shown schematically in Figure 4.50 and in Figure 4.51 was proposed by Jiang *et al.* [71], and is modulated using a current in the range of  $\sim\mu\text{A}$  and a maximum voltage of 7.5 mV, for an active length of 80  $\mu\text{m}$ .

Gu *et al.* [74] followed up on the results obtained by Jiang *et al.* [71] and reported high-speed operation using a PhC-based MZI having similar dimensions (lattice constant  $a = 400 \text{ nm}$ , hole diameter  $d = 220 \text{ nm}$ , overlayer thickness  $t = 260 \text{ nm}$ , and interaction length = 80  $\mu\text{m}$ ). The active element is made up of a *p*–*i*–*n* diode (Figure 4.51) in one arm of the MZI (Figure 4.52) with both *p*-type and *n*-type concentration of  $5 \times 10^{17} \text{ cm}^{-3}$ . The authors reported bit rates of 2 Mb/s and 1 Gb/s (Figure 4.53) with modulation depths of 85% and 20% respectively.



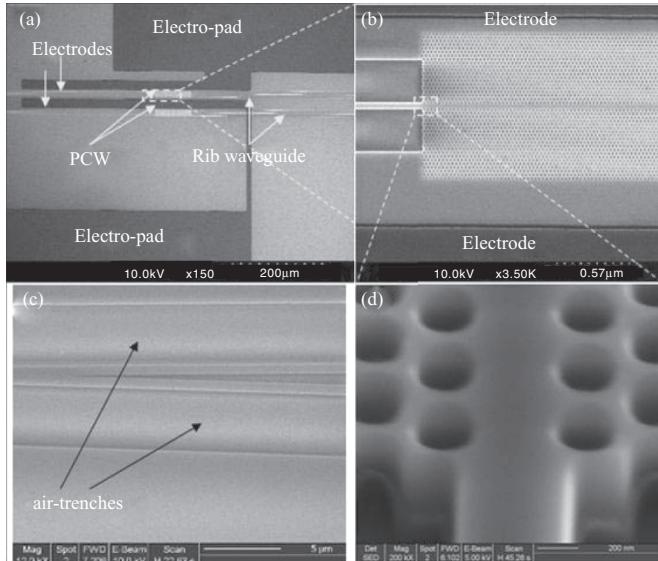
**Figure 4.48** Schematic diagram of a *p*–*i*–*n* showing the structure of strained Ge/Si multiple quantum wells grown on silicon on relaxed SiGe buffers. Reproduced from [69] by permission of MacMillan Publishers Ltd.



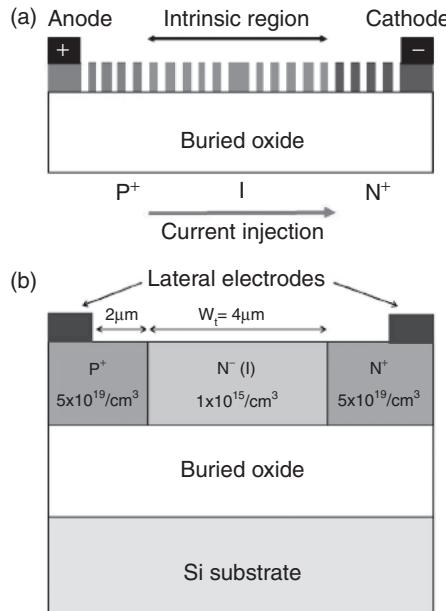
**Figure 4.49** Schematic of the MZI photonic crystal modulator: (a) electrodes; (b) photonic crystals; (c) rib waveguide Y-junction; (d) rib to photonic crystal junction. Reproduced from [71] by permission of American Institute of Physics

Maximum DC modulation depth of 93% was obtained with an injection current of 7.1 mA. Interestingly, this level of injection current is similar to those reported by Tang *et al.* [32] which has multi-micrometre device dimensions. Gu *et al.* [74] reported the device could be further improved by design optimisation, intrinsic region reduction, and fabrication to enhance matching and reduce drive voltage.

Silicon has been limited as an optical material for decades because of a lack of or limited active optical properties. In a paper published in 2006, Jacobsen *et al.* [75] demonstrated a significant linear electro-optic effect induced in silicon by breaking the crystal symmetry. The inversion symmetry of silicon crystal prohibits the existence of a linear electro-optic effect, hence by applying an asymmetric strain on the waveguide, the symmetry can be broken.



**Figure 4.50** SEM micrograph of the silicon PhC modulator with overlayer thickness  $t = 215$  nm: (a) overview picture of the modulator; (b) PhC waveguide with two electrodes; (c) Y-junction; (d) magnified PhC waveguide based on a triangular lattice with lattice constant  $a = 400$  nm, hole diameter  $d = 210$ . Reproduced from [71] by permission of American Institute of Physics

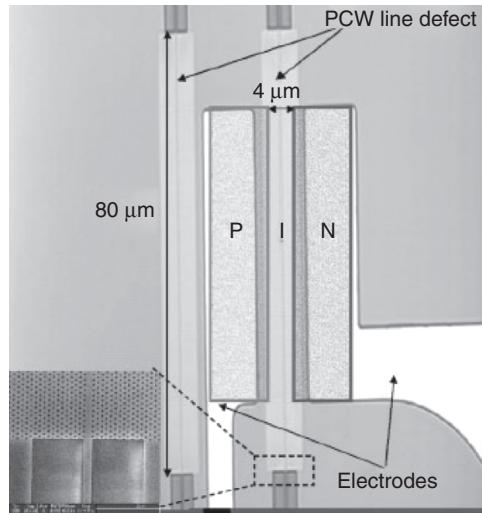


**Figure 4.51** Cross-sectional schematic of the active *p*–*i*–*n* diode (top), 2D model used in the electrical simulation (bottom). Reproduced from [74] by permission of American Institute of Physics

Figure 4.54 shows the principle used for the proposed device [75], where a straining layer is deposited on the top of the waveguide. The proposed structure used a deposited silicon nitride glass layer ( $\text{Si}_3\text{N}_4$ ) as a straining layer [76]. The amorphous  $\text{Si}_3\text{N}_4$  is pre-compressively strained and hence tries to expand the structure underneath in both horizontal directions, hence creating the asymmetry.

It is theoretically predicted [77] that the material nonlinearity [78] is enhanced linearly with the group index. By using photonic crystal waveguides, an enlarged group index can be obtained, hence increasing the material nonlinearity. The proposed device has been designed to achieve values above 230 for the group index. Figure 4.55 shows the use of photonic crystal waveguides inserted in a Mach–Zehnder interferometer.

Although the enhanced nonlinearity can be increased above  $800 \text{ pm V}^{-1}$  for a specific wavelength and a specific photonic crystal waveguide [75], it is important to make a fair comparison between different nonlinear materials. To do that, one must compare the material nonlinearity measured in [75], which is determined to be approximately  $15 \text{ pm V}^{-1}$  and the commonly applied nonlinear material,  $\text{LiNbO}_3$ . In  $\text{LiNbO}_3$  the largest tensor component [79] is approximately  $360 \text{ pm V}^{-1}$ . However, the authors stated that improvements in the material nonlinearity could be obtained if the silicon forming the waveguide could be deformed more freely by using, for example, nanowires. The advantages of using strain-induced nonlinearities in a modulator are fairly obvious as the speed of such a device would not be limited by charge mobility or charge recombination. Furthermore only an electric field is required and no current is needed, which is not the case in devices where changing the carrier concentration often requires a large ac current.



**Figure 4.52** Top view of the  $p-i-n$  diode implemented into the photonic crystal-based MZI. Reproduced from [74] by permission of American Institute of Physics

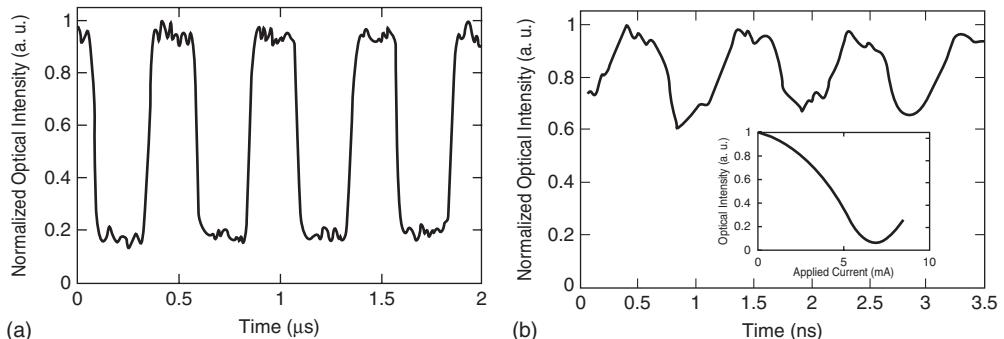
Early in 2007 Liu *et al.* [80] presented a design of monolithically integrated GeSi electroabsorption modulators and photodetectors based on the Franz–Keldysh effect. The GeSi composition was chosen for optimal performance around 1550 nm. The proposed modulator device was butt-coupled to a Si(core)/SiO<sub>2</sub>(cladding) high-index contrast waveguides, and is predicted to have a 3 dB bandwidth of  $>50$  GHz and an extinction ratio of 10 dB. The advantage of this design is that using the same device structure, a waveguide-coupled photodetector can also be integrated to the waveguide. The photodetector proposed by Liu *et al.* has a predicted responsivity of  $>1$  A/W and a 3 dB bandwidth of  $>35$  GHz.

As shown in Figure 4.56, the GeSi electroabsorption (EA) modulator and the photodetector are based on the same structure, potentially allowing efficient monolithic process integration. Both structures are based on a vertical Si/Ge<sub>0.9925</sub>Si<sub>0.0075</sub>/Si  $p-i-n$  diode with a doping level of  $2 \times 10^{19} \text{ cm}^{-3}$  in  $n^+$  and  $p^+$  Si, and their height  $H$  and width  $W$  can be designed to obtain optimal device performance. The only difference in the dimensions of the GeSi EA modulator and the photodetector is that the latter is longer than the former ( $L_2 > L_1$ ) to increase the absorption. The drawbacks of such a structure are the difficulty to obtain efficient butt coupling to minimise losses due to the impedance mismatch between the strip waveguide and the modulator as well as to fabricate the vertical Si/Ge<sub>0.9925</sub>Si<sub>0.0075</sub>/Si  $p-i-n$  diodes.

From the above, it is clear that silicon optical modulators utilising the plasma dispersion effect are capable of reaching multiple gigahertz speed using changes in the optical properties of the material via electrical signals. This is a huge improvement during a period when device operating speeds improved from 20 MHz in 2000 [46] to 1 GHz in 2004 [55], and latterly to 30 Gb/s in 2007 [62].

The question lies herein: is it possible to increase the operating speed of silicon optical devices further? Perhaps utilising other mechanisms?

Recently, Hochberg *et al.* [81] exploited the Kerr phenomenon by inducing optical nonlinearities in the waveguide material using an intense modulation beam. The authors overcame

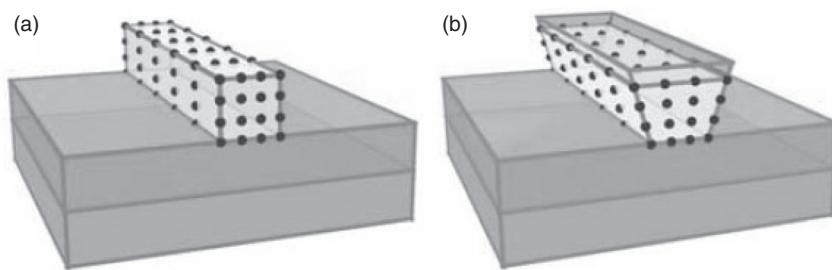


**Figure 4.53** Normalised optical intensity output; (a) at 2 Mb/s; (b) at 1 Gb/s. *Inset* shows the modulator optical output intensity against applied current. Reproduced from [74] by permission of American Institute of Physics

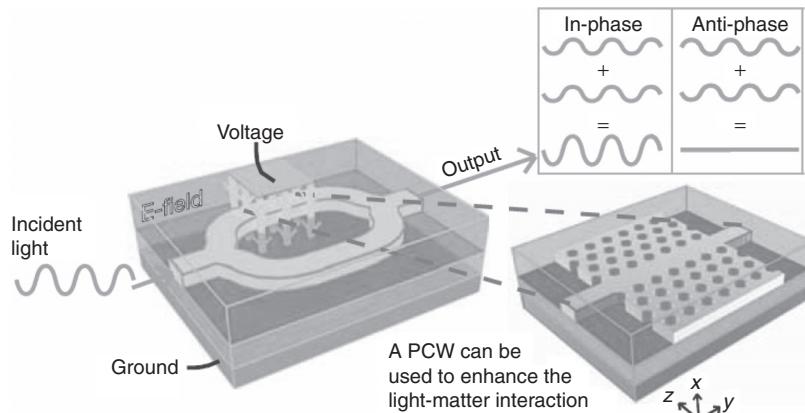
the relatively weak ultrafast nonlinearities in silicon by cladding a silicon waveguide with a specially engineered nonlinear optical polymer (Figure 4.57) based around a Mach–Zehnder interferometer (Figure 4.58).

The source signal is introduced to the MZI by way of a beam splitter and in one of the arms, a modulating gate signal is introduced by way of a 3 dB coupler. The nonlinear Kerr effect allows the gate signal to induce a phase shift in the source signal, which in turn allows the shifted source signal to interfere with the optical signal at the reference arm of the MZI, thereby causing an intensity modulation of the source signal. According to the authors, the demonstrated modulation frequency of 10 GHz is limited by existing measurement equipment and they showed via indirect evidence (spectral measurements) that the device can function into the terahertz range.

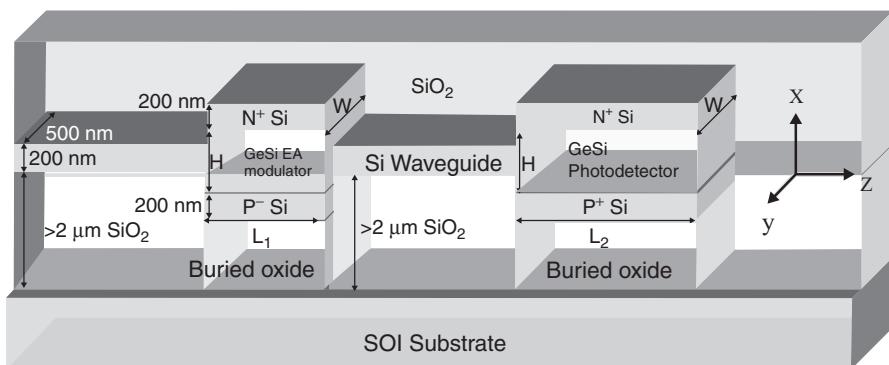
A key point to note is that the authors took advantage of efficient light confinement in the silicon waveguides and the highly nonlinear properties of optical polymer. However, it remains to be seen if such hybrid processes can be fully integrated into a CMOS process.



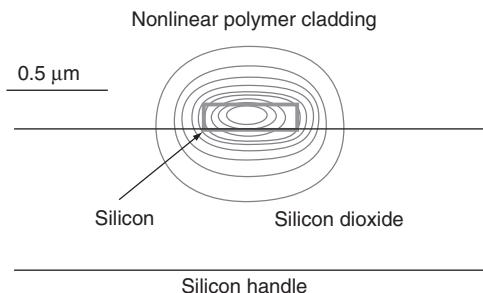
**Figure 4.54** Applying strain to crystalline silicon: (a) waveguide fabricated in the top layer of an SOI wafer; (b) the same waveguide with a straining layer deposited on top. The straining layer breaks the inversion symmetry and induces a linear electro-optic effect. Reproduced from [75] by permission of MacMillan Publishers Ltd.



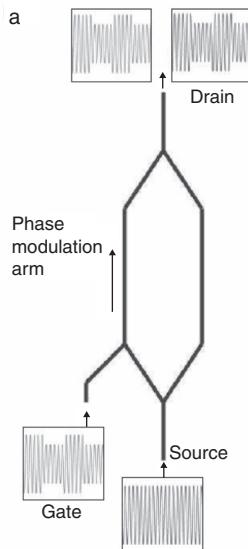
**Figure 4.55** Diagram of a Mach-Zehnder modulator. Incident light is split into two waveguides. The output amplitude depends on the phase difference at recombination. As shown at top right, in-phase recombination gives a ‘1’ bit output while anti-phase recombination gives a ‘0’. Reproduced from [75] by permission of MacMillan Publishers Ltd.



**Figure 4.56** Structure of the proposed monolithically integrated GeSi electroabsorption modulators and photodetectors. Reproduced from [80] by permission of the Optical Society of America



**Figure 4.57** Mode pattern of the optical signal in the silicon waveguide clad with a specially engineered nonlinear polymer. Contours are drawn in 10% increments of power. Reproduced from [81] by permission of MacMillan Publishers Ltd.



**Figure 4.58** The gate signal has its intensity modulation transferred to the source signal via non linear phase modulation in one arm of the Mach Zehnder interferometer. Reproduced from [81] by permission of MacMillan Publishers Ltd.

In July 2007, the Intel group improved the modulation speed of their device to 40 Gb/s [82], although the modulation depth was still modest at 1.6 dB. This improvement was attributed to improved packaging which minimises the parasitic effect; lowering the RF attenuation with better travelling-wave electrode design; and an enhanced modulator termination circuitry. The ability to operate at such a high speed, which matches the fastest devices deployed today, demonstrates the potential of integrated silicon photonic chips to be used in high-speed data rate streams in applications such as providing a cost-effective solutions for future optical interconnects. Furthermore, tera-scale computing may be possible via integrating multiple devices on a single chip.

#### 4.4 Conclusion

It is clear from the foregoing discussion that silicon optical modulators have undergone a significant transformation since the early work of the 1980s. Device dimensions have been reduced to allow operation at the market standard and above up to 40 Gb/s [80], and work is ongoing that predicts improvement of the bandwidth by at least another order of magnitude in the future. This is already a significant achievement in a material that does not intrinsically exhibit a significant electric field based modulation mechanism. Whilst these devices still have ground to make up on modulators in other material technologies, the possibility of producing all-silicon optical circuits, and subsequently integrating them with electronic circuits is a hugely attractive vision. This will not only increase the application areas of silicon, but will

also enhance the electronic circuits of the future. Hence silicon photonics is enabling the emergence of silicon as a candidate to become the photonics material of the future, and perhaps the platform upon which the next technological revolution will be built.

## References

- [1] R. A. Soref, ‘Silicon-based optoelectronics’, *Proc. IEEE*, **81**, 1993, 1687–706.
- [2] R. A. Soref, ‘Silicon-based photonic devices’, Solid state circuits conference, 1995, Digest of technical papers. 41st I SSCC, 1995, 66–67
- [3] G. T. Reed and B. L. Weiss, ‘Electro-optic effect in He<sup>+</sup> implanted optical waveguides in LiNbO<sub>3</sub>’, *Electronics Letters*, **23**, 1987, 424–5.
- [4] A. Polman, G. N. van den Hoven, J. S. Custer, J. H. Shin, R. Serna and P. F. A. Alkemade, ‘Erbium in crystal silicon: optical activation, excitation, and concentration limits’, *Journal of Applied Physics*, **77**, 1995, 1256–62.
- [5] G. T. Reed and C. E. Png, ‘Silicon Optical Modulators’, *Materials Today*, **8**, 2005, 40–50.
- [6] G. T. Reed, ‘Silicon On Insulator Photonics-An Introduction’, In: *School of Electronics, Computing and Mathematics — University of Surrey*, 2001, p. 245.
- [7] C. E. Png, G. T. Reed, W. R. Headley, K. P. Homewood, A. Liu, M. Paniccia, R. M. H. Atta, G. Ensell, A. G. R. Evans, D. Hak and O. Cohen, ‘Design and experimental results of small silicon-based optical modulators’, San Jose, CA, United States, 2004.
- [8] D. E. Aspnes, ‘Properties of Silicon’, in: *Properties of Silicon*. London and N.Y.: INSPEC, IEE, 1988, p. 72.
- [9] T. S. Moss, G. J. Burrell and B. Ellis, *Semiconductor opto-electronics*. London: Butterworth, 1973.
- [10] R. A. Soref and J. P. Lorenzo, ‘All-Silicon Active and Passive Guided-Wave Components for  $\lambda = 1.3$  and 1.6  $\mu\text{m}$ ’, *IEEE J. of Quantum Elect.*, **QE-22**, 1986, 873–879.
- [11] R. A. Soref and B. R. Bennett, ‘Electrooptical Effects in Silicon’, *IEEE J. Quantum Elect.*, **QE-23**, 1987, 123–129.
- [12] E. O. Brigham, *The Fast Fourier transform*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc, 1974.
- [13] D. E. Johnson and J. R. Johnson, *Mathematical Methods in Engineering and Physics*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc, 1982.
- [14] E. B. Saff and A. D. Snider, *Fundamentals of Complex Analysis for Mathematics, Science and Engineering*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc, 1986.
- [15] R. A. Soref and B. R. Bennett, ‘Kramers–Kronig analysis of electro-optical switching in silicon’, *Proc. SPIE: Integrated Optical Circuit Engineering IV*, 16–17 September 1986, Cambridge, MA, USA, 1987.
- [16] S. A. Clark, B. Culshaw, E. J. C. Dawnay and I. E. Day, ‘Thermo-optic phase modulators in SIMOX material’, *Proceedings of SPIE The International Society for Optical Engineering*, **3936**, 2000, 16–24.
- [17] D. J. Albares and R. A. Soref, ‘Silicon-on-sapphire waveguides’, *Proc. SPIE: Integrated Optical Circuit Engineering IV*, 16–17 September 1986, Cambridge, MA, USA.
- [18] R. A. Soref, J. Schmidtchen and K. Petermann, ‘Large single-mode rib waveguides in GeSi-Si and Si-on-SiO<sub>2</sub>’, *IEEE Journal of Quantum Electronics*, **27**, 1991 1971–1974.
- [19] B. L. Weiss, Reed, G. T., Toh, S. K., Soref, R. A. and F. Namavar, ‘Optical waveguides in SIMOX structures’, *IEEE Photonic. Tech. L.*, **3**, 1991, 19–21.
- [20] G. T. Reed, L. Jinhua, C. K. Tang, L. Chenglu, P. L. F. Hemment and A. G. Rickman, ‘Silicon on insulator optical waveguides formed by direct wafer bonding’, *Mat. Sci. Eng. B: Solid*, **B15**, 1992, 156–159.
- [21] J. P. Lorenzo and R. A. Soref, ‘1.3  $\mu\text{m}$  electro-optic silicon switch’, *Appl. Phys. Lett.*, **51**, 1987, 6–8.
- [22] L. Friedman, R. A. Soref and J. P. Lorenzo, ‘Silicon double-injection electro-optic modulator with junction gate control’, *J. Appl. Phys.*, **63**, 1988, 1831–9.
- [23] S. R. Giguere, L. Friedman, R. A. Soref and J. P. Lorenzo, ‘Simulation studies of silicon electro-optic waveguide devices’, *Journal of Applied Physics*, **68**, 1990, 4964–70.
- [24] A. Liu, R. Jones, L. Ling, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu and M. Paniccla, ‘A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor’, *Nature*, **427**, 2004, 615–18.
- [25] G. V. Treyz, P. G. May and J.-M. Halbout, ‘Silicon optical modulators at 1.3 micrometer based on free-carrier absorption’, *IEEE Electron Device Letters*, **12**, 1991, 276–278.
- [26] G. V. Treyz, P. G. May and J. M. Halbout, ‘Silicon Mach-Zehnder waveguide interferometers based on the plasma dispersion effect’, *Applied Physics Letters*, **59**, 1991, 771–3.

- [27] S. M. Jackson, P. D. Hewitt, G. T. Reed, C. K. Tang, A. G. R. Evans, J. Clark, C. Aveyard and F. Namavar, 'Novel optical phase modulator design suitable for phased arrays', *J. Lightwave Technol.*, **16**, 1998, 2016–2019.
- [28] S. M. Jackson, G. T. Reed, C. K. Tang, A. G. R. Evans, J. Clark, C. Aveyard and F. Namavar, 'Optical beamsteering using integrated optical modulators', *J. Lightwave Technol.*, **15**, 1997, 2259–2263.
- [29] U. Fischer, B. Schuppert and K. Petermann, 'Integrated optical switches in silicon based on SiGe-waveguides', *IEEE Photonics Technology Letters*, **5**, 1993, 785–7.
- [30] C. K. Tang, G. T. Reed, A. J. Wilson and A. G. Rickman, 'Low-loss, single-mode, optical phase modulator in SIMOX material', *J. Lightwave Technol.*, **12**, 1994, 1394–1400.
- [31] C. K. Tang, G. T. Reed, A. J. Wilson and A. G. Rickman, 'Simulation of a low loss optical modulator for fabrication in SIMOX material', *Silicon-Based Optoelectronic Materials Symposium*, 12–14 April 1993, San Francisco, CA, USA, 1993.
- [32] C. K. Tang and G. T. Reed, 'Highly efficient optical phase modulator in SOI waveguides', *Electronics Letters*, **31**, 1995, 451–452.
- [33] H. C. Huang and T. C. Lo, 'Simulation and analysis of silicon electro-optic modulators utilizing the carrier-dispersion effect and impact-ionization mechanism', *Journal of Applied Physics*, **74**, 1993, 1521–8.
- [34] Synopsys, Inc. 700 East Middlefield Road, Mountain View, CA 94043, USA.
- [35] A. Cutolo, M. Iodice, P. Spirito and L. Zeni, 'Silicon electro-optic modulator based on a three terminal device integrated in a low-loss single-mode SOI waveguide', *Journal of Lightwave Technology*, **15**, 1997, 505–518.
- [36] A. Cutolo, M. Iodice, A. Irace, P. Spirito and L. Zeni, 'An electrically controlled Bragg reflector integrated in a rib silicon on insulator waveguide', *Applied Physics Letters*, **71**, 1997, 199–201.
- [37] A. Irace, G. Breglio and A. Cutolo, 'All-silicon optoelectronic modulator with 1 GHz switching capability', *Electronics Letters*, **39**, 2003, 232–233.
- [38] G. Breglio, A. Cutolo, A. Irace, P. Spirito, L. Zeni, M. Iodice and P. M. Sarro, 'Two silicon optical modulators realizable with a fully compatible bipolar process', *IEEE Journal of Selected Topics in Quantum Electronics*, **4**, 1998, 1003–10.
- [39] A. Sciuto, S. Libertino, A. Alessandria, S. Coffa and G. Coppola, 'Design, fabrication and testing of an integrated Si-based light modulator', *Journal of Lightwave Technology*, **21**, 2003, 228–235.
- [40] P. D. Hewitt and G. T. Reed, 'Multi micron dimension optical  $p-i-n$  modulators in silicon-on-insulator', San Jose, CA, USA, 1999.
- [41] P. D. Hewitt and G. T. Reed, 'Improving the response of optical phase modulators in SOI by computer simulation', *Journal of Lightwave Technology*, **18**, 2000, 443–450.
- [42] P. D. Hewitt and G. T. Reed, 'Improved modulation performance of a silicon  $p-i-n$  device by trench isolation', *Journal of Lightwave Technology*, **19**, 2001, 387–90.
- [43] I. E. Day, S. W. Roberts, R. O'Carroll, A. Knights, P. Sharp, G. F. Hopper, B. J. Luff and M. Asghari, 'Single-chip variable optical attenuator and multiplexer subsystem integration', Anaheim, CA, United States, 2002.
- [44] I. Day, I. Evans, A. Knights, F. Hopper, S. Roberts, J. Johnston, S. Day, J. Luff, H. Tsang and M. Asghari, 'Tapered silicon waveguides for low insertion loss highly-efficient high-speed electronic variable optical attenuators', Atlanta, GA, USA, 2003.
- [45] R. R. Whiteman, A. P. Knights, D. George, I. E. Day, A. Vonsovici, A. A. House, G. F. Hopper and M. Asghari, 'Recent progress in the design, simulation and fabrication of small cross-section silicon-on-insulator VOAs', San Jose, CA, United States, 2003.
- [46] P. Dainesi, A. Kung, M. Chablotz, A. Lagos, P. Fluckiger, A. Ionescu, P. Fazan, M. Declerq, P. Renaud and P. Robert, 'CMOS compatible fully integrated Mach-Zehnder interferometer in SOI technology', *IEEE Photonics Technology Letters*, **12**, 2000, 660–2.
- [47] T. W. Ang, P. D. Hewitt, A. Vonsovici, G. T. Reed, A. G. R. Evans, P. R. Routley, T. Blackburn and M. R. Josey, 'Integrated optics in Unibond for greater flexibility', Seattle, WA, USA, 1999.
- [48] T. W. Ang, G. T. Reed, A. Vonsovici, A. G. R. Evans, P. R. Routley and M. R. Josey, '0.15 dB/cm loss in Unibond SOI waveguides', *Electronics Letters*, **35**, 1999, 977–8.
- [49] C. E. Png, S. P. Chan, S. T. Lim and G. T. Reed, 'Optical phase modulators for MHz and GHz modulation in silicon-on-insulator (SOI)', *Journal of Lightwave Technology*, **22**, 2004, 1573–1582.
- [50] C. E. Png, G. Masanovic and G. T. Reed, 'Coupling to 1 micron silicon modulators', San Jose, CA, USA, 2002.
- [51] C. E. Png, G. T. Reed, R. M. H. Atta, G. Ensell and A. G. R. Evans, 'Development of small silicon modulators in Silicon-On-Insulator (SOI)', San Jose, CA, United States, 2003.
- [52] R. M. H. Atta, G. J. Ensell, A. G. R. Evans, C. E. Png and G. T. Reed, 'Fabrication of a highly efficient optical modulator based on silicon-on-insulator', Maspalomas, Gran Canaria, Spain, 2003.

- [53] C. E. Png, ‘Silicon-On-Insulator Phase Modulators’, *Section 5.2.13, Ph.D. Thesis*, University of Surrey, 2004.
- [54] C. A. Barrios, V. R. de Almeida and M. Lipson, ‘Low-power-consumption short-length and high-modulation-depth silicon electrooptic modulator’, *Journal of Lightwave Technology*, **21**, 2003, 1089–98.
- [55] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu and M. Paniccia, ‘A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor’, *Nature*, **427**, 2004, 615–18.
- [56] F. Gan and F. X. Kartner, ‘High-speed silicon electrooptic modulator design’, *IEEE Photonics Technology Letters*, **17**, 2005, 1007–1009.
- [57] X. Qianfan, B. Schmidt, S. Pradhan and M. Lipson, ‘Micrometre-scale silicon electro-optic modulator’, *Nature*, **435**, 2005, 325–7.
- [58] Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya and M. Lipson, ‘12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators’, *Opt. Express*, **15**, 2007, 430–436.
- [59] R. Jones, H. Rong, A. Liu, A. W. Fang, M. J. Paniccia, D. Hak and O. Cohen, ‘Net continuous wave optical gain in a low loss silicon-on-insulator waveguide by stimulated Raman scattering’, *Optics Express*, **13**, 2005, 519–525.
- [60] R. Jones, L. Ansheng, R. Haisheng, M. Paniccia, O. Cohen and D. Hak, ‘Lossless optical modulation in a silicon waveguide using stimulated Raman scattering’, *Optics Express*, **13**, 2005, 796–800.
- [61] F. Y. Gardes, G. T. Reed, N. G. Emerson and C. E. Png, ‘A sub-micron depletion-type photonic modulator in silicon on insulator’, *Optics Express*, **13**, 2005, 8845–8854.
- [62] A. Liu, L. Liao, D. Rubin, H. Nguyen, B. Ciftcioglu, Y. Chetrit, N. Izhaky and M. Paniccia, ‘High-speed optical modulation based on carrier depletion in a silicon waveguide’, *Optics Express*, **15**, 2007, 660–668.
- [63] C. Gunn, ‘CMOS Photonics for High-Speed Interconnects’, *Micro, IEEE*, **26**, 2006, 58–66.
- [64] A. Huang, C. Gunn, G. L. Li, S. Liang, S. Mirsaidi, A. Narasimha and T. Pinguet, ‘A 10Gb/s Photonic Modulator and WDM MUX/DEMUX integrated with electronics in 0.13 μm SOI CMOS’, Presented at ISSCC, San Francisco, 2006.
- [65] G. T. Reed, F. Y. Gardes, B. D. Timotijevic, W. R. Headley and G. Z. Mashanovich, ‘Tailoring the Response of Silicon Photonics Devices’ *Laser and Electrooptics Society*, 2006 316–317.
- [66] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood and C. A. Burrus, ‘Band-edge electroabsorption in quantum well structures: the quantum-confined Stark effect’, *Physical Review Letters*, **53**, 1984, 2173–6.
- [67] U. Arad, E. Redmard, M. Shamay, A. Averboukh, S. Levit and U. Efron, ‘Development of a large high-performance 2-D array of GaAs-AlGaAs multiple quantum-well modulators’, *IEEE Photonics Technology Letters*, **15**, 2003, 1531–1533.
- [68] L. Chin-Pang, A. Seeds, J. S. Chadha, P. N. Stavrinou, G. Parry, M. Whitehead, A. B. Krysa and J. S. Roberts, ‘Design, fabrication and characterization of normal-incidence 1.56-μm multiple-quantum-well asymmetric Fabry-Perot modulators for passive picocells’, *IEICE Transactions on Electronics*, **E86-C**, 2003, 1281–9.
- [69] Y.-H. Kuo, Y. K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D. A. B. Miller and J. S. Harris, ‘Strong quantum-confined Stark effect in germanium quantum-well structures on silicon’, *Nature*, **437**, 2005, 1334–1336.
- [70] R. Lewen, S. Irmscher, U. Westergren, L. Thylen and U. Eriksson, ‘Segmented Transmission-Line Electroabsorption Modulators’, *Journal of Lightwave Technology*, **22**, 2004, 172–179.
- [71] Y. Jiang, W. Jiang, L. Gu, X. Chen and R. T. Chen, ‘80-micron interaction length silicon photonic crystal waveguide modulator’, *Applied Physics Letters*, **87**, 2005, 221105–1.
- [72] S. G. Johnson, P. R. Villeneuve, F. Shanhui and J. D. Joannopoulos, ‘Linear waveguides in photonic-crystal slabs’, *Physical Review B (Condensed Matter)*, **62**, 2000, 8212–22.
- [73] M. Soljacic, S. G. Johnson, F. Shanhui, M. Ibanescu, E. Ippen and J. D. Joannopoulos, ‘Photonic-crystal slow-light enhancement of nonlinear phase sensitivity’, *Journal of the Optical Society of America B (Optical Physics)*, **19**, 2002, 2052–9.
- [74] L. Gu, W. Jiang, X. Chen, L. Wang and R. T. Chen, ‘High speed silicon photonic crystal waveguide modulator for low voltage operation’, *Applied Physics Letters*, **90**, 2007, 071105–1.
- [75] R. S. Jacobsen, K. N. Andersen, P. I. Borel, J. Fage-Pedersen, L. H. Frandsen, O. Hansen, M. Kristensen, A. V. Lavrinenko, G. Moulin, H. Ou, C. Peucheret, B. Zsigri and A. Bjarklev, ‘Strained silicon as a new electro-optic material’, *Nature*, **441**, 2006, 199–202.
- [76] J. M. Madou, *Fundamentals of Microfabrication*, Boca Raton ed. Florida: CRC Press, 2002.
- [77] M. Soljacic and J. D. Joannopoulos, ‘Enhancement of nonlinear effects using photonic crystals’, *Nature Materials*, **3**, 2004, 211–219.

- [78] P. N. Butcher and D. Cotter, *The Elements of Nonlinear Optics 5*, Cambridge University Press, Cambridge, 1990.
- [79] G. L. Li and P. K. L. Yu, ‘Optical intensity modulators for digital and analog applications’, *Journal of Lightwave Technology*, **21**, 2003, 2010–30.
- [80] J. Liu, D. Pan, S. Jongthammanurak, K. Wada, L. C. Kimerling and J. Michel, ‘Design of monolithically integrated GeSi electro-absorption modulators and photodetectors on a SOI platform’ *Opt. Express*, **15**, 2007, 623–628.
- [81] M. Hochberg, T. Baehr-Jones, G. Wang, M. Shearn, K. Harvard, J. Luo, B. Chen, Z. Shi, R. Lawson, P. Sullivan, A. K. Y. Jen, L. Dalton and A. Scherer, ‘Terahertz all-optical modulation in a silicon-polymer hybrid system’, *Nature Materials*, **5**, 2006, 703–709.
- [82] L. Liao, A. Liu, D. Rubin, J. Basak, Y. Chetrit , H. Nguyen, R. Cohen, N. Izhaky and M. Paniccia, ‘40 Gbit/s silicon optical modulator for high-speed applications,’ *Electron. Lett.* **43**, (22), 2007.

# 5

## Silicon Lasers

Bahram Jalali, Dimitris Dimitropoulos, Varun Raghunathan  
and Sasan Fathpour

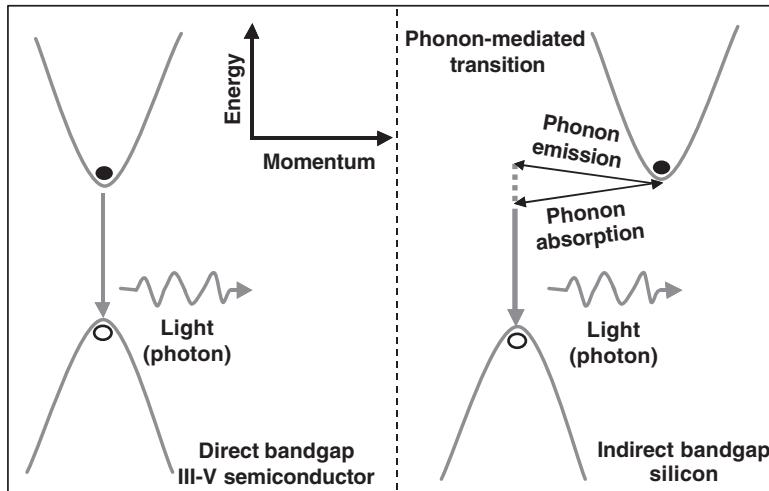
### 5.1 Introduction

Silicon, the wonder material of the 20th century, appears to have one more trick up its sleeve. After dominating the digital electronic industry, the bedrock of the digital world, it is now being considered as a platform of choice for building optoelectronic devices. Silicon's mature and large-scale manufacturing base is believed to be precisely what is needed to effect a much needed reduction in the cost of photonic devices that are currently being manufactured in compound semiconductors such as gallium arsenide and indium phosphide. Such a cost reduction can bring the power of optical networks to the desktop computer and to home entertainment systems. The technology, termed *silicon photonics*, would be compatible not only with CMOS-based electronic integrated circuits, but also with micro-electro-mechanical systems (MEMS). It could enable a new generation of electro-opto-mechanical chips that perform the job of today's complex systems at a fraction of the cost, size, and power dissipation.

For some time now, it has been known that when surrounded by its native oxide ( $\text{SiO}_2$ ), otherwise known as ordinary glass, silicon becomes a low-loss optical waveguide at wavelengths of 1300 and 1550 nm, the regions of the spectrum where optical networks operate. Using such waveguides as basic building blocks, optical filters and switches have been developed, and by adding germanium to the chip, photodetectors have also been produced. In other words, most of the building blocks that make up an optical communication system are already available. If we also had optical amplifiers and lasers, the silicon photonic toolbox would be complete.

### 5.2 Why Bulk Silicon Cannot Amplify Light

Silicon light sources and amplifiers have remained elusive because of the material's band structure. Silicon has an indirect band structure, which means that the upper and the lower electronic states (conduction and valence bands) do not occur at the same value of crystal



**Figure 5.1** Difference of light emission mechanisms between direct and indirect bandgap materials

momentum (Figure 5.1). Because visible or infrared photons have negligible momentum (compared with that of the electron), the de-excitation of an electron (recombination) needs to be mediated by emitting or absorbing a phonon in order to conserve momentum. Such second-order radiative recombination events do not occur frequently, as characterized by a very long lifetime that is of the order of 1 s. The electron in the upper state has to sit and wait until a phonon with the right momentum shows up. While waiting, it becomes prone to nonradiative recombination with the energy being dissipated as heat. The experimentally measured lifetime in silicon is in the millisecond to microsecond range, depending on the impurity or defect concentration. This suggests that the desired radiative processes are insignificant compared with undesired nonradiative recombination. Even when using the highest purity silicon, devices have an electrical to optical conversion efficiency of only  $10^{-4}$ – $10^{-3}$ . The situation would be different if one could break free of the low radiative recombination rate of bulk silicon. This has served as the premise for recent attempts at creating a silicon laser.

### 5.2.1 Physics of Optical Gain in Bulk Silicon

There are two types of electronic transitions that govern optical amplification in bulk silicon: the interband transitions and the intraband transitions. As will be shown, the first would lead to optical gain despite the indirect bandgap, if the second was absent. This section provides a derivation of optical gain in silicon inspired by the classic work of Dumke [1, 2].

We will first discuss interband transitions. We start with quantum mechanical expressions, resulting from second-order time-dependent perturbation theory, that describe the probability for transitions from the valence to the conduction band (and *vice versa*) with the simultaneous

absorption/emission of photons and phonons. These probabilities are found to be:

$$W_{V \rightarrow C} = \frac{2\pi}{\hbar} |M_{CV}|^2 f(E_V)(1 - f(E_C))n_{OPT}(E) \begin{cases} n_{VIB}(E_P)\delta(E_C - E_V - E - E_P) \\ + (1 + n_{VIB}(E_P))\delta(E_C - E_V - E + E_P) \end{cases} \quad (5.1)$$

$$W_{C \rightarrow V} = \frac{2\pi}{\hbar} |M_{CV}|^2 f(E_C)(1 - f(E_V))(1 + n_{OPT}(E)) \begin{cases} (1 + n_{VIB}(E_P))\delta(E_C - E_V - E - E_P) \\ + n_{VIB}(E_P)\delta(E_C - E_V - E + E_P) \end{cases} \quad (5.2)$$

where

$$M_{CV} = \sum_C \frac{\langle c | H_{E-VIB} | c' \rangle \langle c' | H_{E-RAD} | v \rangle}{E_{C'} - E_V}$$

The stimulated transition probability  $W_{V \rightarrow C}^{\text{stim}}$  in Equation (5.1) is proportional to the photon number  $n_{OPT}$  and describes absorption of photons. The part of Equation (5.2) that is proportional to  $n_{OPT}$  describes the stimulated emission,  $W_{C \rightarrow V}^{\text{stim}}$ , and the rest describe spontaneous emission,  $W_{C \rightarrow V}^{\text{spont}}$ . The two terms in the curly brackets describe the processes involving phonon absorption and emission, respectively. The term  $H_{E-VIB}$  is the Hamiltonian for the interaction of electrons and phonons,  $H_{E-RAD}$  is the Hamiltonian of the interaction of electrons with photons,  $|c\rangle$ ,  $|c'\rangle$  denote conduction band states,  $|v\rangle$  the valence band states. The number of vibration quanta is  $n_{VIB}$  and  $n_{OPT}$  denotes the number of photons in a given frequency. The electron distribution function,  $f(E)$ , and the phonon distribution function,  $n(E)$ , are given by:

$$f(E_C) = \frac{1}{\exp((E - E_{F,C})/kT) + 1} \quad (5.3a)$$

$$f(E_V) = \frac{1}{\exp((E - E_{F,V})/kT) + 1} \quad (5.3b)$$

$$n_{VIB}(E_P) = \frac{1}{\exp(E_P/kT) - 1} \quad (5.3c)$$

The first two equations are the Fermi-Dirac distribution function and the third is the Bose-Einstein function. Invoking the effective mass approximation, the energies of conduction and valence band electrons are equal to  $E_C(\vec{k}_C) = E_G + \hbar^2(k_C - k_{C0})^2/(2m_e^*)$  and  $E_V(\vec{k}_V) = \hbar^2k_V^2/(2m_h^*)$ , the photon energy is  $E(\vec{k}) = \hbar\omega = \hbar k(c/\sqrt{\epsilon_r})$  ( $\epsilon_r$  is the dielectric permittivity) and the phonon energy is denoted as  $E_P(\vec{q})$ .

As with all processes that occur in crystals, momentum must be conserved. Let  $\vec{k}_C$  and  $\vec{k}_V$  denote the wavevectors of the electrons in the conduction and the valence band, and  $\vec{k}, \vec{q}$  denote the wavevectors of the photons and the phonons. Then for the case where an electron makes a transition from the conduction to the valence band with the emission of a photon and a phonon the condition gives:

$$\vec{k}_C = \vec{k}_V + \vec{k} + \vec{q} \cong \vec{k}_V + \vec{q} \quad (5.4)$$

The last approximation describes the fact that the photon momentum (at infrared frequencies) is much smaller than the momentum mismatch,  $(\vec{k}_C - \vec{k}_V)\hbar/2\pi$ , for the transition from the top of the valence band to the bottom of the conduction band. This stems from the fact that

the optical wavelength is much larger than the lattice constant. Therefore the difference in momentum must be made up by the phonon wavevector  $\vec{q}$ . The momentum conservation Equation (5.4) ensures that  $M_{\text{CV}}$  is nonzero in the perturbation theory expressions (5.1 and 5.2).

Consider first the total spontaneous emission rate and a given optical frequency  $\omega$ . To obtain the total rate we must sum up all the possible processes that can lead to spontaneous emission of a photon at this frequency. To do so, we must sum up over all initial and final electron states  $\vec{k}_C$  and  $\vec{k}_V$ . Then the total rate of spontaneous emission per second is:

$$\frac{1}{\tau(\omega)} = \int \frac{d^3 \vec{k}_C}{(2\pi)^3 / V} \int \frac{d^3 \vec{k}_V}{(2\pi)^3 / V} W_{C \rightarrow V}^{\text{spont}} \quad (5.5)$$

Now, we turn our attention to the stimulated emission and absorption transitions, processes that depend on photon number  $n_{\text{OPT}}$ . The objective is to calculate the optical gain coefficient. To do so, we calculate the net stimulated emission rate and divide by the photon number and the speed of light in the medium. This gives:

$$g(\omega) = \frac{\sqrt{\epsilon_r}}{c} \int \frac{d^3 \vec{k}_C}{(2\pi)^3 / V} \int \frac{d^3 \vec{k}_V}{(2\pi)^3 / V} \frac{W_{C \rightarrow V}^{\text{stim}} - W_{V \rightarrow C}^{\text{stim}}}{n_{\text{OPT}}(E)} \text{ (cm}^{-1}) \quad (5.6)$$

This coefficient describes the amplification of optical intensity,  $I(\omega)$ , per unit length of travel in the crystal:

$$dI(\omega)/dx = g(\omega)I(\omega) \quad (5.7)$$

Expressions (5.1) and (5.2) describe two sets of transitions. In one set a phonon of energy  $E_p$  is annihilated and in the other it is created. If we denote the electronic bandgap of the material by  $E_G$ , then from expressions (5.1) and (5.2) we see that the first set of transitions starts at energies  $E_G - E_p$  and the second set at the slightly higher energy  $E_G + E_p$ . We note that the upward transitions ( $V \rightarrow C$ ) involving the lower optical frequency  $E = E_C - E_V - E_p$  cannot occur if  $n_{\text{VIB}}(E_p) = 0$ , a situation that occurs at absolute zero. In this situation, the lower absorption edge is inverted. The situation is of course different for the higher absorption edge.

To obtain the absorption/emission spectrum we set the temperature to zero. This assumption results in an upper bound on the gain, a result that is entirely acceptable as we are interested in the fundamental limits in the present discussion. The spectrum is governed by the electron energy–wavevector relation corresponding to the lower absorption edge:

$$\int d^3 \vec{k}_C d^3 \vec{k}_V \delta \left( E_G + \frac{\hbar^2 k_C^2}{2m_C} + \frac{\hbar^2 k_V^2}{2m_V} - E - E_p \right) \quad (5.8)$$

The first summation results in:

$$16\pi^2 (m_V/\hbar^2) \int_0^{\frac{1}{\hbar} \sqrt{2m_C(E+E_p-E_G)}} k_C^2 k_V dk_C$$

And a change of variable leads to:

$$2\sqrt{2}\pi^2 (m_V m_C / \hbar^4)^{3/2} \int_{-a}^{a=(E+E_p-E_G)/2} \sqrt{a^2 - x^2} dx$$

The final integration computes to:

$$2\sqrt{2}\pi^2(m_V m_C/\hbar^4)^{3/2} \times \frac{\pi}{8}[E - (E_G - E_P)]^2$$

The indirect absorption edge has a quadratic dependence on the photon energy. For the second absorption edge a similar result holds, the absorption now starting from  $E_G + E_P$ . Lumping all the constants together, the total absorption coefficient can be written as:

$$\alpha_{Si}(\hbar\omega) = B_1(\hbar\omega - E_1)^2 + B_2(\hbar\omega - E_2)^2 \quad (5.9)$$

The constants  $B_1$ ,  $B_2$  and the energies of the absorption edges  $E_1 = E_G - E_P$ ,  $E_2 = E_G + E_P$  can be obtained from a fit of the measured absorption data. We obtain these values from absorption data in reference [3]:

$$\begin{aligned} B_1 &= 400 \text{ cm}^{-1}(\text{eV})^{-2}, & E_1 &= 1.05 \text{ eV} \\ B_2 &= 1300 \text{ cm}^{-1}(\text{eV})^{-2}, & E_2 &= 1.15 \text{ eV} \end{aligned}$$

The absorption edges of silicon suggested by Equation (5.9) should be contrasted to the shape of the absorption edge of a direct semiconductor such as GaAs. In that case the absorption has a square-root dependence on energy (the derivation is similar to the one for silicon) tracking the energy dependence of the density of states:

$$\alpha_{GaAs}(\hbar\omega) = K\sqrt{\hbar\omega - E_o} \quad (5.10)$$

where  $K = 1.17 \times 10^4 \text{ cm}^{-1}(\text{eV})^{-1/2}$  and  $E_o = 1.5 \text{ eV}$  [4].

The problem now is two relate the absorption to gain in silicon. For comparison, we prefer to do this for both silicon and GaAs. We start with silicon considering first the lowest absorption edge. The quantity:

$$f(E_C)[1 - f(E_V)](1 + n_{VIB}) - [1 - f(E_C)]f(E_V)n_{VIB}$$

enters in the  $W^{stim}$  terms in expression (5.6) for the stimulated gain/loss. In the equilibrium case where loss (absorption) dominates ( $f(E_C) \ll 1$  and  $f(E_V) \cong 1$ ) the expression becomes  $-n_{VIB}$ . Under inversion ( $f(E_C) \cong 1$  and  $f(E_V) \ll 1$ ) where gain dominates and the expression becomes  $1 + n_{VIB}$ . This means that, for the lower edge, the gain equals  $(1 + n_{VIB})/n_{VIB}$  times the optical loss. In the same way we find that for the upper edge the ratio is  $n_{VIB}/(1 + n_{VIB})$ . Assuming room temperature and using  $E_P = 50 \text{ meV}$  we have  $n_{VIB} = 0.17$ , suggesting a gain of:

$$g_{Si}(\hbar\omega) = 6.88 \times B_1(\hbar\omega - E_1)^2 + 0.145 \times B_2(\hbar\omega - E_2)^2 \quad (5.11)$$

For GaAs, in a similar manner we find that the gain is  $g_{GaAs}(\hbar\omega) = -\alpha_{GaAs}(\hbar\omega)$ .

The next step is to relate the gain to the carrier density in the conduction band. To do so, we again consider the low-temperature limit where the distribution function of electrons can be taken to equal 1 for energy states below the quasi-Fermi level and equal to 0 for states above the Fermi level. (the opposite holds for electrons in the valence band). In this case the density, increasing monotonically with frequency, determines the energy and hence the frequency dependence of gain. The frequency at which the maximum gain occurs for

degenerate statistics is given by:  $E_{F,C} + E_{F,V} = \hbar\omega - E_G$ . For the quasi-Fermi levels, in the case of low temperature, we have a simple relation to the electron and hole densities [4]:

$$E_{F,C} = \frac{\hbar^2}{2m_e^*}(3\pi^2 N_e)^{2/3}, \quad E_{F,V} = \frac{\hbar^2}{2m_h^*}(3\pi^2 N_h)^{2/3} \quad (5.12)$$

and

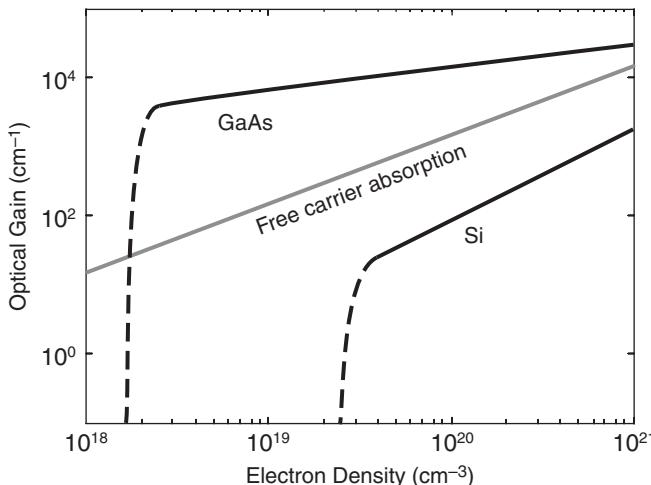
$$E_{F,C} + E_{F,V} = \frac{\hbar^2}{2m^*}(3\pi^2 N)^{2/3} \quad (5.13)$$

where  $N = N_e = N_h$ , and  $1/m^* = (1/m_e^*) + (1/m_h^*)$ . It is evident from Equations (5.11) and (5.13) that the maximum gain in silicon scales with injected carrier density as  $\sim N^{4/3}$ . This is in contrast to the GaAs where we find from Equations (5.10) and (5.13) that the gain scales as  $\sim N^{1/3}$ .

Unfortunately, when carriers are injected to create population inversion, and hence optical gain, there is optical absorption present due to intraband transitions of carriers, often referred to as ‘free-carrier absorption’ (FCA). The free-carrier absorption is described by the Drude model and can be found to depend linearly on the electron–hole pair density  $\sim N$ . A derivation of the resulting optical loss can be found in [5] and will not be repeated here. The numerical constant of proportionality is roughly the same for GaAs and silicon. At the wavelength of 1550 nm the loss per unit length computes to [5]:

$$\alpha_{FCA} = 1.45 \times 10^{-17} \times N(\text{cm}^{-3})(\text{cm}^{-1})$$

In Figure 5.2 we show the optical gain in silicon and GaAs, as well as the competing FCA loss. The values are plotted as a function of injected free carrier density. We see that for GaAs the



**Figure 5.2** Calculated optical gain for silicon and for GaAs as well as the competing free carrier absorption (FCA). Auger recombination, a nonradiative process, has not been included in the calculations. It increases rapidly beyond a carrier density  $\sim 10^{19} \text{ cm}^{-3}$  in silicon

high rate of direct transitions easily overwhelms the FCA loss resulting in significant net optical gain. In silicon however, the low rate of indirect radiative recombination, while constituting a finite gain, is insufficient to overcome the FCA. Consequently, bulk silicon exhibits a net optical loss when population inversion is created. We note that the carrier density at which gain appears depends on the effective density of states in the bands, a quantity that is much lower for GaAs than silicon. This difference also favors GaAs as it relates to the semiconductor's ability to overcome the FCA. We note that the carrier density at which gain appears depends on the effective density of states in the bands, a quantity that is much lower for GaAs than Si. This difference also favors GaAs. Auger recombination, a nonradiative process, becomes significant at carrier densities approaching  $10^{19} \text{ cm}^{-3}$ , with a lifetime that decreases as  $N^{-2}$ . Hence, achieving carrier densities of  $10^{19} \text{ cm}^{-3}$  and beyond will not be practical as it will require excessively large current densities.

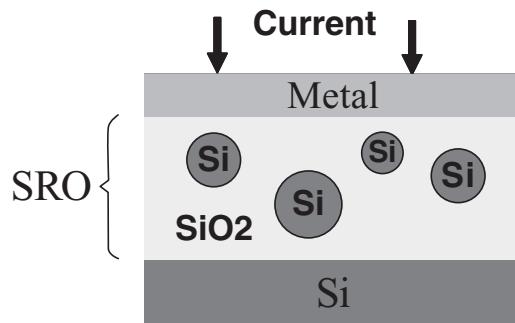
The lesson we draw from the above results is that FCA plays the pivotal role in optical amplification and hence lasing in silicon. In attempt so create a silicon laser, one is well advised not to lose sight of this detrimental, and yet central, effect.

### 5.3 Approach 1: Quantum Confinement

In the next two sections, we describe two novel approaches to overcoming the low radiative emission rate of silicon by using quantum confinement and erbium doping. A number of research groups around the world have made important contributions to this topic, resulting in a large number of publications. The discussion below is not a comprehensive historical survey, but rather a summary of key concepts. For a more comprehensive review, the reader is referred to excellent review papers by Fauchett [6] and Pavesi [7].

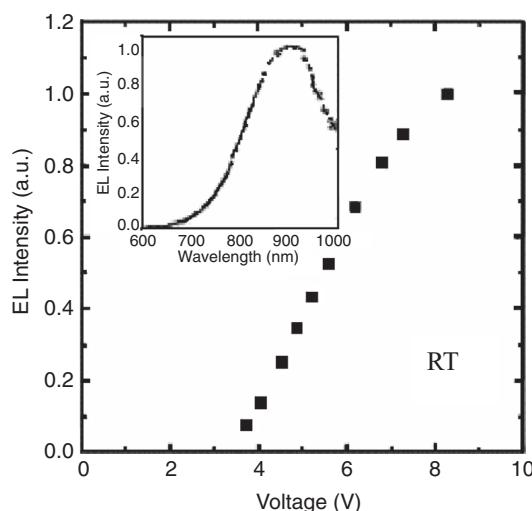
According to the Heisenberg uncertainty principle, when the electron is localized, its momentum becomes uncertain. This phenomenon may offer a solution to the indirect bandgap of silicon. An interesting case study is the semiconductor GaP, which is used for light emitting devices despite its indirect bandgap. The trick is that the momentum conservation requirement is relaxed when an electron is localized at a nitrogen impurity site.

A popular approach to quantum confinement has been the use of silicon nanocrystals that occur naturally in a Silicon-rich oxide (SRO) thin film. When a  $\text{SiO}_x$  ( $x < 2$ ) film is subjected to high temperature anneal, the excess silicon leaves the oxide matrix and forms nanometer-size grains of crystalline silicon dispersed throughout the oxide (Figure 5.3). The nanocrystals are excited by pumping the material with a high-intensity light beam. Electrical pumping has also been demonstrated by creating a metal–oxide–semiconductor structure, shown in Figure 5.3. Electroluminescence is observed in these structures with unipolar (Figure 5.4) [8] and bipolar injection [9]. Several research groups, notably at the University of Trento in Italy and at the University of Rochester, New York, have reported optically pumped gain in these films [10, 11]. However, the observed characteristics cannot be explained on the basis of electron localization in the nanocrystals, and the observations are highly dependent on how the sample was prepared, with some samples showing gain and others exhibiting loss. Consequently, questions remain regarding the nature of the observed optical gain. Additionally, the emission wavelength is in the 800–900 nm range, i.e., outside the two standard telecommunication bands centered at 1320 and 1550 nm.

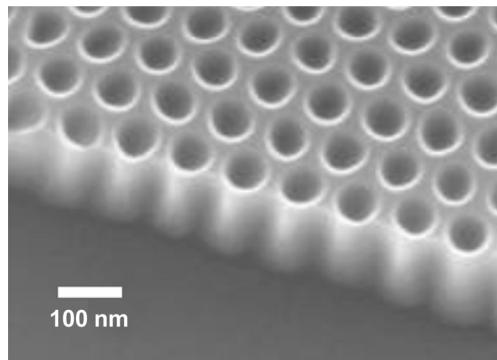


**Figure 5.3** Silicon-rich oxide (SRO) embedded in metal–oxide–semiconductor (MOS) structure used for electrical pumping. The conduction mechanism is electron tunneling through the  $\text{SiO}_2$ . By increasing the conductivity of the  $\text{SiO}_2$ , silicon nano particles reduce the required voltage and hence reduce device degradation and failure over time

Other attempts at exploiting quantum confinement in defect sites or in intentionally formed quantum structures are showing promise. Researchers at Brown University have reported evidence of lasing at cryogenic temperatures ( $\leq 70\text{ K}$ ) [12]. A two-dimensional array of nanometer-size holes was etched into a thin film of silicon that resides on the oxide layer (Figure 5.5). The sample was cleaved (forming mirrors) and pumped optically. As shown in Figure 5.6 the optical input–output characteristics show a threshold behavior for emission at a wavelength of 1278 nm. Attributed to lasing, the threshold behavior occurs at temperatures below 70 K. No such behavior was observed at the secondary emission wavelength of 1135 nm, that originates from the substrate. The emission spectrum measured at a temperature of 10 K



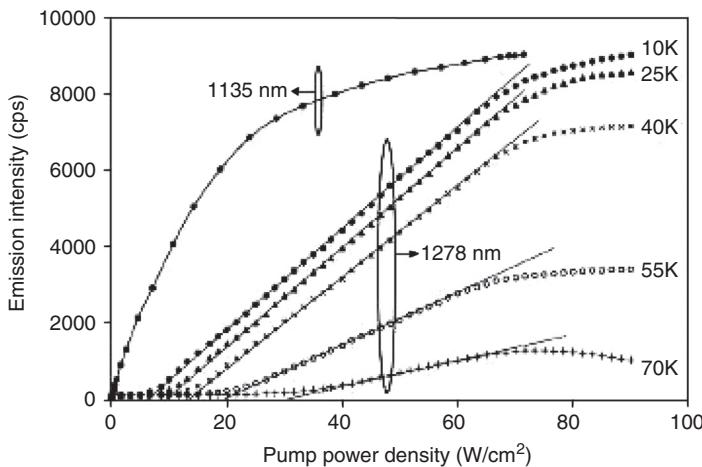
**Figure 5.4** Electroluminescence from silicon-rich oxide (SRO) pumped with an MOS junction. Reproduced from [8] by permission of Elsevier Limited



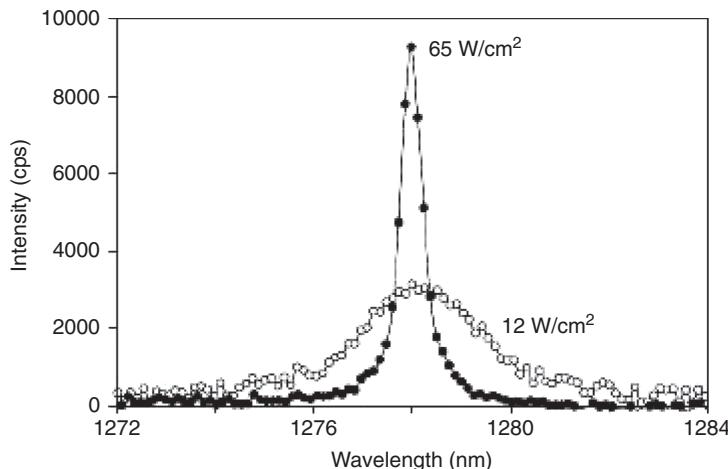
**Figure 5.5** Scanning electron photo of a silicon on insulator film patterned via dry etching. Evidence of lasing was observed at temperatures below 70 K when the film was optically pumped from the edge. Reproduced from [12] by permission of MacMillan Publishers Ltd.

supports the threshold behavior. The spectrum (Figure 5.7) narrows when the pump intensity is increased, a phenomenon that is a characteristic feature of stimulated emission. The mechanism for light emission (1270 nm wavelength) is believed to be from defects on etched silicon surfaces.

In another report, a team consisting of researchers at the University of California, Irvine, and in Taiwan used a junction diode in which the dopants were confined to nanometer size regions (Figures 5.8 and 5.9) [13]. They reported compelling evidence of stimulated emission at room temperature. Optical amplification has not been demonstrated yet, and electrical-to-optical conversion efficiency is low ( $\sim 10^{-4}$ ). Nevertheless, the room-temperature operation and the



**Figure 5.6** Optical input–output characteristics of the nano-patterned silicon on insulator film showing a threshold behavior for emission at a wavelength of 1278 nm. Attributed to lasing, the threshold behavior occurs at temperatures below 70 K. No such behavior was observed at the secondary emission wavelength of 1135 nm that originates from the substrate. Reproduced from [12] by permission of MacMillan Publishers Ltd.

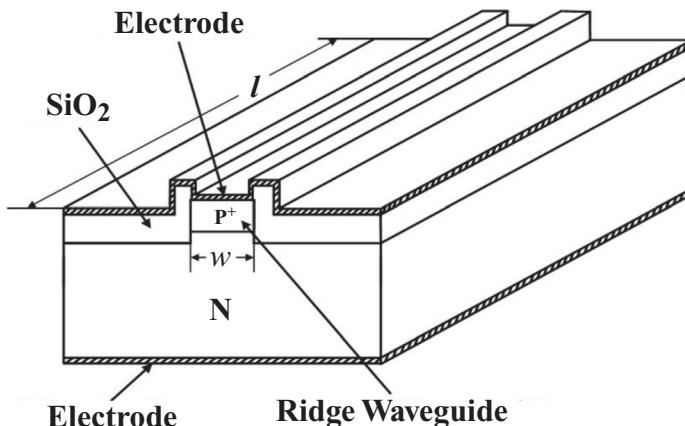


**Figure 5.7** The spectrum of emission from nano-patterned silicon-on-insulator film at a temperature of 10 K. The spectrum narrows when the pump intensity is increased; a classic characteristic of stimulated emission. Reproduced from [12] by permission of MacMillan Publishers Ltd.

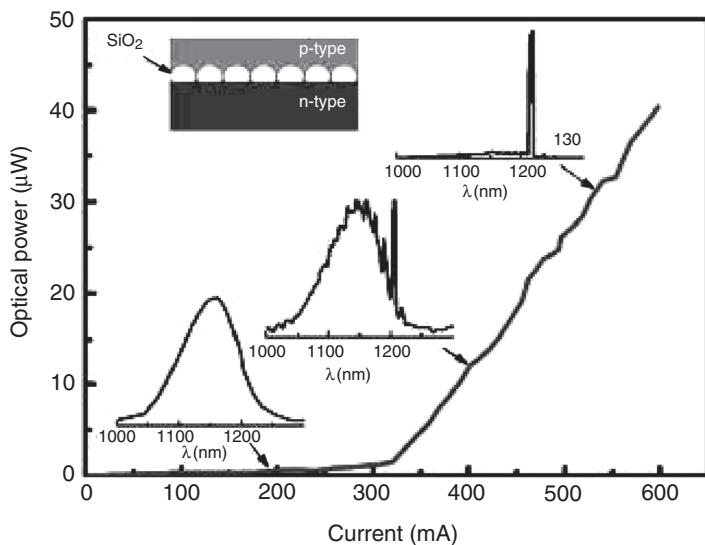
electrical pumping are very important features that warrant additional studies, aimed at reproducing the results and embedding the structure in a tightly confined optical waveguide or cavity.

#### 5.4 Approach 2: Erbium Doping and Other Approaches

The successful realization of light emission and amplification in optical fibers that are doped with erbium has motivated efforts aimed at erbium in silicon. However, silicon is not a good



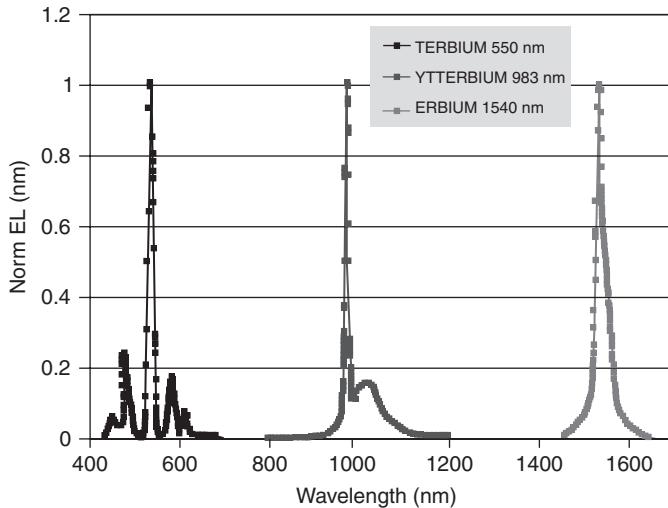
**Figure 5.8** A rib  $p$ - $n$  junction structure where  $p$ -type dopants were diffused from a source with embedded  $\text{SiO}_2$  nanoparticles. The latter are believed to mask the diffusion and results in highly localized nano-scale junctions. Reproduced from [13] by permission of American Institute of Physics



**Figure 5.9** Measured light–current characteristics of an LED with nanoscale  $p$ – $n$  junction regions. The spectral narrowing and the threshold behavior is typically associated with stimulated emission. The inset shows diffusion of  $p$ -type dopants into the  $n$ -type silicon whose surface is masked by  $\text{SiO}_2$  nanoparticles. Reproduced from [13] by permission of American Institute of Physics

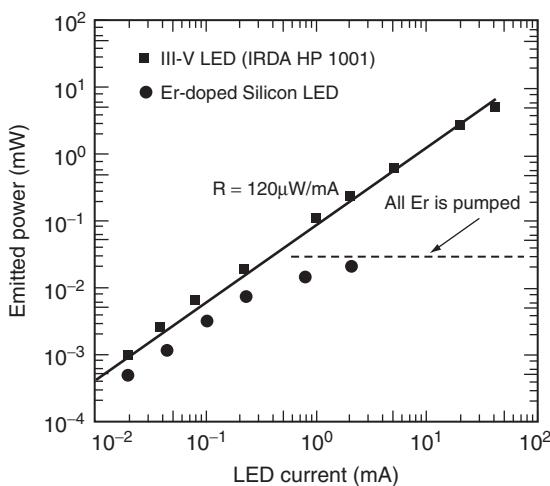
host for erbium, resulting in poor emission at room temperature. The reason is believed to be the back-transfer of energy from the excited Er ions to silicon and also the low concentration of erbium that can be accommodated by silicon.

Optical fiber is made from  $\text{SiO}_2$  (glass) the main constituent of SRO, suggesting that that erbium-doped SRO should be considered as an active optical material. Erbium exhibits light emission at the technologically important wavelength of 1550 nm and other rare earth dopants can be used to achieve emission at different wavelengths (Figure 5.10) [14]. The material can be excited electrically, by sandwiching the film between a silicon and a metal film. A voltage applied to the so-called metal–oxide–semiconductor (MOS) structure causes electrons to tunnel through the oxide, and in the process, to excite the erbium atoms. Silicon nanocrystals also get excited and transfer their energy to nearby erbium ions. Light emitting diodes with efficiencies of about 10%, as high as commercial gallium arsenide devices, have been reported by STMicroelectronics in Italy (Figure 5.11), although with much lower maximum output power of tens of microwatts [14]. Energetic electrons injected into the oxide (by tunneling) cause premature device failure (a link between ‘hot’ electrons and device failure is well established in the electronics industry). Initially, it was believed that the presence of silicon nanocrystals increase the emission cross-section of erbium. However, it has recently been shown that this is not the case and the cross-section is the same as erbium in glass [15]. Nanocrystals then play a more modest role. By increasing the conductivity of the  $\text{SiO}_2$ , silicon nanoparticles reduce the required voltage and hence reduce device degradation and failure over time, albeit at the expense of reduced emission efficiency.

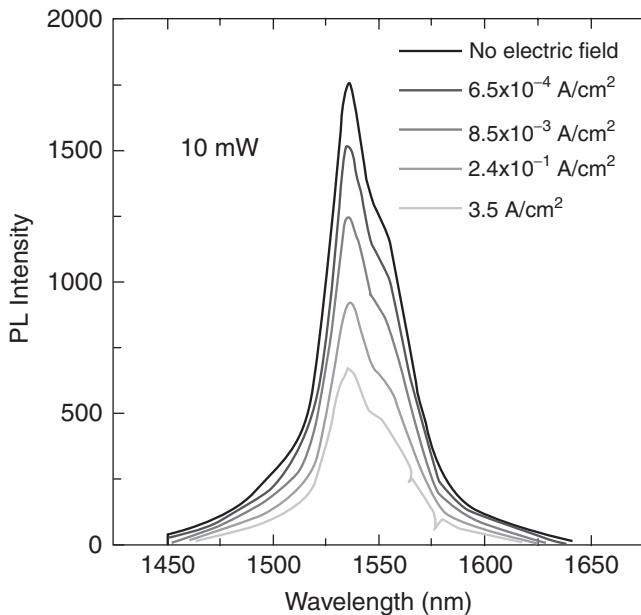


**Figure 5.10** Emission spectrum for SRO doped with various rare earth impurities. Reproduced from [14] by permission of IEEE © (2002)

The attractive features of this technology are the emission in the telecommunication band of 1550 nm as well as the electrical pumping. The main challenge that remains in the path to an electrically pumped laser is highlighted in Figure 5.12 [16] that shows the dependence of photoluminescence intensity on the pump current density. The emission is observed to diminish with current density, a phenomenon that may be due to free carrier absorption and Auger processes, and one that dampens the prospects for achieving electrically pumped optical



**Figure 5.11** Light-current characteristics for erbium-doped SRO diode. The diode exhibits high efficiency, comparable to a GaAs LED, however the output saturates at a much lower value of tens of microwatt. Reproduced from [14] by permission of IEEE © (2002)



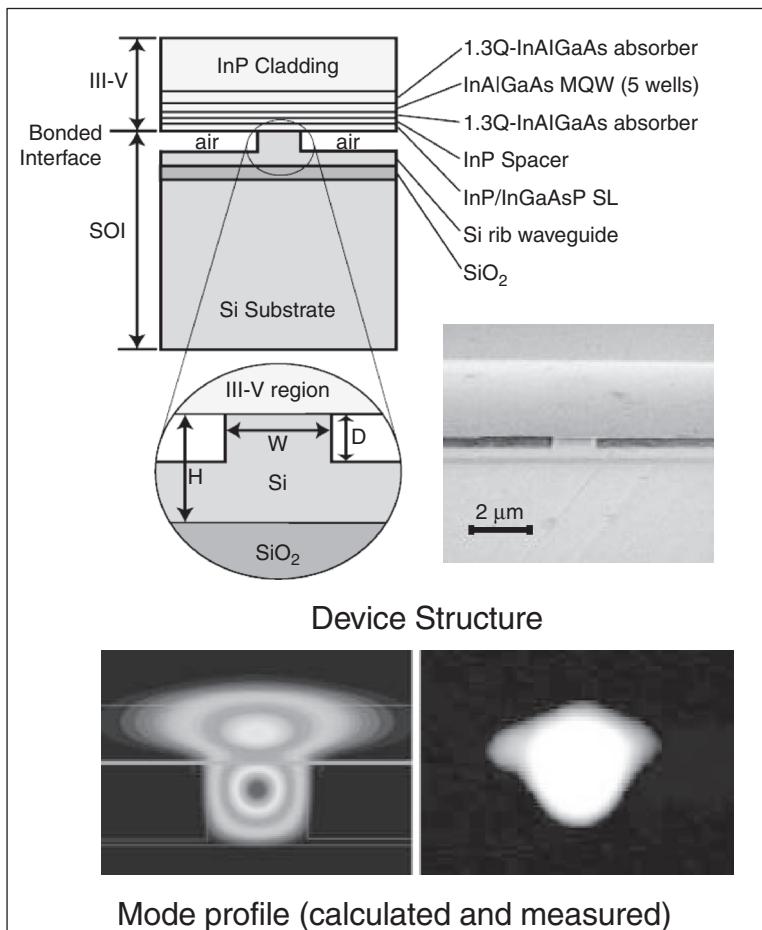
**Figure 5.12** Dependence of photoluminescence intensity on pump current density in an erbium-doped SRO diode. The reduction in emission with current density may be due to free-carrier absorption and Auger effects. Reproduced from [16] by permission of IEEE © (2005)

amplification. This effect notwithstanding, another challenge is achieving a high enough current density using a tunnel junction without subjecting the device to excessive heating and premature device failure. To be sure, the highest current density demonstrated in these devices is an order of magnitude lower than the threshold current density in III–V quantum dot lasers. Nonetheless, the observation of efficient electrical to optical conversion, and optically pumped *internal gain* [17] are important developments and research on means to better understand and to overcome the current limitations is underway.

A simpler approach, and one that mimics erbium-doped waveguide amplifiers (EDWA), is the erbium-doped glass waveguide, which can be deposited on a silicon substrate. In this case, the material is similar to the SRO, except it contains no nanocrystals. To compensate for the low gain per unit length of erbium-doped glass, high-Q (quality factor) microdisk resonators have been used, leading to very low threshold in optically pumped lasers [18].

#### 5.4.1 Hybrid Integration of III–V and Silicon

A near term approach for realizing silicon-based lasers is the heterogeneous integration of III–V and silicon. Direct growth of GaAs on silicon was extensively studied in the 1980s and has recently attracted renewed interest [19]. A noteworthy recent development is the demonstration of an InGaAs quantum dot laser grown directly onto a silicon substrate [20] by Bhattacharya's group at the University of Michigan. Another one is a laser realized by bonding a III–V gain element into a silicon waveguide cavity demonstrated by Bowers' group at UCSB [21]. In

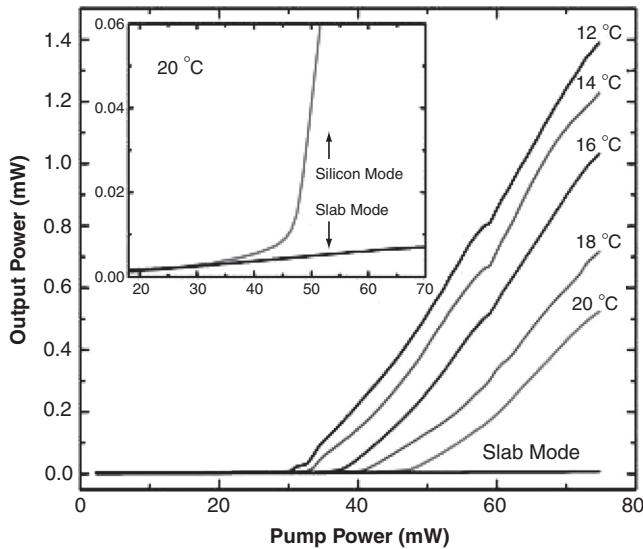


**Figure 5.13** A laser consisting of a III–V gain elements bonded onto a silicon waveguide. Reproduced from [21] by permission of the Optical Society of America

this approach, the III–V semiconductor is the active element and silicon performs a passive function. The laser, shown in Figure 5.13 and 5.14, was optically pumped although electrical pumping is indeed possible. A more established technology, and one that has been widely employed for realizing ultra-compact optical data transmission modules, is to bond not just the gain element, but the complete III–V laser diode (gain element plus high-reflectivity mirrors and metal electrodes) onto a SOI wafer containing silicon waveguide couplers or WDM filters. A recent embodiment that represents the state-of-the-art is the silicon-based 10 Gbit/s optical transceivers being developed by Luxtera Corporation in Carlsbad, California [22].

## 5.5 Approach 3: Raman Effect

In a proposal to the Defense Advance Research Projects Agency (DARPA) in 2001, the authors' group at UCLA proposed the use of stimulated Raman scattering (SRS) as a means to realize



**Figure 5.14** Optically pumped lasing observed in the hybrid III–V/Si laser. Reproduced from [21] by permission of the Optical Society of America

amplification and lasing in silicon. This approach eventually led to the first demonstration of lasing in silicon [23]. Raman back-scattering, using visible and near infrared wavelengths, was used in the late 1960s and early 1970s as an analytical tool to study the vibrational properties of silicon. More recently, SRS has been exploited in optical fibers to create amplifiers and lasers. However, several kilometers of fiber are typically required to create a useful device, suggesting that the approach is not applicable to silicon. Often overlooked was the fact that the gain coefficient for SRS in silicon is approximately  $10^3$ – $10^4$  times higher than that in silica fiber. Additionally, owing to the large refractive index, silicon waveguides can confine the optical field to an area that is approximately 100–1000 times smaller than the modal area in a standard single-mode optical fiber, resulting in proportionally higher Raman gain. When combined, these facts make it possible to observe SRS over the interaction lengths encountered on a chip.

The initial demonstration of spontaneous Raman emission from silicon waveguides in 2002 [24] was followed by the first demonstration of stimulated Raman scattering [25] and parametric Raman wavelength conversion [26] both in 2003. Other merits of Raman effect include the fact that it occurs in pure silicon and hence does not require rare earth dopants (such as erbium), and that the spectrum is widely tunable through the pump laser wavelength.

### 5.5.1 Physics of Raman Scattering in Silicon

Classical electrodynamics provides a simple and intuitive macroscopic description of the Raman scattering process [27]. In the spontaneous scattering, thermal vibrations of lattice at frequency  $\omega_v$  (15.6 THz in silicon) produce a sinusoidal modulation of the susceptibility. The incident pump field induces an electric polarization that is given by the product of the susceptibility and the incident field. The beating of the incident field oscillation ( $\omega_p$ ) with oscillation of the susceptibility ( $\omega_v$ ) produces induced polarizations at the sum frequency,  $\omega_p + \omega_v$ ,

and at the difference frequency  $\omega_p - \omega_v$ . The radiation produced by these two polarization components is referred to as anti-Stokes and Stokes waves, respectively. Quantum statistics dictates that the ratio of Stokes power to anti-Stokes power is given by  $(1 + N)/N$ , where  $N = [\exp(\hbar\omega_v/kT) - 1]^{-1}$  is the Bose occupancy factor and has a value of  $\sim 0.1$  for silicon at room temperature.

The same model can be extended to describe stimulated Raman scattering [27]. Here, one assumes that pump and Stokes fields are present, with a frequency difference equal to the atomic vibrational frequency. The latter can be due to spontaneous emission in a Raman laser, or in the case of a Raman amplifier, it is the input Stokes signal that is to be amplified. The two fields (pump and Stokes) create a force that stimulates atomic vibrations, even in the absence of a dipole moment. This can be understood as follows. If  $E$  is the total field comprising of pump and Stokes, and  $\chi$  is the susceptibility, the energy stored in the field

$$V = \frac{(1 + \chi)}{2} EE^*$$

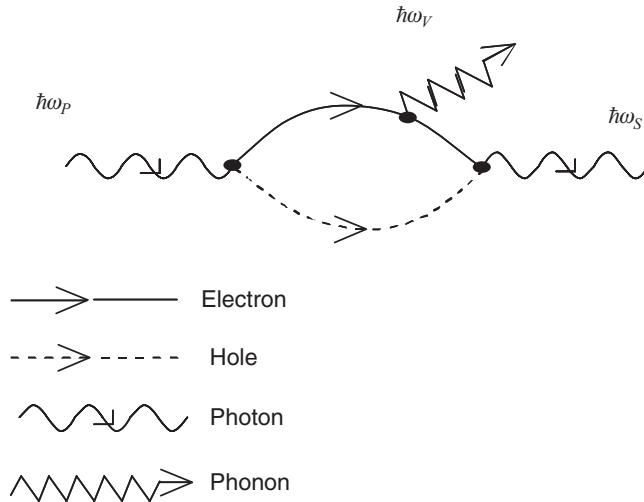
will have a component oscillating at  $\omega_p - \omega_s = \omega_v$ . Through the modulation of the susceptibility with displacement  $Q$ , this will produce a force

$$F \propto \frac{\partial V}{\partial Q} \approx \frac{\partial \chi}{\partial Q} E_p E_s \exp(-\omega_v t)$$

This driving force will enhance atomic oscillations which in turn will increase the amplitude of the Stokes field,  $E_s$ . This positive feedback phenomenon is called stimulated Raman scattering and results in the amplification of the Stokes field.

While providing an intuitively appealing description of Raman scattering, the macroscopic model described above does not account for detailed processes responsible for Raman scattering in silicon. The microscopic picture reveals that the direct coupling of light with atomic vibrations, described by the interaction Hamiltonian involving photons and phonons, is very weak. This is generally true in semiconductors owing to the large atomic mass that appears (squared) in the denominator of the cross-section. In silicon, the lack of dipole moment further underscores this fact. Electrons mediate the Raman scattering process in silicon. Microscopically, the scattering proceeds in three steps [28]. In step one, the incident photon excites the semiconductor into an intermediate step by creating an electron–hole pair. In step two, the pair is scattered into another state by emitting a phonon via the electron-phonon interaction Hamiltonian. In step three, the electron–hole pair in the intermediate step recombines radiatively with emission of a scattered photon. While electrons mediate the process, they remain unchanged after the process. Furthermore, transitions involving electrons are virtual and hence do not have to conserve energy, although momentum must be conserved. The Feynman diagram for the process is shown in Figure 5.15.

The Raman scattering process involves the optical phonon branches of atomic vibrations (as opposed to Brillouin which describes scattering involving acoustic phonons). In first-order scattering, only one phonon is involved and momentum conservation implies that only zone-center phonons can participate. Higher-order Raman scattering involves multiple phonons which can be from any point in the Brillouin zone, as long as their total momentum equals the (negligible) photon momentum, provided that the selection rules allow it. In silicon, the zone center optical phonon is triply degenerate with a frequency of 15.6 THz. The first-order



**Figure 5.15** The Feynman diagram for the Stokes Raman scattering process characterized by the emission of a phonon. The diagram for the anti-Stokes scattering is the same as above, but with the phonon being absorbed rather than emitted

resonance, which is of primary importance here, has a FWHM of approximately 100 GHz [29]. This imposes a maximum information bandwidth of approximately 100 GHz that can be amplified. The Raman linewidth becomes broader when a broadband pump is used.

Crystal symmetry imposes a selection rule that dictates which scattering geometries are allowed. The spontaneous scattering efficiency,  $S$  is given by:

$$S = S_0 \sum_{k=1,2,3} |\hat{e}_s R_k \hat{e}_p|^2$$

Unit vectors  $\hat{e}_p$  and  $\hat{e}_s$  denote the polarization of the pump and Stokes electromagnetic fields.  $S_0$  contains intrinsic microscopic property of silicon including derivatives of the polarizability, and the absolute amplitude of the displacement of the zone-center optical phonons. The sum runs over the three Raman matrices, each corresponding to the phonon displacement along one of the three principle axis of the crystal [30]:

$$\vec{R}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \vec{R}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \vec{R}_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

In the above representation, the correspondence between vectors and crystal directions is  $\hat{e}_1 = (1,0,0) \equiv [100]$ ,  $\hat{e}_2 = (0,1,0) \equiv [010]$  and  $\hat{e}_3 = (0,0,1) \equiv [001]$ . In standard wafer technology the waveguides are fabricated on a (001) surface and are parallel to the [110] direction, so that in an SOI rib waveguide that supports a TE and a TM mode and is fabricated

along [110] the polarization directions are

$$\hat{e}_{TE} = \frac{1}{\sqrt{2}}(1, -1, 0)$$

and

$$\hat{e}_{TM} = (0, 0, 1).$$

The Raman gain coefficient  $g_R$  can be obtained from the spontaneous efficiency  $S$  using the Einstein relation, as [31]:

$$g_R = \frac{8\pi c^2 \omega_p}{\hbar \omega_s^4 n^2(\omega_s)(N+1)\Delta\omega} S$$

Substituting the appropriate values, the gain coefficient is obtained as  $\sim 76$  cm/GW [24]. This uses  $S = 8.4 \times 10^{-7}$  cm $^{-1}$  Sr $^{-1}$  which was obtained by extrapolating the values measured 1.1  $\mu$ m wavelength to 1.55 using a  $\lambda^{-4}$  relation. This is in the same order of magnitude, but several times larger than the values extracted from Raman gain measurements ( $\sim 20$  cm/GW) at 1.55  $\mu$ m [25], nonetheless, when compared with silica ( $0.93 \times 10^{-2}$  cm/GW), the Raman gain in silicon is  $10^3$ – $10^4$  times larger. Such a large difference has its origin in the much narrower linewidth of the Raman spectrum in *crystalline* silicon, compared with the *amorphous* fiber.

It is customary to describe the induced polarization for the case of stimulated Raman scattering through the nonlinear susceptibility,  $\chi_{ijmn}^R$ , defined by the following expression:

$$P_i^{\text{NL}}(\omega_s) = \epsilon_0 \chi_{ijmn}^R E_j(\omega_p) E_m(-\omega_p) E_n(\omega_s)$$

On the other hand, the atomic displacement can be obtained using a classical harmonic oscillator model [27] with the driving force described above. By comparing the induced polarization suggested by the displacement with the definition above, one arrives at the following expression for the induced Raman susceptibility,

$$\chi_{ijmn}^R = 2\Gamma \omega_v \frac{2ncg_R}{\omega_s(\mu_0/\epsilon_0)^{1/2}} \frac{\sum_{k=1,2,3} (R_{ij})_k (R_{mn})_k}{(\omega_v^2 - (\omega_p - \omega_s)^2 - 2i\Gamma(\omega_p - \omega_s))}$$

Where  $\Gamma$  is the dissipative term in the harmonic oscillator equation and  $n$  is the refractive index. Crystal symmetry considerations, described by the Raman tensor,  $R$ , lead to a total of 12 equal nonvanishing components that have the indices of the form:

$$1221 = 1212 = 2112 = 2121 = 1331 = 1313 = 3113 = 3131 = 2332 = 2323 = 3223 = 3232$$

The induced susceptibility is related to the Raman gain coefficient as [30]:

$$\chi_{1221}^R(\omega_p - \omega_s = \Omega) = \frac{i}{(\mu_0/\epsilon_0)^{1/2}} \frac{2ncg_R}{\omega_s} = 11.2 \times 10^{-14} i \frac{\text{cm}^2}{\text{V}^2}$$

Two-photon absorption (TPA) is another nonlinear optical effect that is particularly strong in semiconductors. This effect results in pump depletion and generation of the free carriers

that, through the free carrier plasma effect, give rise to a broadband absorption spectrum. TPA has been shown to be negligible from the point of view of pump depletion [25]. This is plausible since the TPA coefficient in silicon  $\beta$  is relatively small,  $\sim 0.5 \text{ cm/GW}$ . On the other hand, absorption by TPA-generated free carriers is a broadband process that competes with the Raman gain. The effect has been identified as a limiting factor in all-optical switching in III–V semiconductor waveguides [32–36]. It has also been discussed as a potential limit to achievable Raman gain in GaP waveguides [37], although a Raman gain of 24 dB was demonstrated in these waveguides [38]. More recently, TPA-induced FCA has been studied in silicon waveguides in the context of Raman process [39, 40] and in transmission of ultra-short pulses in silicon waveguides [41, 42].

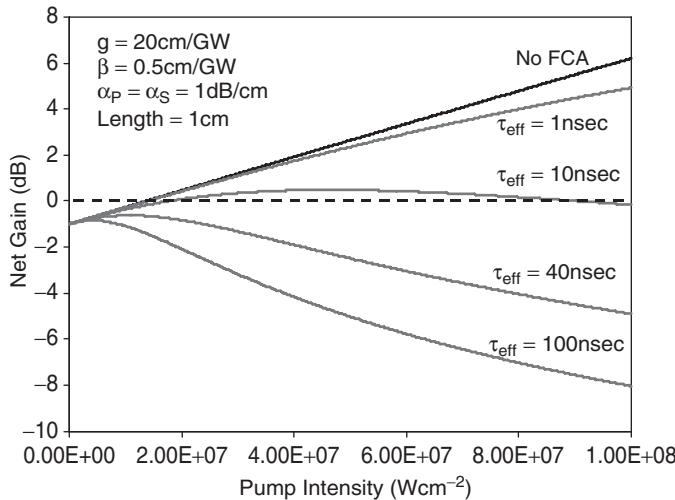
The magnitude of TPA-induced free-carrier absorption depends on the free carrier concentration through the relation:  $\alpha^{\text{FCA}} = 1.45 \times 10^{-17} (\lambda/1.55)^2 \Delta N$ , where  $\lambda$  is the wavelength, in microns, and  $\Delta N$  is the density of electron–hole pairs [43]. The latter is related to the pump intensity  $I_p$  by:

$$\Delta N = \beta I_p^2 \tau_{\text{eff}} / (2h\nu)$$

Where  $h\nu$  is the pump photon energy, and  $\tau_{\text{eff}}$  is the effective recombination lifetime for free carriers. This equation neglects the contribution to free carrier generation due to pump–signal TPA and hence is valid in the regime where Stokes intensity is much smaller than the pump intensity.

The fundamental parameter that governs the TPA-induced loss, and hence the success of Raman-based devices is the recombination lifetime  $\tau_{\text{eff}}$ . It is well known that the recombination lifetime in SOI is much shorter than that in a bulk silicon sample with comparable doping concentration. This lifetime reduction is due to the presence of interface states at the boundary between the top silicon and the buried oxide layer. This effect depends on the method used for preparation of the SOI wafer and the film thickness, with measured and expected values ranging between 10 and 200 ns [45, 46]. In SOI waveguides the lifetime is further reduced to a few nanoseconds, or even below in the case of submicrometer waveguides, due to the recombination at the etched waveguide facets and, in the case of rib waveguides, due to diffusion into the slab regions [47, 48]. The lifetime can be further reduced by application of a reverse-bias  $p$ – $n$  junction [39, 40, 47] or by introduction of midgap states through high-energy irradiation, and gold or platinum doping. A modest amount of CW gain has been observed in deep submicrometer waveguides by Osgood's group at Columbia University [48], where the impact of surface and interface recombination plays a critical role in reducing the lifetime. CW gain has also been demonstrated by sweeping the free carriers using a reverse-bias  $p$ – $n$  junction [49]. This approach is further discussed later in a separate subsection below.

The plot of the net Raman gain as a function of CW pump intensity for a waveguide of length  $L = 1.9 \text{ cm}$ , and propagation loss  $1 \text{ dB/cm}$ , is shown in Figure 5.16 [40] for different free-carrier lifetime values. The plot shows that more than 5dB of gain can be obtained with a pump intensity of less than  $100 \text{ MW/cm}^2$ . Gain increases with intensity while the loss rises as intensity squared and dominates when lifetime is long. The pump is assumed to be a monochromatic source. The finite linewidth of the pump laser will result in a lower gain than



**Figure 5.16** Impact of carrier lifetime on achievable CW Raman gain. Gain increases with intensity while loss rises as intensity squared and dominates when lifetime is long

what is predicted in Figure 5.16. It is clear that to create a successful amplifier, an effective lifetime of  $\leq 1$  ns is required.

A detailed theoretical study on Raman scattering in silicon waveguides using the coupled-mode theory formalism was published in 2003 [30]. In this work, waveguides of varying cross-sections were analyzed to obtain the gain coefficient for the Raman amplification process. In addition to this, it was found that the lack of interaction between cross-polarized Stokes being amplified by the pump could be used to build a polarization-multiplexed Raman amplifier. Following this, the CW Raman lasers were studied theoretically with the main saturation mechanisms attributed to free-carrier absorption processes [50]. The variation of carrier lifetime in silicon waveguides for varying dimensions and the role of a reverse-biased carrier sweeping scheme were studied using two-dimensional electronic simulations [47]. Following this, the limitations imposed by the reverse-biased carrier sweeping scheme in effectively sweeping carriers due to carrier screening effects were analyzed [51]. This is discussed in detail in Sections 5.5.2 and 5.5.3. More recently, there has also been theoretical work on understanding the pulsed Raman amplification process in silicon wire waveguides [52] and the prospects of bidirectional pumping to mitigate the free-carrier absorption processes [53]. Additionally, it has been shown that a cladding pump waveguide geometry offers a higher total gain, albeit at the expense of increased device length [54].

### 5.5.2 Limitation of Active Carrier Removal Using a p–n Junction

In this section, the physical limitations of carrier lifetime in Si are discussed. In particular, the intensity dependence of the carrier lifetime in a rib waveguide in which a reverse-biased  $p$ – $n$  junction is used to sweep out the carriers is studied. The main challenge in Raman lasers and amplifiers is the nonlinear optical loss that competes with the gain. The optical absorption

loss is due to free carriers that are created in the waveguide because of two-photon absorption (TPA) of the pump beam. The rate of free-carrier losses for pump and signal intensities scale as  $I_p^3$  and  $I_p^2 I_s$ , respectively. Since the rate of Raman amplification is proportional to  $I_p I_s$ , the nonlinear loss can be detrimental. The magnitude of the nonlinear optical loss is proportional to the carrier density that has been created in the waveguide, which is in turn proportional to the time it takes for carriers to be removed from the waveguide core. The carrier lifetime is therefore the critical parameter for CW operation, or for pulsed operation when the pulse repetition period is shorter than the carrier lifetime.

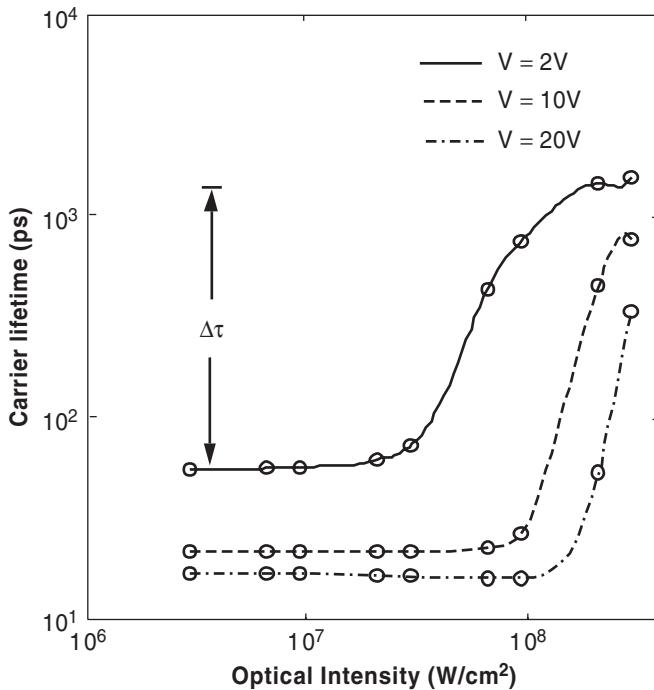
The lifetime is determined by the combination of diffusion and interface/surface recombination currents. In a bare SOI waveguide, the geometry of the waveguide plays a significant role in determining the carrier lifetime, as an order of magnitude reduction can be achieved by varying the slab height and the rib-to-slab height ratio [47]. The carrier lifetime also depends on the generation rate in the waveguide and therefore on the intensity of the pump beam.

We have used a commercial software (ATLAS, ATHENA v. 5.6.0.R Silvaco International) to find the carrier lifetime in a rib waveguide with a straddled  $p-n$  junction. The waveguide supports an optical mode with an effective area  $A_{\text{eff}}$ , of  $2.3 \mu\text{m}^2$ . The details of the structure and parameters used can be found elsewhere [47]. The carrier lifetime  $\tau$  was determined as a function of the applied voltage  $V$  and the optical intensity. The total number of carriers generated by TPA per second per unit length is:  $(\beta I_p^2 / 2E_p) A_{\text{eff}}$ , where,  $E_p$  is the photon energy. The generation rate  $G$  in the simulation is assumed to be uniform over the rib cross-section such that the total number of generated carriers is  $GwH$ . In this framework, the carrier lifetime is given by:

$$\tau = \frac{1}{GwH} \int_0^H \int_{-w/2}^{w/2} N dx dy$$

where  $N$  is the carrier concentration. In Figure 5.17, we show the effective lifetime  $(\tau_e + \tau_h)/2$ , for applied voltages of 2, 10 and 20 V, as a function of the intensity. The most important feature in this plot is that, for a given voltage, when the intensity exceeds a critical value, the lifetime rapidly increases. As the intensity is further increased the lifetime growth saturates. Furthermore, an increase in the reverse bias voltage shifts the critical intensity that must be exceeded for the lifetime to vary to higher values [51].

We can understand the essential features by examining the case of a  $p-n$  junction where there is a uniform generation rate in the depletion region. When no generation is taking place, the applied field  $E_o$ , is proportional to  $V$  and inversely proportional to the width of the depletion region. Now if an electron-hole pair is generated, the particles will drift under the action of the applied electric field. We note that the carrier drift velocity in silicon has a nonlinear dependence on the applied electric field due to velocity saturation. For electric fields smaller than  $E_{\text{sat}} \approx 2 \times 10^4 \text{ V/cm}$ , the carrier velocities increase linearly with the applied field as  $v_e = \mu_e E$ , and  $v_h = \mu_h E$  ( $\mu_e = 1500 \text{ cm}^2 \text{ s}^{-1} \text{ V}^{-1}$  and  $\mu_h = 450 \text{ cm}^2 \text{ s}^{-1} \text{ V}^{-1}$ ). For higher electric fields the carrier velocity is saturated. For electrons, this saturation velocity is  $v_{e,\text{sat}} \approx 10^7 \text{ cm/s}$ . The electric field can exceed  $E_{\text{sat}}$ , but it cannot exceed the material breakdown field  $E_B \sim 3 \times 10^5 \text{ V/cm}$ . When the generation rate is low enough, the generated carriers are swept out of waveguide core and the carrier lifetime is therefore the mean transit time through the rib region with width  $w$ , i.e.,  $\tau_e = w/(2\mu_e E_o)$  when  $E_o < E_{\text{sat}}$  and  $\tau_e = w/(2v_{e,\text{sat}})$  when  $E_o > E_{\text{sat}}$ . Similar expressions hold for the holes.

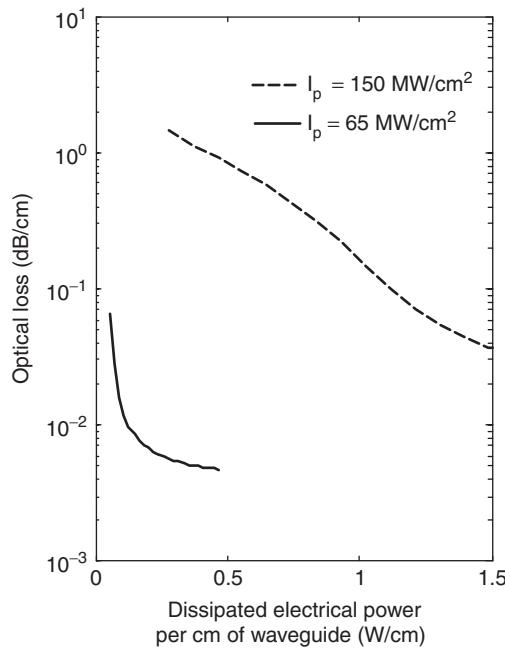


**Figure 5.17** Dependence of the effective carrier lifetime on the optical intensity for various values of the applied voltage. Reproduced from [51] by permission of American Institute of Physics

Furthermore, since the electrons and holes drift in opposite directions, a secondary electric field will be set up in the depletion region with direction opposite to the applied field. The applied electric field will be screened by the secondary field, leading to an increase, of the carrier lifetime, as observed in Figure 5.17. The screening of the applied field by the carriers is the same effect that is responsible for the saturation of responsivity in high-power photodetectors. In the regime in Figure 5.17 where saturation has set in, drift is negligible and carrier transport is governed by diffusion and recombination. In this case the lifetime should approach the lifetime of carriers in a waveguide with no junction. Therefore, the difference in lifetime at the low- and high-intensity regimes in Figure 5.17 ( $\Delta\tau$ ) is the lifetime reduction enabled by the junction. We note that the generation rate along the waveguide varies slowly in comparison to that in the transverse direction. Therefore, in calculations of carrier lifetime, longitudinal carrier diffusion can be neglected.

### 5.5.3 Electrical Power Dissipation Caused by Active Carrier Sweep-out

The reduction in lifetime and the concomitant reduction in optical loss that is achieved with the  $p-n$  junction come at the expense of on-chip electrical power dissipation [51]. The power dissipation per unit length is a strong function of pump intensity which varies along the waveguide length. Neglecting the recombination of carriers (this is permissible when the



**Figure 5.18** Dependence of the FCA loss on the dissipated electrical power that is necessary to achieve that loss. Results are shown for optical intensities of 150 and 250 MW/cm<sup>2</sup>. Simulations were limited to a maximum applied voltage of 25 V. Reproduced from [51] by permission of American Institute of Physics

dominant current is the drift current) the current (per unit length of the waveguide) is simply determined by the generation rate as  $j = e(G \times w \times H)$ . The power dissipation per unit length of the waveguide is  $W_{\text{diss}} = jV$  where  $V$  is the applied voltage. Figure 5.18 shows how the FCA loss varies with the power dissipated to sweep out the carriers for intensities 65 and 150 MW/cm<sup>2</sup>. Simulations were limited to a maximum applied voltage of 25 V. The FCA loss is determined from the average electron–hole density  $N = G(\tau_e + \tau_h)/2$  as  $a_{\text{FCA}} = \sigma N$  where  $\sigma = 1.45 \times 10^{-17} \text{ cm}^2$  is the FCA loss cross-section. We note that the power dissipation per unit length is a strong function of pump intensity which varies along the waveguide length. Nonetheless, Figure 5.18 indicates that the reduction in nonlinear absorption comes at the expense of significant on-chip electrical dissipation. The most direct path to power reduction is lateral scaling of device dimensions, as indicated in the expression for the lifetime when the generation rate can be neglected.

To summarize this section, the lifetime dependence on the intensity is due to the screening of the applied field by the photo-generated carriers, and can result in an increase in the lifetime by more than an order of magnitude. Hence this effect is expected to have significant implications for operation of silicon Raman lasers and amplifiers that employ carrier sweep-out via a  $p$ – $n$  junction. In addition, the benefits gained by free-carrier loss reduction with reverse-bias come at the expense of increased on-chip electrical power dissipation.

A novel means to avoid power dissipation is discussed later in Section 5.9. This is done by realizing a  $p$ – $n$  junction diode in which sufficient carrier sweep-out occurs at zero voltage.

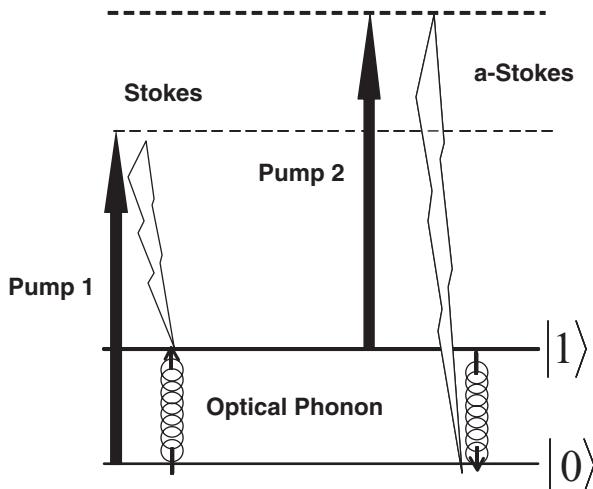
In fact, we will show that negative power dissipation and CW Raman gain can be achieved simultaneously. This approach invokes the photovoltaic effect caused by TPA, and is applicable to any photonic device based on nonlinear interactions in silicon. These include Raman amplifiers/lasers, as well as devices that exploit Raman- or Kerr-based four-wave mixing.

#### 5.5.4 Raman Wavelength Conversion

The Raman scattering spectrum also contains an anti-Stokes wave that is up-shifted from the pump by the 15.6 THz phonon frequency. The gain coefficient for the anti-Stokes wave will have a negative sign, indicating that an incident anti-Stokes wave will be attenuated. However, an anti-Stokes signal can be generated through four-wave mixing (FWM) induced through the Raman susceptibility, in much the same way that conventional FWM takes place via the electronic third-order electronic nonlinear susceptibility (responsible for the Kerr effect) [26, 55, 56]. In the Raman process, the energy conservation dictates that  $\omega_{\text{as}} = 2\omega_p - \omega_s$ , and momentum conservation results in the so-called phase-matching condition, with the total phase mismatch defined as:

$$\Delta\beta = 2\beta_p - \beta_s - \beta_{\text{as}}$$

where,  $\beta$ , is the wavevector for the given wavelength, and the corresponding mode of polarization could be TE or TM. As  $\Delta\beta$  approaches zero, the pump, Stokes and anti-Stokes waves experience coherent interaction. This phenomenon is referred to as coherent anti-Stokes Raman scattering (CARS) in the spectroscopy literature and was first observed in silicon in 2003 [26]. As shown in Figure 5.19, the process the creation of the anti-Stokes photon is accompanied



**Figure 5.19** Energy-level representation of Raman wavelength conversion process.  $|1\rangle$  and  $|0\rangle$  are vibrational states and arrow represent virtual transitions. A phonon is created and annihilated, leaving the phonon population unchanged

by creation and annihilation of a coherent zone center phonon which transfers information between the Stokes and anti-Stokes waves.

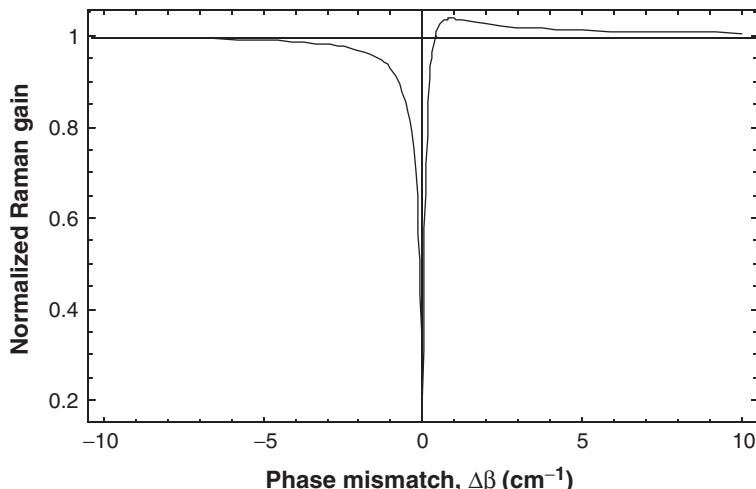
The conversion efficiency is highly sensitive to the phase mismatch and, in general, the efficiency has a  $\text{sinc}^2$  dependence on phase mismatch. In silicon, phase mismatch is dominated by the material dispersion, as waveguide dispersion is relatively negligible, except in waveguides with submicrometer modal dimensions. In such devices, waveguide dispersion provides a mean to compensate for material dispersion [56]. Other means of phase matching include the use of waveguide birefringence and/or strain [56]. At phase matching, the evolution of Stokes and anti-Stokes fields  $E(z)$  along the waveguide length  $z$  is given by,

$$E_S(z) = E_S(0) + (E_S(0) + E_{\text{aS}}^*(0))g_R I_P z / 2$$

$$E_{\text{aS}}^*(z) = E_{\text{aS}}^*(0) - (E_S(0) + E_{\text{aS}}^*(0))g_R I_P z / 2$$

The above equation predicts a linear increase in the fields with distance which holds true for small propagation lengths. Once the Stokes and anti-Stokes fields become equal in amplitude, no further change takes place, leading to a saturation effect. The characteristic length is very long, therefore, this regime is not expected to occur in chip-scale devices.

Figure 5.20 shows the normalized Raman gain as a function of phase mismatch. At large values of  $|\Delta\beta|$ , stimulated Raman amplification is the predominant effect and leads to the amplification of the Stokes signal with an effective gain given by  $g_R$ . Under the phase matching condition, Raman gain is suppressed and the parametric coupling of Stokes and anti-Stokes waves dominates. In this region, Stokes, anti-Stokes, and pump fields are strongly coupled and parametric conversion dominates. For small, positive values of  $\Delta\beta$ , the normalized gain slightly exceeds unity due to modulation instability.



**Figure 5.20** The variation of normalized Raman gain as a function of phase mismatch. At zero phase mismatch, Raman gain is suppressed in favor of parametric Stokes to anti-Stokes conversion

**Table 5.1** A summary of pertinent nonlinear optical effects in silicon and comparison of their relative strengths

Nonlinear effect	Description	Strength of non-linearity	Spectral Properties
Stimulated Raman scattering	Stimulated interaction of optical phonons with the Stokes and a strong pump field	$g_R = 20 \text{ cm/GW}$ (at 1550 nm)	Shift = 15.6 THz (fixed) Bandwidth = 105 GHz (fixed) Gain scales as $1/\lambda$
Stimulated Brillouin scattering	Stimulated interaction of acoustic phonons with the Stokes field and a strong pump field	cm/MW(at 1550 nm)	Shift $\sim$ 40 GHz (at 1550 nm), scales as $1/\lambda$ Bandwidth = $\Delta f$ (MHz) not measured
<i>Nonlinear Absorption</i>			
Two-photon absorption	Absorption of 2 photons in succession	$\beta_2 = 0.5 \text{ cm/GW}$ (at 1550 nm)	Similar profile as linear absorption. Cutoff at 2.2 $\mu\text{m}$ (half bandgap)
Three-photon absorption	Absorption of 3 photons in succession	$\gamma_3 = \text{much weaker than } \beta_2$	Similar profile as linear absorption. Cutoff at 3.3 $\mu\text{m}$ (1/3 bandgap)
Free-carrier absorption	Absorption due to free carriers present/created in the semiconductor	$\alpha_{FCA} = \sigma \Delta N \sim 1 \text{ cm}^{-1}$ (at $\Delta N$ of $10^{17}/\text{cm}^3$ )	Broadband absorption process. Scales as $\lambda^2$

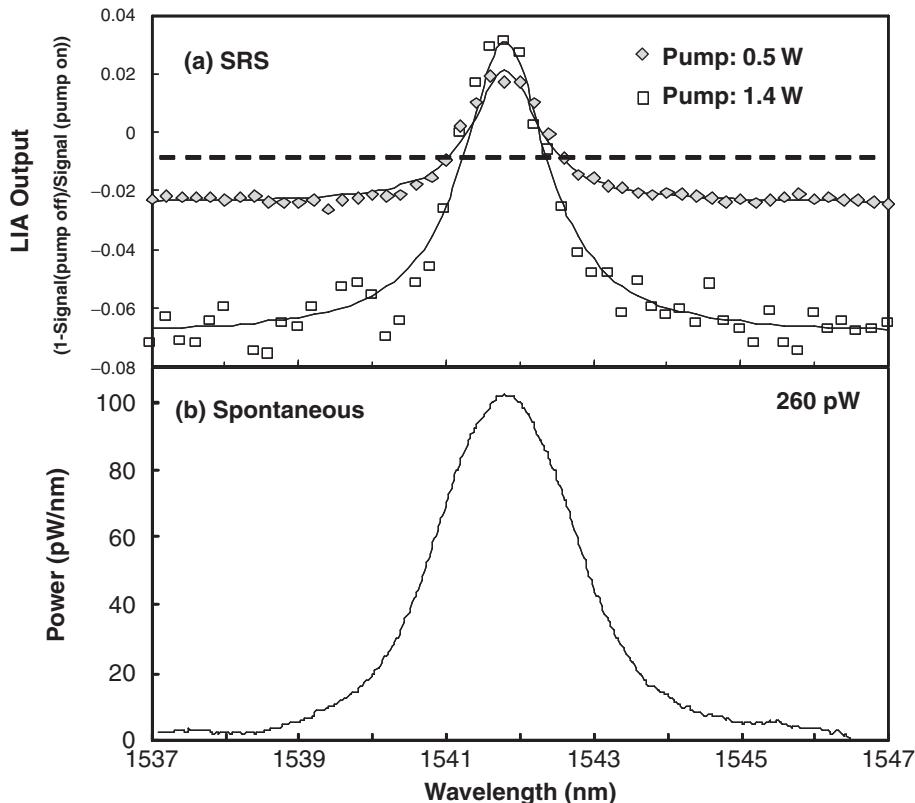
### 5.5.5 Comparison of Raman with Other Optical Nonlinearities in Silicon

A summary of the relevant nonlinear optical effects along with their strength and bandwidth is presented in Table 5.1. It is interesting to note that, serendipitously, the Raman scattering process is the strongest among these effects.

## 5.6 Experimental Realization of Raman-based Silicon Photonic Devices

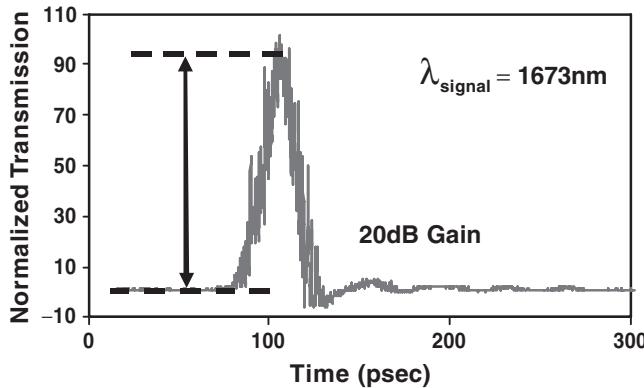
Experimental results shown in Figure 5.21 highlight the competition between Raman gain and TPA-induced free-carrier absorption [57]. Here, a CW pump emitting at 1427 nm was used along with a tunable CW signal laser. The plot show the amplification of the signal laser as it is tuned across the Raman resonance. The data clearly show the competition between the resonant Raman amplification and the broadband pump-induced absorption. It is clear that further increase in pump intensity is futile as the intensity-squared increase of TPA-induced loss dominates over Raman gain which increases linearly with pump intensity. Also shown in Figure 5.21 is the measured spontaneous emission spectra from the same waveguides. The observed FWHM is approximately twice the 100 GHz intrinsic Raman linewidth, a feature caused by the finite linewidth of the pump laser.

One method for avoiding free-carrier accumulation and the concomitant loss is pulsed pumping. As long as the pump pulse period is longer than the free-carrier lifetime the



**Figure 5.21** (a) Measured spectral characteristic of Stimulated Raman scattering (SRS) in an SOI waveguide, under CW pumping. Two different input pump powers are shown, to illustrate the effect of TPA-induced free carrier absorption. (b) Measured spontaneous emission spectrum. LIA: lock-in amplifier

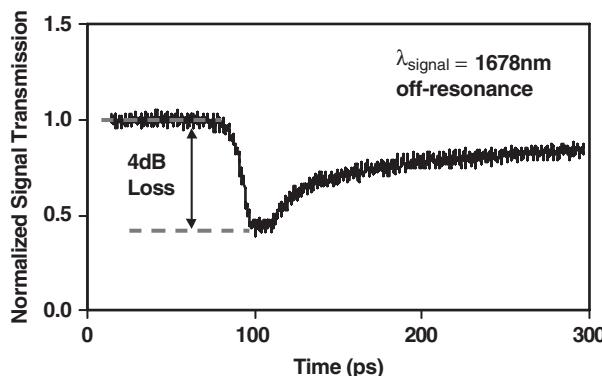
free-carrier accumulation can be mostly eliminated. Figure 5.22 shows the measured change in CW signal beam (tuned to the peak of the Stokes resonance) due to the pump pulse. The pump source was a mode-locked fiber laser with 25 MHz repetition rate and 1 ps output pulse width. Since phonon response time in silicon is more than 3 ps, the pulse width is broadened by using a spool of standard single-mode optical fiber. A maximum pump on-off gain of 20 dB has been obtained. Taking into account the losses in the waveguide, we obtain a net waveguide gain of 13 dB [58]. This gain includes waveguide propagation losses, but not the fiber-waveguide coupling loss. In separate devices that have adiabatic mode tapers, we have shown net fiber-to-fiber gain of 11 dB under similar pulse pumping scheme [59]. The net free-carrier loss and the free-carrier lifetime can be measured by performing the same measurement with the signal laser tuned away from the Raman resonance. This is shown in Figure 5.23. Maximum loss due to the combined FCA loss is measured to be 4 dB. Thus, the intrinsic Raman gain in the Silicon waveguide is 24 dB. By extrapolating the exponential decay of the carriers we estimate a free carrier lifetime of  $\sim 4$  ns.



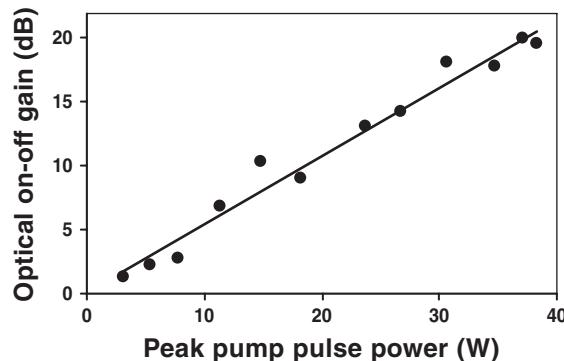
**Figure 5.22** Time-resolved Raman amplification with the signal laser at 1673 nm. A pump on–off gain of 20 dB is obtained. Pump pulse wavelength is 1540 nm

The variation of optical gain as a function of peak pump power coupled into the waveguide is shown in Figure 5.24. Optical gain is found to saturate around 37 W of peak pump power. This can be attributed to the pulse breakup and excessive spectral broadening of pump laser in the fiber pigtail preceding the waveguide [59].

Figure 5.25 shows the three-dimensional plot of the Stokes-to-anti-Stokes conversion through the CARS process [60]. The pump laser (at 1427 nm) is coupled into the  $\text{TE}_0$  mode, and the signal laser is coupled into the  $\text{TM}_0$  mode. The Stokes signal laser is scanned in a range from 1530 to 1560 nm. The figure shows the anti-Stokes spectra ( $\text{TM}_0$  polarized) measured as a function of the Stokes signal wavelength. The CW pump power in the waveguide was 0.7 W. There is a clear peak at 1328.8 nm of anti-Stokes emission when the Stokes laser is tuned to Stokes wavelength of 1542.3 nm. The nature of the ‘satellite’ peaks maybe due to the  $\text{sinc}^2$  dependence of the conversion efficiency with phase mismatch.



**Figure 5.23** Time-resolved signal loss with the signal laser at 1678 nm, i.e. outside the Raman resonance. Maximum loss of 4 dB and carrier lifetime of 4 ns is obtained

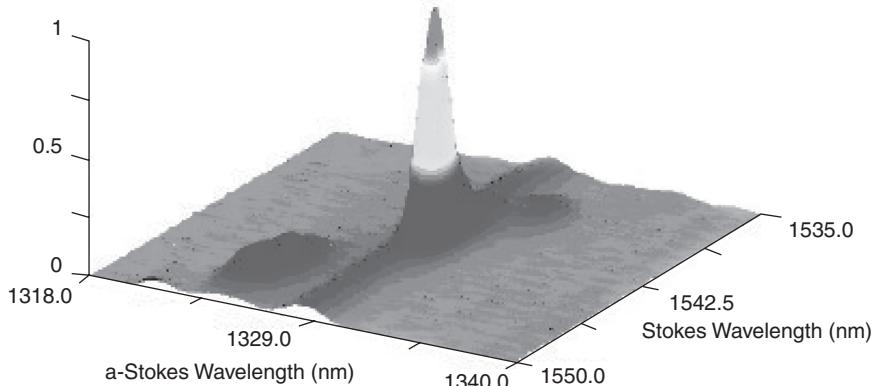


**Figure 5.24** On–off optical gain as a function of peak pump pulse power. Maximum gain of 20 dB is obtained

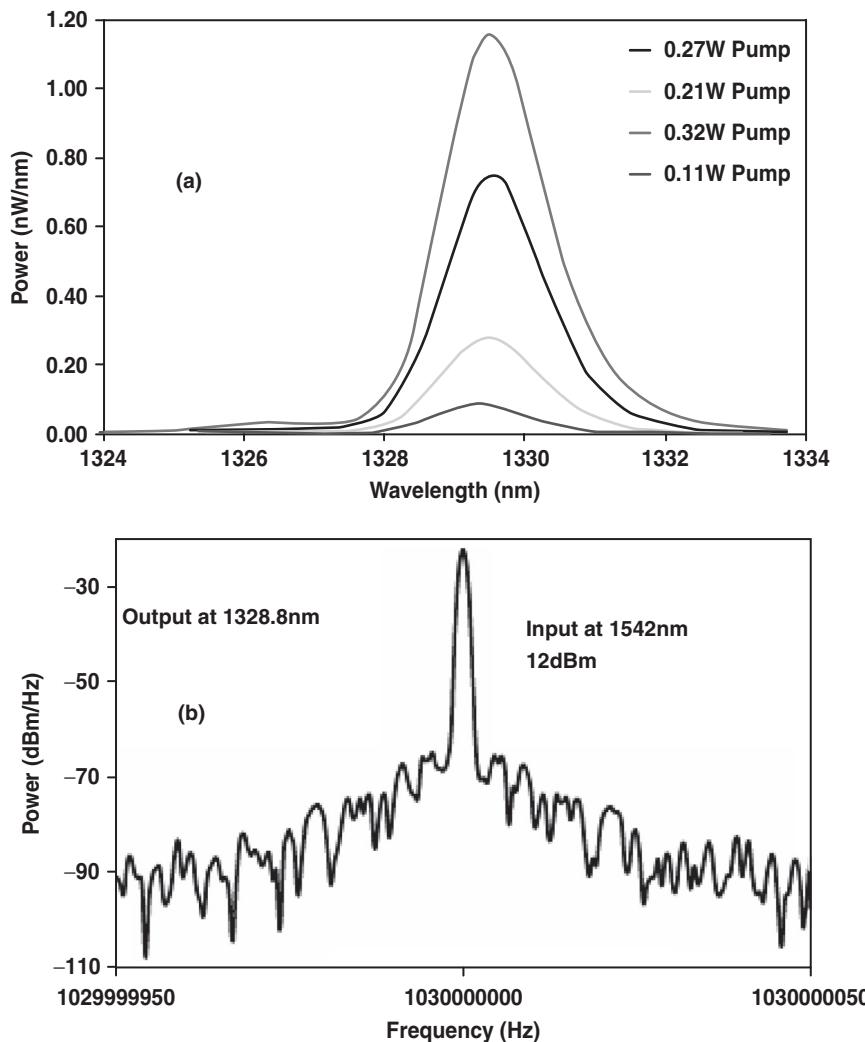
Figure 5.26a shows the converted anti-Stokes signal spectrum [61]. The FWHM for wavelength conversion, which is approximately 250 GHz, is determined solely by the pump laser linewidth. Figure 5.26(b) shows the conversion of 1.03 GHz RF modulation from 1542 to 1328.5 nm [61]. The input RF signal power applied to the Stokes wavelength is shown in the inset. The measured electrical signal-to-noise ratio (SNR) is 47 dBe. We note that from the application point of view, the 1328 and 1550 nm bands are the two most important bands in optical communication. The measured conversion efficiency was approximately  $10^{-5}$ . As mentioned previously, a number of design approaches are available for phase matching a silicon waveguide and for realizing high conversion efficiency [61].

Figure 5.27 shows the measured input–output curve for the first silicon Raman laser [62]. The laser, demonstrated in 2004, operated in the pulse mode and consisted of a 1.7 cm-long silicon waveguide gain medium and an external cavity formed via a fiber loop. The laser has

Stokes/a-Stokes conversion



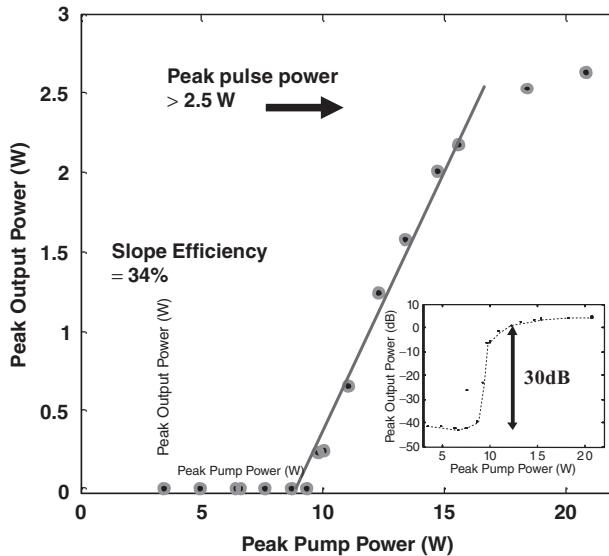
**Figure 5.25** Measured anti-Stokes spectra vs Stokes signal wavelength. The z-axis represents the conversion efficiency, normalized to unity. CW pump wavelength is 1428 nm



**Figure 5.26** (a) Anti-Stokes spectrum of converted signal at varying pump powers; (b) RF spectrum of wavelength converted analog data signal at 1.03 GHz. SNR of 47 dBe is obtained over 100 Hz bandwidth

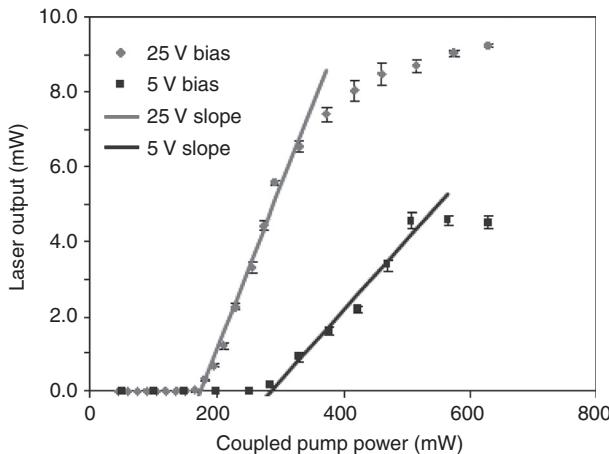
a threshold of 9 W peak pulse power (corresponding to a few mW of average power) and was able to produce output pulses with over 2.5 W peak power. Plotted on a logarithmic scale (inset) the output power abruptly increases by 30 dB when threshold is reached. The demonstration was a major milestone as it clearly showed that silicon can indeed lase. The strong lasing characteristics and high conversion efficiency of  $\sim 34\%$  showed that silicon Raman lasers must be considered as a practical and compact alternative to fiber Raman lasers.

In 2005, Intel Corporation demonstrated the first CW silicon Raman laser [63]. The device was a 5-cm-long silicon waveguide with a cavity formed by HR coating the chip facets. CW



**Figure 5.27** Measured output power with respect to peak pump power for the first silicon laser. Reproduced from [23] by permission of the Optical Society of America. The slope efficiency obtained by dividing the output peak pulse power by that of the input is  $\sim 34\%$

operation was achieved by using a reverse-bias  $p-n$  junction to sweep out the TPA generated free carriers, an approach that was previously proposed in 2004 [64, 65]. Figure 5.28 shows the laser input-output behavior at reverse-bias voltages of 5 and 25 V. The laser produced a maximum output power of  $\sim 9$  mW at 25 V reverse-bias with 600 mW of CW pump power inside the waveguide. The CW operation is an important step in the development of silicon



**Figure 5.28** Threshold characteristics of CW silicon Raman laser demonstrated in a 5 cm silicon waveguide and with using reverse biased  $p-i-n$  diode for carrier sweep-out. Reproduced from [63] by permission of MacMillan Publishers Ltd.

Raman lasers. One drawback of the reverse-bias carrier sweep-out approach is the electrical power dissipated on the chip. With a reverse current of approximately 50 mA expected for this device, the laser dissipates an on-chip electrical power of more than 1W. From this perspective, methods that can drastically reduce carrier lifetime, and hence mitigate the need for active carrier removal, are desirable. As a compromise, reduction of the required sweep-out voltage will reduce the electrical power dissipation.

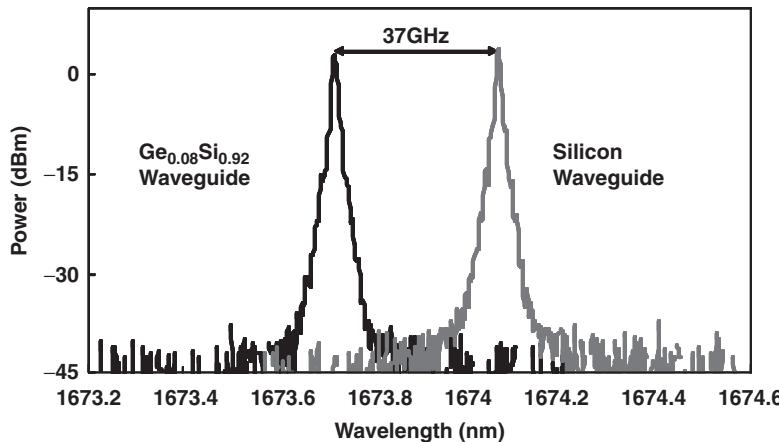
The ability to integrate a *p*-*n* junction along with the gain medium offers the exciting possibility of intra-cavity gain switching [66]. By injecting free carriers into the gain medium, cavity loss can be modulated leading the direct electrical modulation of the laser output. Using this technique, direct laser switching with 30 dB extinction ratio has been demonstrated [66]. This is a unique feature of silicon Raman lasers that is not possible in fiber Raman lasers. It allows the laser to be interfaced with on-chip electronic circuitry in all-silicon optoelectronic integrated circuits. This idea, first demonstrate by UCLA, has been extended by Intel in a demonstration of an electrically switched Raman amplifier, with the device representing a lossless modulator [67].

## 5.7 GeSi Raman Devices

The introduction of germanium in the overall scheme of nonlinear Raman processes in silicon offers new avenues for tailoring the device characteristics. In particular [68]:

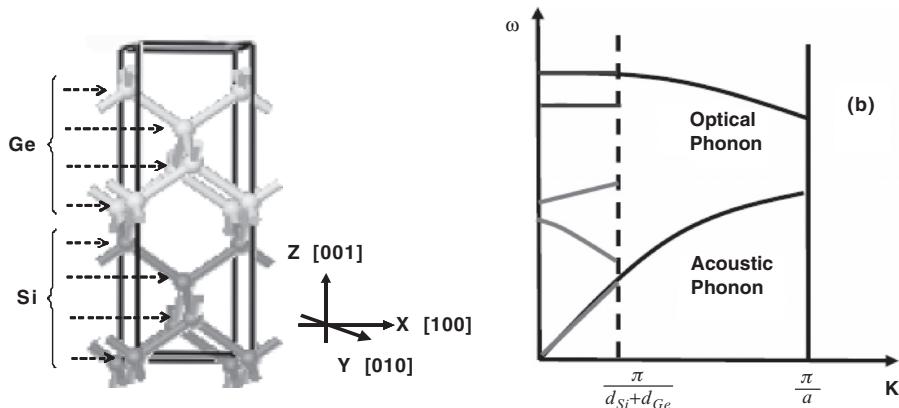
- (i) The strain caused by the difference in the lattice constants of Si and Ge along with the composition effect, provide mechanisms for tuning the Stokes shift associated with the dominant Si–Si ( $500\text{ cm}^{-1}$ ) vibrational mode. In addition, the presence of Si–Ge modes ( $400\text{ cm}^{-1}$ ) and Ge–Ge modes ( $300\text{ cm}^{-1}$ ) provide flexibility in pump and signal wavelengths.
- (ii) Spectral broadening can be achieved by via graded Ge composition.
- (iii) The strain resulting in birefringence can provide an addition degree of freedom for phase matching in the wavelength conversion process. Stress also results in broadening of the gain spectrum, via splitting of the degenerate optical phonon modes.
- (iv) When grown on an SOI substrate, the use of double cladding in the vertical direction can improve fiber waveguide coupling efficiency.
- (v) Higher carrier mobility, and hence diffusion constant, in SiGe reduces the effective lifetime in the waveguide. This reduces the losses associated with the free carriers that are generated by TPA. However, this benefit will be countered by the higher TPA coefficient in GeSi.

Recently, the first GeSi optical amplifier and laser was demonstrated [68]. A pulsed gain of 14 dB and lasing with sharp threshold characteristics were observed for  $\text{Ge}_{0.08}\text{Si}_{0.92}$  rib waveguides. The Stokes spectrum, shown in Figure 5.29, exhibits a 37 GHz red shift which is in qualitative agreement with a model that takes into account the effect of composition and strain on the optical phonon frequency [68]. These results suggest that the spectrum of Raman scattering can be engineered using the GeSi material system. As a result, Ge Si Raman devices represent an exciting topic for future research and development. For example, using Ge/Si superlattices, it is possible to fold the acoustic modes of vibrations into the center of the



**Figure 5.29** Stimulated Raman spectra of GeSi waveguides compared with that of silicon waveguide. A 37 GHz red shift in the Stokes wavelength is observed. Pump wavelength was 1539 nm. Reproduced from [68] by permission of the Optical Society of America

Brillouin zone (Figure 5.30). In this fashion, one can create Raman active vibrational modes with nearly arbitrary vibration frequencies. The pump-Stokes frequency shift (15.6 THz in bulk silicon) is no longer fixed to just one single value. A number of Stokes resonances with the desired shift from the pump frequency can be engineered by proper design of the superlattice structure [69].



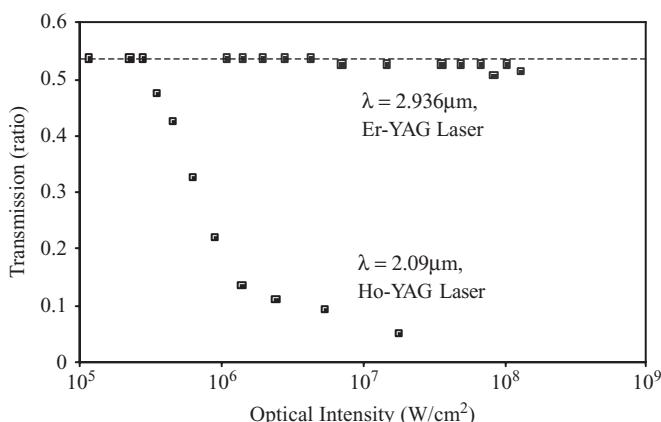
**Figure 5.30** A Si/Ge superlattice (left) and the resulting folding of acoustic modes of vibration into the center of the Brillouin zone. This technique, the pump-Stokes frequency shift is no longer fixed to a single value (15.6 THz in bulk silicon). New Stokes frequencies are created throughout the range of frequencies of the acoustic branch of lattice vibration

## 5.8 Mid-wave IR (MWIR) Silicon Photonics

In this section, we present experimental results of linear nonlinear absorption in silicon at the mid-infrared (IR) wavelengths [70]. Nonlinear losses due to two-photon and free-carrier absorption that are found to degrade near-IR silicon Raman devices become negligible at photon energies less than half the bandgap (i.e.  $\lambda > 2.2 \mu\text{m}$ ). Moreover, the low loss window for linear absorption in silicon extends from 1.2 to  $6.5 \mu\text{m}$ . These factors, along with the excellent thermal conductivity and high optical damage threshold, render silicon an ideal material for building Raman lasers and amplifiers that operate at mid-wave infrared (MWIR) wavelengths. This new technology will expand the application space of silicon photonics beyond data communication and into biochemical sensing, laser medicine, and LIDAR.

The measurement of nonlinear losses in Silicon was performed with the objective of ascertaining the wavelength dependence of free-carrier and two-photon absorption. Pulsed pump laser sources were used in these measurements and were coupled into bulk silicon sample. silicon samples with [1 1 1] orientation, 1 inch diameter, 1 inch length and resistivity of  $2000 \Omega \text{ cm}$  were used in these measurements. A standard  $\text{CaF}_2$  lens with 15 cm focal length was used to focus the laser beam and the silicon sample was moved towards the focus of the lens to increase the intensity of the coupled optical beam. The following solid state pump sources were used in this work: (i) Ho-YAG crystal operating at  $2.09 \mu\text{m}$ , free-running mode with a pulse width of  $100 \mu\text{s}$  and energy of  $\sim 1\text{J}$ ; (ii) Er-YAG laser operating at  $2.936 \mu\text{m}$ , Q-switched with a pulse width of  $100 \text{ ns}$  and energy of  $\sim 25 \text{ mJ}$  [71]. Both these lasers had a spot diameter of  $\sim 2.1 \text{ mm}$ . At the output end, a slow photodetector was used to measure the energy of the pulse.

Figure 5.31 shows the transmission through a silicon sample at pump wavelengths of  $2.09 \mu\text{m}$  and  $2.936 \mu\text{m}$ . The silicon sample was double-side polished and the reflection loss per facet is  $\sim 29\%$ . Hence, the maximum transmission was measured to be  $\sim 53\%$ . At  $2.09 \mu\text{m}$  pump wavelength which corresponds to photon energy of more than half the bandgap, the transmission



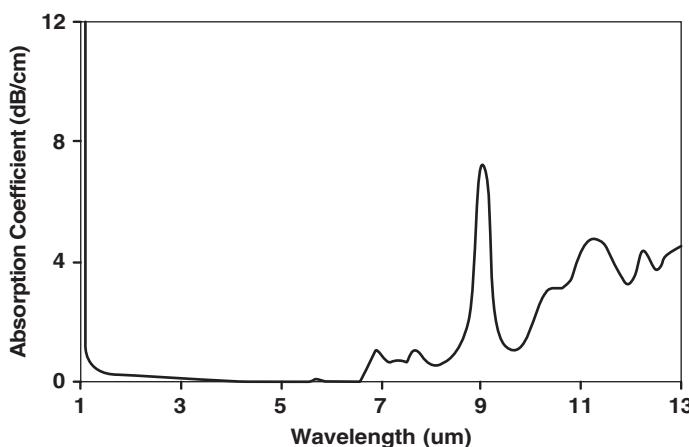
**Figure 5.31** Optical transmission in Silicon as a function of intensity. Two different pump sources at  $2.09 \text{ mm}$  and  $2.936 \text{ mm}$  were used in these experiments. The enhanced nonlinear losses at  $2.09 \text{ mm}$  due to TPA and FCA and the absence of these losses at  $2.936 \text{ mm}$  are clearly seen. Reproduced from [70] by permission of Wiley-VCH

reduces considerably with increasing pump intensity. This loss can be attributed to the two-photon absorption (TPA) process and the losses due to free-carriers generated by TPA. However, as pump photons are reduced in energy below half the bandgap, the two-photon absorption process vanishes. This is observed in the transmission results corresponding to  $2.936\text{ }\mu\text{m}$  pump wavelength. Although the pulse width of the pump sources is different, the peak intensity achieved in the latter case is much higher to compensate the reduced pulse widths. The slight decrease in transmission with increasing intensities could be due to the three-photon absorption (3PA) process. However this process is expected to be extremely weak to cause any significant free-carrier losses. Thus, the absence of nonlinear absorption at pump photon energies less than half the bandgap eliminates a key loss mechanism in silicon Raman devices.

The negligible linear absorption in silicon is another key attribute that makes it attractive for building efficient mid-infrared nonlinear optical devices. The linear absorption in silicon has been extensively studied over the years [72, 73]. Here, we measured the absorption coefficient in our samples using a FTIR apparatus. Figure 5.32 shows the absorption coefficient of silicon in units of dB/cm as a function of wavelength in the range of  $1\text{--}13\text{ }\mu\text{m}$ . The low-loss window following the indirect band gap absorption extends from  $1.2$  to  $\sim 6.5\text{ }\mu\text{m}$  wavelength range. This broad low-loss window clearly underscores the potential of Silicon as the guiding medium for realizing active and passive functionalities in the near infrared (Telecom) as well as the mid-wave infrared wavelengths. Beyond  $7\text{ }\mu\text{m}$  the increase in losses could be due to impurities and multiphonon absorption processes.

### 5.8.1 Prospects of Mid-infrared Silicon Raman Devices

The experimental measurements presented in the previous section clearly suggest that the nonlinear loss mechanisms become insignificant in the mid-infrared wavelength regions, hence eliminating the main problem with near-infrared silicon Raman lasers and amplifiers. This combined with: (i) the unsurpassed quality of commercial silicon crystals; (ii) the low cost and wide availability of the material; (iii) extremely high optical damage threshold of  $1\text{--}4\text{ GW/cm}^2$



**Figure 5.32** Linear absorption in Silicon measured using FTIR [70]

**Table 5.2** Comparison of relevant parameters of silicon and the best Raman crystals in use today

Property	Silicon	Ba(NO <sub>3</sub> ) <sub>2</sub>	LiIO <sub>3</sub>	KGd(WO <sub>4</sub> ) <sub>2</sub>	CaWO <sub>4</sub>
Optical damage threshold (MW/cm <sup>2</sup> )	~1000–4000	~400	~100		
Thermal conductivity (W/m k)	148	1.17		2.6 [1 0 0] 3.8 [0 1 0] 3.4 [0 0 1]	16
Raman gain (cm/GW)	20 (1550 nm)	11 (1064 nm)	4.8 (1064 nm)	3.3 (1064 nm)	
Transmission range (μm)	1.1–6.5	0.38–1.8	0.38–5.5	0.35–5.5	0.2–5.3
Refractive index	3.42	1.556	1.84	1.986–2.033	1.884
Raman shift at 300 K (cm <sup>-1</sup> )	521	1047.3	770 822	901 768	910.7
Spontaneous Raman linewidth (cm <sup>-1</sup> )	3.5	0.4	5.0	5.9	4.8

- 10 × higher optical damage threshold
- 100 × higher thermal conductivity
- High Raman gain, large excellent crystals

(depending on the crystal resistivity); and (iv) excellent thermal conductivity, renders silicon an ideal Raman crystal for MWIR wavelengths. When compared with other popular Raman crystals such as Ba(NO<sub>3</sub>)<sub>2</sub>, LiIO<sub>3</sub>, KGd(WO<sub>4</sub>)<sub>2</sub> and CaWO<sub>4</sub> [74] silicon clearly has the best combination of relevant parameters. Table 5.2 validates this hypothesis by providing numerical values for relevant parameters of silicon and popular Raman crystals.

These factors indicate that silicon Raman lasers should be considered as a tool for covering the technologically important MWIR spectrum. Exploiting the mature silicon fabrication technology, low-loss integrated cascaded microcavities can be employed to realize higher-order Stokes emission. This represents a low-cost and practical approach to extending the wavelength coverage of existing pump lasers to wavelength bands that are outside their reach. This new technology will expand the application space of silicon photonics beyond data communication and into biological, medical and military systems.

A suitable low-index cladding is necessary in order to realize silicon Raman amplifiers and lasers. From work on optical fibers it is known that silica becomes highly lossy beyond 1.8 μm. Thus, SOI is not a suitable for MWIR applications. Another requirement of the cladding layer for high-power lasers and amplifiers is good thermal conductivity. Sapphire is attractive from the point of view of low index (~1.6) and low loss in the MWIR and good thermal conductivity of 23 W m<sup>-1</sup> K<sup>-1</sup> (roughly 20 × higher than silica). Sapphire exhibits low-loss transmission at MWIR wavelengths [75]. Thus silicon-on-sapphire (SOS) structures could be potentially used to build integrated silicon MWIR Raman devices.

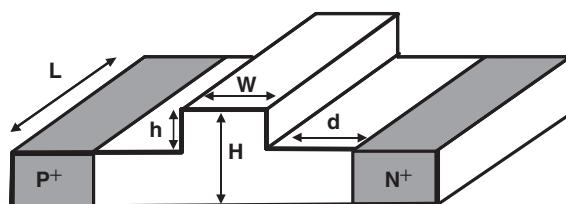
## 5.9 Energy Harvesting in Silicon Raman and Other Nonlinear Optical Devices

The prevailing vision for silicon photonics has been the introduction of photonics into the CMOS VLSI manufacturing process. This vision has been motivated by the compatibility of silicon photonics with the CMOS process. While much progress has been made in Si-based

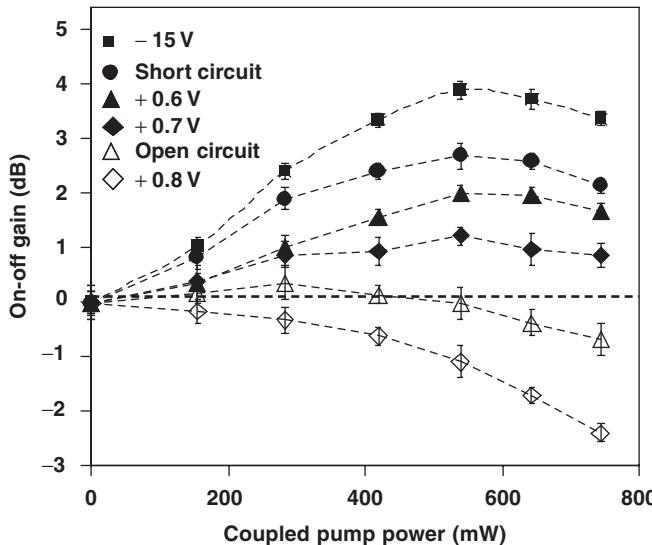
photronics so far, no attention has been paid to the power dissipation of these devices. At the same time, the problem of power dissipation in silicon VLSI is so severe that it threatens to bring to a halt the continued advance of the technology, as described by Moore's law [76]. This fact is highlighted by the recent momentous shift of the microprocessor industry away from increasing the clock speed and in favor of multi-core processors [77]. Evidently, realization of low-power silicon photonic devices is essential for opto-electronic integration.

Typically, lasers are the most power-hungry photonic devices. However, the lack of an electrically-pumped Si laser, to date, dictates an architecture where the light source remains off-chip. In such architecture, an off-chip source empowers the chip, whereas modulators, amplifiers, photodetectors, and perhaps wavelength converters, are integrated on the chip. Among these devices the optical amplifier has the highest power dissipation. As we saw in Sections 5.5 and 5.6, Raman amplification has been the most successful approach for achieving both amplification and lasing in silicon. As discussed therein, to overcome the nonlinear losses, a reverse-biased  $p-n$  junction is used to remove the free carriers. This solution can come at the expense of a large amount of heat being dissipated onto the chip. In reference [63], about 1 W of electrical power had to be dissipated (25 V at  $\sim$ 40 mA) to achieve  $\sim$ 4 dB of CW optical gain and to produce  $\sim$ 8 mW of output from the Raman laser. Recently, it has been demonstrated that similar values of CW gain can be obtained at zero bias, resulting in zero power dissipation [78]. Moving beyond this, it has also been shown that it is even possible to obtain CW optical gain while harvesting electrical power from the device [79, 80]. In this mode, the TPA-generated carriers are swept out by the built-in field of the junction, yet the device delivers electrical power. The concept becomes clear if one considers the device as a nonlinear optical equivalent of a solar cell. To the extent that TPA generation occurs in other semiconductor nonlinear photonic devices (such as wavelength converters), the approach is applicable to those devices as well. This approach works at low to moderate pump intensities, beyond which the junction electric field is unable to sweep out the free carriers, leading to high optical loss, as discussed in Section 5.5.2.

A schematic of the fabricated device is shown in Figure 5.33 [79, 80]. Figure 5.34 presents the measured on-off Raman gain at different coupled pump powers and biasing conditions. As expected, a net optical loss is observed for an open-circuit  $p-n$  junction. When a reverse bias is applied, the induced electric field removes the TPA-induced free carriers from the waveguide region, hence reducing the carrier effective lifetimes. A maximum on-off gain of  $\sim$ 4 dB is obtained at a reverse-bias of 15 V. Considering the measured linear loss of 0.5 dB/cm, the highest achievable net Raman gain is thus  $\sim$ 2.5 dB in the 3-cm-long device. Higher reverse voltages do not increase the gain considerably, which may be attributed to saturation of the



**Figure 5.33** Schematic of the fabricated silicon waveguide with  $p-n$  junction; the values of the geometrical dimensions are:  $H = 2.0 \mu\text{m}$ ,  $h = 0.9 \mu\text{m}$ ,  $d = 1.9 \mu\text{m}$ ,  $W = 1.5 \mu\text{m}$ , and  $L = 3 \text{ cm}$ . Reproduced from [79, 80] by permission of American Institute of Physics



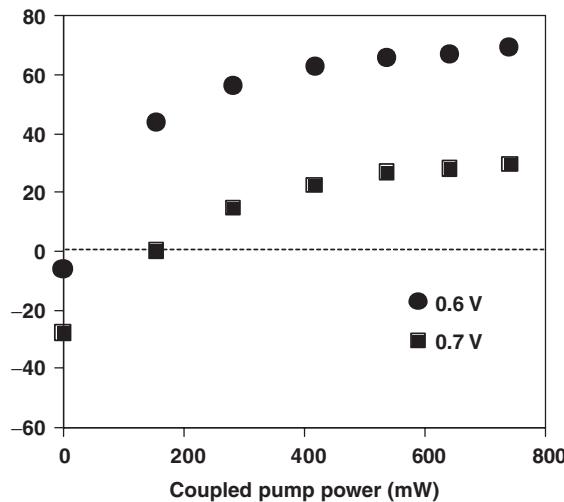
**Figure 5.34** Measured on–off Raman gain vs coupled pump power in the devices of Figure 5.33 at different biasing conditions. Reproduced from [79, 80] by permission of American Institute of Physics

drift velocity [79, 80]. It is also observed that an on–off Raman gain of 2.7 dB is attained when the diode is short-circuited (0 V).

As seen in Figure 5.34, on–off Raman gains as high as 2 dB are measured when the device is forward-biased at voltages  $\leq 0.7$  V. The importance of this biasing regime is that power dissipation is negative, in other words, electrical power is generated by the photovoltaic effect. Figure 5.35 shows the *generated* power for biases of 0.6 and 0.7 V at different pump intensities, extracted from the measured current and voltage drop across the diode. The TPA-induced photovoltaic effect is also clearly evident in the measured *I*–*V* characteristics presented in Figure 5.36. The attenuated pump intensity via TPA creates free carriers in the waveguide. The collected TPA-generated carriers contribute to a current component that delivers electrical power to the external circuitry. Therefore, the sweep-out of TPA generated free carriers can be exploited to generate electrical power.

Sweep-out of carriers and reduction of the effective lifetime has been the original motivation for utilizing a reverse-biased *p*–*n* junction in Raman amplifiers [39, 40, 63]. The forward-bias discussed above might seem counterintuitive. However, it should be reminded that reduction of the effective lifetime can be achieved as long as the diode current is negative. In other words, a negative voltage is not a prerequisite for carrier sweep-out. Numerical drift-diffusion simulations show that the effective lifetime of biases in the fourth quadrant is about one order of magnitude lower than the open-circuit value [78, 79]. Therefore, collection of photogenerated carriers in the fourth quadrant of the *I*–*V* characteristics by the built-in field can sufficiently decreases the free carrier loss and results in CW optical gain.

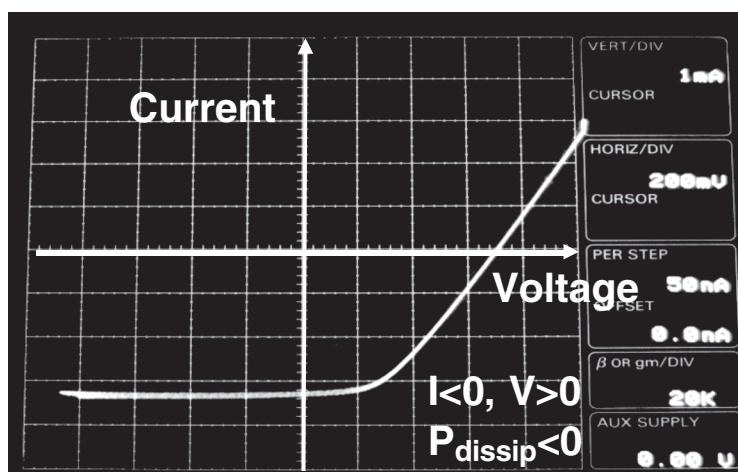
The main limitation of this approach is that at very high optical intensities, the built-in field of the junction is insufficient to remove the high density of TPA-generated carriers. In such cases, a reverse-bias needs to be applied to increase the field, resulting in positive electrical



**Figure 5.35** Generated electrical power at different coupled pump powers for two biases in the fourth quadrant. Reproduced from [79, 80] by permission of American Institute of Physics

power dissipation. Therefore, there exists a trade-off between the amount of gain (and hence output power) and the electrical power generation/dissipation.

Finally, high pump intensities, and hence TPA, are also encountered in other silicon photonic devices that operate based on third-order nonlinear effects such as Raman lasers as well as in Raman- and Kerr-based wavelength converters [60, 81, 82]. Hence, the present approach is also applicable in these devices.



**Figure 5.36** Current–voltage ( $I$ – $V$ ) characteristics of the diode measured with a curve-tracer for an input (uncoupled) optical illumination of  $\sim 1.1$  W. Reproduced from [79, 80] by permission of American Institute of Physics

To summarize this section, TPA and the resulting free-carrier scattering are omnipresent problems in silicon photonic devices that operate based on nonlinear optical interactions. Active removal of these carriers can be achieved, not only without suffering electrical power dissipation and heating, but while generating a modest amount of electrical power.

## 5.10 Summary

Economic incentives as well as intellectual challenges have fueled the race to create silicon lasers. While the indirect bandgap of silicon plays a major role in its inability to amplify light, the prerequisite to lasing, one is well advised to consider the effect of free carrier absorption, the phenomenon that prevents net optical gain from the material in the first place. Silicon-rich oxide (SRO) has been the most widely studied platform in which electroluminescence and optically pumped net gain have been demonstrated. However, some questions remain regarding this technology because the observation of gain is highly dependent on the material preparation conditions, with some samples showing net gain and others showing net loss. In addition, no laser has been demonstrated yet. The addition of erbium to SRO has the advantage of emission in the technologically important 1550 nm band. Electroluminescence and internal gain have been reported, but not lasing. The role of nanocrystals appear to be two-fold: (1) as sensitizers for erbium ions and relaxing the requirement for pump wavelength, and (2) by increasing the conductivity of the oxide, they lower the voltage required to establish current and hence improve the mean time to failure of LEDs (at the cost of reduced efficiency). Erbium-doped SRO tunnel diodes have exhibited efficiencies on par with GaAs LEDs but with low output powers that are limited to tens of microwatts. The reason for this appears to be Auger recombination and free-carrier absorption, phenomena that dampen the prospects for observation of electrically pumped gain and lasing.

Optically pumped lasing has been demonstrated at temperatures below 70 K in two-dimensional etched periodic structures that resemble photonic bandgap geometries. The mechanism for emission is believed to be due to defect states at the etched silicon surfaces.

So far, the most successful approach has been stimulated Raman scattering, an approach that has produced optically pumped amplifiers and lasers. The success of this approach originates from the fact that it does not rely on electronic transitions for producing light. Nevertheless, the pesky free carrier absorption, in the form of scattering from carriers that are generated by two-photon absorption, limits the performance of these devices. However, two-photon absorption vanishes at wavelengths above 2.25  $\mu\text{m}$  and because of this, Raman-based silicon photonics will have its highest impact in the mid-wave IR (MWIR) portion of spectrum. With its high Raman gain coefficient, high thermal conductivity and high optical damage threshold, silicon can compete with the best Raman crystals for MWIR operation. Therefore, it is expected that silicon will find applications in *high-power* amplifiers and lasers that are needed for biochemical detection, LIDAR systems, and fog-resistant free-space communication. Of course, the use of silicon as a passive MWIR material is well established. As a case study, pyroelectric detectors used as motion sensors use silicon windows because of the material's excellent transmission at MWIR wavelengths.

## References

- [1] W. P. Dumke, ‘Interband transitions and maser action’, *Phys. Rev.*, **128**, 1962, 1559.
- [2] W. P. Dumke, ‘Quantum theory of free carrier absorption’, *Phys. Rev.*, **124**, 1961, 1813.

- [3] M. Cardona, *Fundamentals of Semiconductors*, Springer, 2001.
- [4] A. Yariv, *Quantum Electronics*, 3rd edn, Wiley, 1989.
- [5] R. A. Soref and B. R. Bennett, ‘Electrooptical Effects in Silicon’, *IEEE J. Quant. Electron.*, **QE-23**, 123–129, 1987.
- [6] P. M. Fauchet, ‘Light emission from Si quantum dots’, *Materials Today*, **8**, 2005, 26–33.
- [7] L. Pavesi, ‘Routes towards silicon-based lasers’, *Materials Today*, **8**, 2005, 18–25.
- [8] A. Irrera, D. Pacifici, M. Miritello, G. Franzo, F. Priolo, F. Iacona, D. Sanfilippo, G. Di Stefano and P. G. Fallica, ‘Electroluminescence properties of light emitting devices based on silicon nanocrystals’, *Physica E*, **16**, 2003, 395–399.
- [9] R. J. Walters, G. I. Bourianoff and A. Atwater, ‘Field-effect electroluminescence in silicon nanocrystals’, *Nature Materials*, **4**, 2005, 143–146.
- [10] L. Pavesi, L. DalNegro, C. Mazzoleni, G. Franzo and F. Priolo, ‘Optical gain in silicon nanocrystals’, *Nature*, **408**, 2000, 440–444.
- [11] J. Ruan, P. M. Fauchet, L. Dal Negro, M. Cazzanelli and L. Pavesi, ‘Stimulated emission in nanocrystalline silicon superlattices’, *Applied Physics Letters*, **83**, 2003, 5479–5481.
- [12] S. G. Cloutier, P. A. Kossyrev and J. Xu, ‘Optical gain and stimulated emission in periodic nanopatterned crystalline silicon’, *Nature Materials*, **4**, 2005, 887–891.
- [13] M. J. Chen, J. L. Yen, J. Y. Li, J. F. Chang, S. C. Tsai and C. S. Tsai, ‘Stimulated emission in a nanostructured silicon pn junction diode using current injection’, *Applied Physics Letters*, **84**, 2004, 2163–2165.
- [14] M. E. Castagna, S. Coffa, L. Carestia, A. Messian and C. Buongiorno, ‘Quantum Dot Materials and Devices for Light Emission in Silicon’, *Proceedings of 32nd European Solid-State Device Research Conference (ESSDERC 2002)*, Firenze, Italy, 2002, D21.3.
- [15] H. Mertens, A. Polman, I. M. P. Aarts, W. M. M. Kessels and M. C. M. van de Sanden, ‘Absence of the enhanced intra-4f transition cross section at 1.5 μm of Er<sup>3+</sup> in Si-rich SiO<sub>2</sub>’, *Applied Physics Letters*, **86**, 2005, 241109.
- [16] C. D. Presti, A. Irrera, G. Franzò, F. Priolo, F. Iacona, D. Sanfilippo, G. Di Stefano, A. Piana and P.G. Fallica, ‘Light Emitting Devices Based on Silicon Nanoclusters’, *Proceedings of 2nd IEEE International Conference on Group IV Photonics*, Antwerp, Belgium, 2005, 45–47.
- [17] J. Lee, J. H. Shin and N. Park, ‘Optical Gain at 1.5 um in Nanocrystal Si-Sensitized Er-Doped Silica Waveguide Using Top-Pumping 470 nm LEDs’, *IEEE Journal of Lightwave Technology*, **23**, 2005, 19–25.
- [18] A. Polman, B. Min, J. Kalkman, T. J. Kippenberg and K. J. Vahala, ‘Ultra-low threshold erbium-implanted toroidal microlaser on silicon’, *Appl. Phys. Lett.*, **84**, 2004, 1037.
- [19] M. T. Bulsara, *Optical Interconnects Promised by III-V On-Silicon Integration, Solid State Technology*, August 2004, p. 22, 2004.
- [20] Z. Mi, P. Bhattacharya, J. Yang and K.P. Pipe, ‘Room-temperature self-organized In<sub>0.5</sub>Ga<sub>0.5</sub>As quantum dot laser on silicon’, *Electronics Letters*, **41** (13), 2005, 742–744.
- [21] H. Park, A. W. Fang, S. Kodama and J. E. Bowers, *Optics Express*, **13**(23), 2005, 9460–9464.
- [22] A. Huang, G. Li, Y. Liang, S. Mirsaidi, A. Narasimha, T. Pinguet and C. Gunn, ‘A 10Gb/s Photonic Modulator and WDM MUX/DEMUX Integrated with Electronics in 0.13μm SOI CMOS’, *International Solid State Circuits Conference, ISSCC 2006*, 5–9 February, 2006, paper # 13.7.
- [23] O. Boyraz and B. Jalali, ‘Demonstration of a silicon Raman laser’, *Opt. Exp.*, **12**, 2004, 5269–5273.
- [24] R. Claps, D. Dimitropoulos and Y. Han and B. Jalali, ‘Observation of Raman emission in silicon waveguides at 1.54 μm’, *Opt. Exp.*, **10**, 2002, 1305–1313.
- [25] R. Claps, D. Dimitropoulos, V. Raghunathan, Y. Han and B. Jalali, ‘Observation of stimulated Raman scattering in Silicon waveguides’, *Opt. Exp.*, **11**, 2003, 1731–1739.
- [26] R. Claps, V. Raghunathan, D. Dimitropoulos and B. Jalali, ‘Anti-Stokes Raman conversion in silicon waveguides’, *Opt. Exp.*, **11**, 2003, 2862–2872.
- [27] A. Yariv, *Quantum Electronics*, 3rd edn., Wiley, New York, 1988. ISBN 0-471-60997-8.
- [28] Peter Y. Yu and Manuel Cardona, *Fundamentals of Semiconductors*, 3rd edn, Springer.
- [29] P. A. Temple and C. E. Hathaway, ‘Multiphonon Raman Spectrum of Silicon’, *Phys. Rev.*, **A 7**, 1973, 3685.
- [30] D. Dimitropoulos, B. Houshmand, R. Claps and B. Jalali, ‘Coupled-mode theory of Raman effect in silicon-on-insulator waveguides’, *Opt. Lett.*, **28**, 2003, 1954–1956.
- [31] J. M. Ralston and R. K. Chang, ‘Spontaneous-Raman-scattering efficiency and stimulated scattering in silicon’, *Phys. Rev.*, **B 2**, 1970, 1858–1862.
- [32] J. H. Yee and H. H. M. Chau, ‘Two-Photon indirect transition in GaP crystal’, *Opt. Comm.*, **10**, 1974, 56–58.

- [33] K. W. DeLong and G. I. Stegeman, ‘Two-photon absorption as a limitation to all-optical waveguide switching in semiconductors’, *Appl. Phys. Lett.*, **57**(20), 1990, 2063–2064.
- [34] A. Villeneuve, C. C. Yang, G. I. Stegeman, C. N. Ironside, G. Scelsi and R. M. Osgood, ‘Nonlinear Absorption in a GaAs Waveguide Just Above Half the Band Gap’, *IEEE J. Quant. Electron.*, **30**, 1994, 1172–1175.
- [35] A. M. Darwish, E. P. Ippen, H. Q. Lee, J. P. Donnelly and S. H. Groves, ‘Optimization of four-wave mixing conversion efficiency in the presence of nonlinear loss’, *Appl. Phys. Lett.*, **69**, 1996, 737–739.
- [36] Y. -H. Kao, T. J. Xia and M. N. Islam, ‘Limitations on ultrafast optical switching in a semiconductor laser amplifier operating at transparency current’, *J. Appl. Phys.*, **86**, 1999, 4740–4747.
- [37] K. Suto, T. Kimura, T. Saito and J. Nishizawa, ‘Raman amplification in GaP-Al<sub>x</sub>Ga<sub>1-x</sub>P waveguides for light frequency discrimination’, *IEE Proc.-Optoelectron.*, **145**, 1998, 105–108.
- [38] S. Saito, K. Suto, T. Kimura and J. I. Nishizawa, ‘80-ps and 4-ns Pulse-Pumped Gains in a GaP-AlGaP Semiconductor Raman Amplifier’, *IEEE Photon. Technol. Lett.*, **16**, 2004, 395–397.
- [39] T. K. Liang and H. K. Tsang, ‘Role of free carriers from two-photon absorption in Raman amplification in silicon-on-insulator waveguides’, *Appl. Phys. Lett.*, **84** (15), 2004, 2745–2747.
- [40] R. Claps, V. Raghunathan, D. Dimitropoulos and B. Jalali, ‘Influence of nonlinear absorption on Raman amplification in Silicon waveguides’, *Opt. Exp.*, **12**, 2004, 2774–2780.
- [41] A. R. Cowan, G. W. Rieger and J. F. Young, ‘Nonlinear transmission of 1.5 μm pulses through single-mode silicon-on-insulator waveguide structures’, *Opt. Exp.*, **12**, 2004, 1611–1621.
- [42] O. Boyraz, P. Koonath, V. Raghunathan and B. Jalali, ‘All optical switching and continuum generation in silicon waveguides’, *Optics Express*, **12** (17), 2004 4094–4102.
- [43] R. A. Soref and B. R. Bennett, ‘Electrooptical Effects in Silicon’, *IEEE J. Quant. Electron.*, **QE-23**, 1987, 123–129.
- [44] R. J. Bozeat, S. Day, F. Hopper, F. P. Payne, S. W. Roberts and M. Asghari, ‘Silicon Based Waveguides’, in: L. Pavesi, D. J. Lockwood (eds) *Silicon Photonics*, Chap. 8, 269–294, 2004.
- [45] M. A. Mendicino, ‘Comparison of properties of available SOI materials’, in: R Hull (ed.) *Properties of Crystalline Silicon*, **18.1** p 992–1001, 1998.
- [46] J. L. Freeouf and S. T. Liu, *Proc, IEEE Int. SOI Conf.* Tucson, AZ, USA, 3–5 October 1995, p. 74–5.
- [47] D. Dimitropoulos, R. Jhaveri, R. Claps, J. C. S. Woo and B. Jalali, ‘Lifetime of photogenerated carriers in silicon-on-insulator rib waveguides’, *Appl. Phys. Lett.* **86**, 2005, 071115.
- [48] R. Espinola, J. Dadap, R. Osgood, S. J. McNab and Y. A. Vlasov, ‘Raman amplification in ultrasmall silicon-on-insulator wire waveguides’, *Opt. Exp.* **12**(16), 2004, 3713–3718.
- [49] A. Liu, H. Rong, M. Paniccia, O. Cohen and D. Hak, ‘Net optical gain in a low loss silicon-on-insulator waveguide by stimulated Raman scattering,’ *Optics Exp.*, **12**(18), 2004, 4261–4268.
- [50] M. Krause, H. Renner and E. Brinkmeyer, ‘Analysis of Raman lasing characteristics in silicon-on-insulator waveguides’, *Opt. Express*, **12**, 2004, 5703–5710.
- [51] D. Dimitropoulos, S. Fathpour and B. Jalali, ‘Limitations of active carrier removal in Silicon Raman Lasers and amplifiers’, *Opt. Lett.*, **30**, 2005, 261108 .
- [52] X. Chen, N. C. Panoiu and R. M. Osgood Jr., ‘Theory of Raman-Mediated Pulsed Amplification in Silicon-Wire Waveguides’, *Opt. Lett.*, **42**, 2005, 160–170.
- [53] M. Krause, H. Renner and E. Brinkmeyer, ‘Mitigation of Free-carrier absorption and efficiency increase in Silicon waveguide lasers by bidirectional pumping’, *Group IV Photonics*, Antwerp, Belgium, 2005.
- [54] M. Krauss, H. Renner, E. Brinkmeyer, S. Fathpour, V. Raghunathan and B. Jalali, ‘Efficient Raman Amplification in Cladding-Pumped Silicon Waveguides’, *3rd International Conference on Group IV Photonics* 13–15, September 2006, Ottawa, Canada, Paper P6.
- [55] V. Raghunathan, R. Claps, D. Dimitropoulos and B. Jalali, ‘Wavelength conversion in silicon using Raman induced four-wave mixing’, *Applied Physics Letters*, **85**(1), 2004, 34–36.
- [56] V. Raghunathan, R. Claps, D. Dimitropoulos and B. Jalali, ‘Parametric Raman wavelength conversion in scaled silicon waveguides’, *Journal of Lightwave Technology*, **23**(6), 2005, 2094–2102.
- [57] B. Jalali, O. Boyraz, D. Dimitropoulos, V. Raghunathan and R. Claps, ‘Silicon Raman laser, amplifier and wavelength converter’, in: *Optical Interconnects : The Silicon Approach* (Springer Series in Optical Sciences) by L. Pavesi and G. Guillot, eds.
- [58] V. Raghunathan, O. Boyraz and B. Jalali, ‘20 dB on-off Raman amplification in silicon waveguides’, *CLEO 2005*, Baltimore, MD, May 2005, CMU1.
- [59] O. Boyraz and B. Jalali, ‘Demonstration of 11dB fiber-to-fiber gain in a silicon Raman amplifier’, *IEICE Electron. Exp.*, **1**(14), 2004, 429–434.

- [60] R. Claps, V. Raghunathan, D. Dimitropoulos and B. Jalali, ‘Anti-Stokes Raman conversion in silicon waveguides’, *Opt. Exp.*, **11**, 2003, 2862–2872.
- [61] V. Raghunathan, R. Claps, D. Dimitropoulos and B. Jalali, ‘Parametric Raman wavelength conversion in scaled Silicon waveguides’ *IEEE J. Lightwave Tech.*, **23**, 2005, 2094–2012.
- [62] O. Boyraz and B. Jalali, ‘Demonstration of a silicon Raman laser’, *Opt. Exp.*, **12**, 2004, 5269–5273.
- [63] H. Rong, R. Jones, A. Liu, O. Cohen, D. Hak, A. Fang and M. Paniccia, ‘A continuous-wave Raman silicon laser’, *Nature* **433**, 2005, 725–728.
- [64] T. K. Liang and H. K. Tsang, ‘Role of free carriers from two-photon absorption in Raman amplification in silicon-on-insulator waveguides’, *Appl. Phys. Lett.*, **84**(15) 2004, 2745–2747.
- [65] R. Claps, V. Raghunathan, D. Dimitropoulos and B. Jalali, ‘Influence of nonlinear absorption on Raman amplification in Silicon waveguides’, *Opt. Exp.*, **12**, 2004, 2774–2780.
- [66] O. Boyraz and B. Jalali, ‘Demonstration of directly modulated silicon Raman laser’, *Opt. Exp.*, **13**, 2005, 796–800.
- [67] R. Jones, A. Liu, H. Rong, M. Paniccia, O. Cohen and D. Hak, ‘Lossless optical modulation in a silicon waveguide using stimulated Raman scattering’, *Optics Exp.*, Vol. 13(5), 2005, 1716–1723.
- [68] R. Claps, V. Raghunathan, O. Boyraz, P. Koonath, D. Dimitropoulos and B. Jalali ‘Raman amplification and lasing in SiGe waveguides’, *Optics Exp.*, **13**, 2005, 2459–2466.
- [69] R. Claps, D. Dimitropoulos, B. Jalali and B. Jusserand, ‘Raman scattering from acoustic modes in Ge/Si superlattice waveguides’, accepted for publication in *Superlattices and Microstructures*, 2006.
- [70] V. Raghunathan, R. Shori, O. M. Stafsudd and B. Jalali, Nonlinear absorption in silicon and the prospects of mid-infrared silicon Raman lasers, *Phys. Stat. Sol. (a)* **203**(5), 2006, R38–R40.
- [71] R. K. Shori *et al.*, *Proc. SPIE*, **3929**, 2000, 216.
- [72] C. D. Salzberg and J. J. Villa, *J. Opt. Soc. Am.*, **47**, 1957, 244.
- [73] ‘Properties of Silicon’, Ed. T.H. Ning, INSPEC, IEE New York, pp. 70–79, 1988.
- [74] H. M. Pask, *Progress in Quant. Electr.*, **27**, 2003, 3.
- [75] M. E. Thomas, R. I. Joseph and W J. Tropf, *Applied Optics*, **27**(2), 1988, 239–245.
- [76] D. J. Frank, ‘Power-constrained CMOS scaling limits’, *IBM J. Research and Development*, **46**, 2002, 235–244.
- [77] International Technology Roadmap for Semiconductors, 2005 Edition, <http://www.itrs.net/>.
- [78] S. Fathpour, O. Boyraz, D. Dimitropoulos and B. Jalali, ‘Demonstration of CW Raman gain with zero electrical power dissipation in p-i-n silicon waveguides’, *CLEO*, Long Beach, CA, May 2006.
- [79] S. Fathpour, K. K. Tsia and B. Jalali, ‘Energy scavenging in silicon Raman amplifiers’, E-print available online at ArXiv: <http://lanl.arxiv.org/abs/physics/0605143>
- [80] S. Fathpour, K. K. Tsia and B. Jalali, ‘Energy scavenging in silicon Raman amplifier’, submitted to *Appl. Phys. Lett.*
- [81] R. Espinola, J. Dadap, R. Osgood, Jr., S. McNab and Y. Vlasov, ‘C-band wavelength conversion in silicon photonic wire waveguides’, *Opt. Express* **13**, 2005, 4341–49.
- [82] H. Fukuda, K. Yamada, T. Shoji, M. Takahashi, Tai Tsuchizawa, T. Watanabe, J. Takahashi and S. Itabashi, ‘Four-wave mixing in silicon wire waveguides’, *Opt. Express* **13**, 2005, 4629–37.

# 6

# Optical Detection Technologies for Silicon Photonics

A.P. Knights and J.D.B. Bradley

## 6.1 Introduction

In addition to optical emission and modulation, and the design of diffractive elements for functions such as waveguiding and wavelength (de)multiplexing, the development of power monitoring presents one of the key challenges for any optical communication technology. In this chapter we outline the current approaches aimed at the integration of optical detection functionality in silicon photonic circuits. We begin by describing the physical principles of detection using semiconductor junctions and follow with an outline of detector performance criteria. Both of these topics will provide useful background for those unfamiliar with semiconductor detector theory, although for those wishing to explore in greater detail we refer to a number of excellent texts [1–7]. The subsequent section deals with the problem of sensitizing silicon to sub-bandgap photons (such as those with an associated wavelength around 1550 nm) and how one might integrate such detection with optical integrated circuits (OICs). The approaches reviewed include hybrid integration of III–V-based detectors (an approach also used successfully for the integration of infrared emitters in silicon OICs), however greater emphasis is placed on efforts toward a truly monolithic fabrication technology. For example, a considerable amount of recent work has been dedicated to the development of integrated optical detectors, sensitive to wavelengths around 1550 nm, the fabrication of which is completely compatible with standard silicon (CMOS) processing technology. This includes Ge/Si heterostructures, the incorporation of optical dopants (such as erbium) into the silicon matrix, and the controlled introduction of mid-bandgap energy levels via defect engineering. We conclude the chapter with some summary statements on each approach providing the authors' views on their comparative advantages and disadvantages.

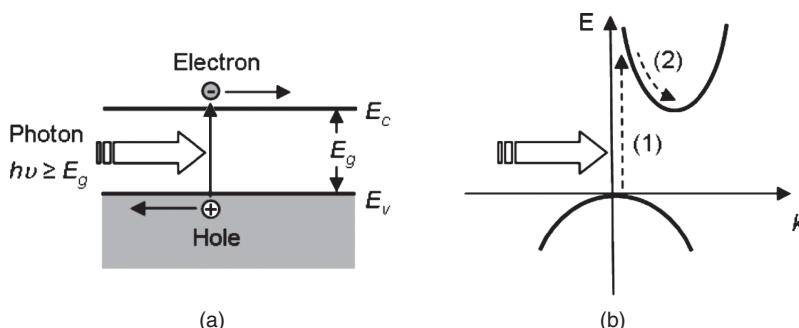
## 6.2 Photodetector Theory

### 6.2.1 Band-edge Photo-excitation of Charge Carriers

It is well established that light has a dual wave–particle nature and may be characterized as individual packets or quanta of energy, called photons. The energy in an individual photon is proportional to its frequency, and can be simply expressed by the following formula:

$$E = h\nu = \frac{hc}{\lambda} \quad (6.1)$$

where  $c$  is the speed of light ( $2.9979 \times 10^8$  m/s in free space),  $\nu$  is the frequency of the light ( $s^{-1}$ ),  $\lambda$  is the wavelength (m) and  $h$  is Planck's proportionality constant ( $6.6262 \times 10^{-34}$  J s). Here, we are primarily interested in the interaction of such photons with a crystalline semiconductor material such as silicon. In a solid crystalline material, the proximity and periodicity of the individual atoms causes the atomic electronic energy levels to overlap and interact, forming a certain electron energy band structure for the entire medium. In a semiconductor, this band structure consists of certain allowed energy levels as well as forbidden energy bandgaps, including that between the highest filled valence level and the conduction band. In order for an electron to move from its stable position in the valence level to the higher energy conduction band where it is free to move it must acquire energy greater than or equal to that of the bandgap energy  $E_g$ . Due to thermal excitations at the given temperature, a certain number of free electrons are always present in the conduction band in the semiconductor's equilibrium state. Likewise, some amount of holes, or empty electrons states left behind by the excited electrons, always exist in the valence band. Once electrons and holes are free in their respective bands, they become mobile and can drift under an applied electric field. Such electron–hole pairs may also be generated by the absorption of a photon of energy  $E > E_g$ , as illustrated in the one-dimensional energy band diagram in Figure 6.1(a). The reverse process may also occur, whereby an electron–hole pair recombines and releases a photon of energy  $E > E_g$ . Photons which have sub-bandgap energy will, in general, simply pass through the medium without being absorbed.



**Figure 6.1** (a) Optical absorption of a photon and generation of an excited mobile electron–hole pair; (b) approximate band structure in  $k$ -space of an indirect bandgap semiconductor showing: (1) direct excitation of an electron to a higher energy state via photon absorption followed by (2) rapid transition to the bottom of the conduction band, releasing energy in the form of a phonon

**Table 6.1** Bandgap type, energy and wavelength (at 300 K) for selected common semiconductor materials

Semiconductor	Group	Type	Bandgap energy $E_g$ (eV)	Bandgap wavelength $\lambda_g$ ( $\mu\text{m}$ )
Silicon	IV	Indirect	1.11	1.15
Germanium	IV	Indirect	0.67	1.88
Aluminum arsenide	III–V	Indirect	2.16	0.57
Gallium arsenide	III–V	Direct	1.43	0.87
Indium arsenide	III–V	Direct	0.36	3.5
Indium phosphide	III–V	Direct	1.35	0.92

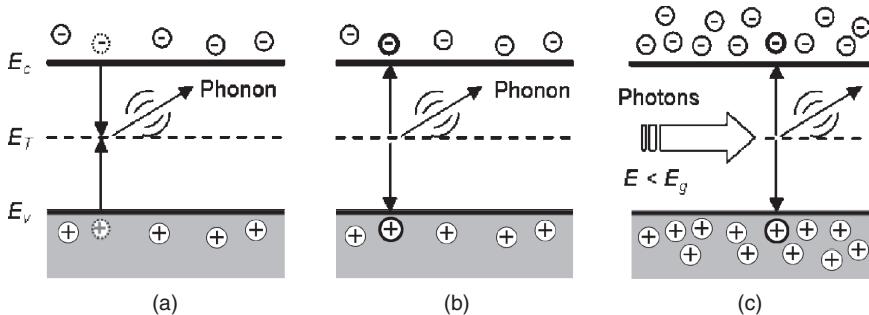
The simple one-dimensional model of photo-excitation is extended in Figure 6.1(b). Here the two dimensional band-structure in  $k$ -space (where  $k$  is the wavenumber and is related to carrier momentum) is shown for a material such as silicon. Silicon is an indirect semiconductor, which means that the conduction band minimum is displaced in  $k$ -space from the valence band maximum. Essentially, this means that the excitation of an electron from valence to conduction band must occur in two stages: direct excitation to a higher energy level at the same position in  $k$ -space, followed by a rapid transition to a lower energy level. This transition, which requires a shift in  $k$ -space, is facilitated by a transfer of momentum to the crystal in the form of a phonon, a quantized kinetic vibration of the crystal lattice structure. Although two separate steps are involved, the absorption process is not unlikely because each step is rapid and happens sequentially. The reverse process, band-to-band emission, is unlikely because it requires the simultaneous presence of three bodies (an electron, a photon and a phonon). This is essentially why silicon can be a highly efficient material for photon detection, while being a poor emitter (light source).

A comparison of silicon with other common semiconductors in terms of bandgap type, energy and corresponding wavelength can be found in Table 6.1. Pure undoped silicon, in particular, has a bandgap energy of about 1.11 eV at 300 K, which means that it is transparent to photons with wavelengths beyond about 1.15  $\mu\text{m}$ . Pure germanium on the other hand, is sensitive to photons with wavelengths as high as 1.88  $\mu\text{m}$ . In order to sensitize crystalline silicon to wavelengths longer than 1.15  $\mu\text{m}$ , the intrinsic structure must be disrupted by introducing dopants or point defects, thus altering the intrinsic periodic band structure. Due to the fact that silicon and germanium are both group IV elements with similar properties, special layers incorporating both materials may also be grown to promote absorption at longer wavelengths.

### 6.2.2 Recombination and Generation via Deep Levels

Two important processes in semiconductors, especially in indirect semiconductors such as silicon and germanium, are carrier generation and recombination via trap levels. Such mid-bandgap (deep levels) are introduced by impurities in the regular atomic crystal structure. Examples of impurities include metals (Au, Cu, etc.) and lattice point defects such as interstitials or vacancies.

So-called Shockley, Read, Hall (SRH) recombination involves the separate capture of an electron and a hole by a mid-bandgap trap level, as well as momentum transfer to the lattice through the impurity or defect, a process which is illustrated in Figure 6.2 (a). The inverse



**Figure 6.2** (a) Recombination and (b) generation of electrons and holes via Shockley, Read, Hall (SRH) defect centers; (c) additional SRH generation via absorption of sub-bandgap energy photons and the resulting excess of mobile electron–hole pairs

process involving carrier generation from a SRH trap site is shown in Figure 6.2 (b). Like the recombination process, SRH generation involves the excitation of first an electron and then a hole (or vice versa), followed by phonon emission. Both processes can be dominant in an indirect material such as silicon because the probability of direct transitions is very low.

Figure 6.2 (c), shows the generation process occurring in the presence of photons, which are absorbed by the trap. In this case, optical absorption at longer wavelengths in the material can be enhanced because less energy is required to excite charge carriers from the mid-bandgap trap level. Furthermore, under illumination the increased rate of generation results in an excess of free carriers in the conduction and valence bands.

In an *n*-type material, the Fermi level lies above the deep level within the bandgap, and the trap is more likely to encounter an electron. Similarly, in a *p*-type material the trap is more likely to encounter a hole. It is more probable then that the recombination (or generation) process will begin with the capture of a majority carrier (i.e. electron in *n*-type semiconductor), followed by minority carrier capture and electron–hole pair recombination (or generation). Therefore, the minority carrier lifetime,  $\tau$ , is the rate-limiting step, and it can be expressed as follows:

$$\tau = \frac{1}{\sigma v_{\text{th}} N_{\text{T}}} \quad (6.2)$$

where  $\sigma$  is the capture cross-section of the defect (of the order of the atomic cross-section or  $1 \times 10^{-16} \text{ cm}^2$ ),  $v_{\text{th}}$  is the thermal velocity of minority carriers ( $\text{cm/s}$ ) and  $N_{\text{T}}$  is the trap density ( $\text{cm}^{-3}$ ). The net carrier recombination or generation rate  $U$  is then given by

$$U = \frac{np - n_i^2}{\tau (p + n + 2n_i \cosh (\frac{E_{\text{T}} - E_{\text{i}}}{kT}))} \quad (6.3)$$

where  $E_{\text{i}}$  is the intrinsic Fermi energy and  $E_{\text{T}}$  is the trap energy,  $p$  and  $n$  are the hole and electron concentrations ( $\text{cm}^{-3}$ ), and  $n_i$  is the intrinsic concentration for both electrons and holes. Evidently,  $U$  is maximized when  $E_{\text{T}} - E_{\text{i}} = 0$ , or when the trap level is close to the center of the bandgap. In equilibrium  $np = n_i^2$  and there is no net recombination or generation. Under nonequilibrium conditions, the magnitude of  $np - n_i^2$  in the numerator determines whether net recombination ( $U > 0$ ) or generation ( $U < 0$ ) is taking place.

In a silicon photodiode, carrier generation via deep defect levels can be utilized in order to increase sensitivity at sub-bandgap wavelengths. Under illumination, the photodiode is no longer in equilibrium and there will be a surplus of excited electron–hole pairs, as shown in Figure 6.2 (c). If carriers are generated within or close to the depletion region of the photodiode, most will be rapidly swept out by the large electric field and collected as photocurrent before they can recombine at trap sites. Furthermore, by introducing point defects (such as di-vacancies) which are known to result in energy levels close to the center of the bandgap, the carrier generation rate can be maximized.

### 6.2.3 p–n Junction Theory of Semiconductors

In order to control the flow of excited charge carriers and produce a measurable photocurrent, a suitable device structure is required. In semiconductors, the basic *p*–*n* junction provides a simple and effective means of achieving this result. A *p*–*n* junction functions as a rectifying structure, which has many applications in electronics and optoelectronics. A brief review of the relevant *p*–*n* junction theory is presented here.

The properties of silicon are significantly altered when specific dopant atoms, such as boron or phosphorous, are introduced into the lattice. In intrinsic (undoped) silicon, the Fermi energy  $E_F$ , which is the energy level at which the probability of occupation by an electron is 0.5, lies near the center of the bandgap  $E_i$ . Below the Fermi level the probability of occupancy rapidly increases towards 1, while above, the probability drops quickly to 0. In an *n*-type material, where *donor* impurities such as phosphorus contribute mobile electrons, the Fermi energy lies closer to the conduction band. Thus, the probability of electrons existing in the conduction band is much higher and there are more mobile electrons than holes. The opposite is true for a *p*-type material, whereby the Fermi level is shifted closer to the valence band and there is a resultant surplus of mobile holes within this band.

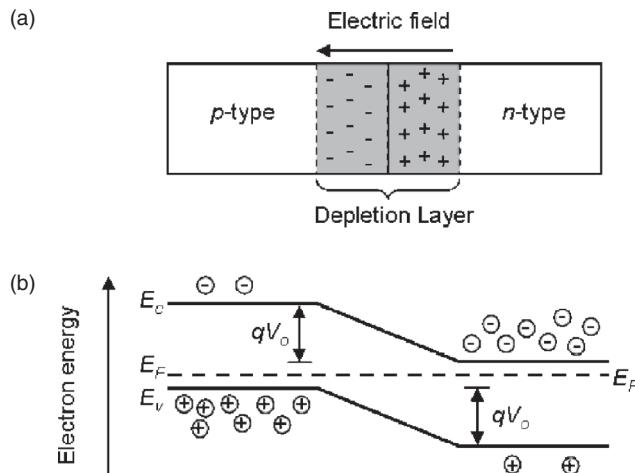
In general,  $E_F$  is related to each of the carrier concentrations as follows:

$$n = n_i e^{\left(\frac{E_F - E_i}{kT}\right)} \quad (6.4)$$

$$p = n_i e^{\left(\frac{E_i - E_F}{kT}\right)} \quad (6.5)$$

where  $n$  is the density of electrons in the conduction band,  $p$  is the density of holes in the valence band and  $n_i$  represents both the number of electrons and the number of holes in their respective bands in the intrinsic material (for silicon,  $n_i = 1.45 \times 10^{10} \text{ cm}^{-3}$  at 300 K). The number of mobile holes and electrons are also related to the intrinsic carrier concentration by the mass action relation  $np = n_i^2$ . In *n*-type material it is usually reasonable to approximate the number of mobile electrons as equal to the number of donors  $N_D$ . Likewise, in *p*-type material, it can be estimated that the number of mobile holes is equal to the number of acceptors  $N_A$ .

When a *p*- and an *n*-region are formed adjacent to each other, the resulting structure is called a *p*–*n* junction. In contact, excess holes in the *p*-region (majority carriers) diffuse towards the *n*-region and excess electrons from the *n*-region diffuse towards the *p*-region. When the electrons diffuse into the *p*-region they recombine with the abundant holes and leave behind positively ionized donor impurity atoms within the crystal structure on the *n*-side of the junction. Similarly, holes from the *p*-region recombine with electrons on the *n*-side and leave behind



**Figure 6.3** A  $p$ - $n$  junction in equilibrium, showing: (a) the fixed charge due to ionized donors and acceptors and the resulting electric field across the despletion layer; (b) the electron energy band diagram, including the built-in junction potential and a representation of the relative number of mobile electrons and holes across the junction

negatively charged acceptor atoms. The result is an area between the  $p$ - and  $n$ -regions where majority carriers are depleted and there is an electric field present due to the positive and negative acceptor and donor impurities. This built-in electric field opposes the diffusion of majority carriers and eventually a state of equilibrium is established. The area which is depleted of mobile majority carriers in the middle of the junction is referred to as the depletion layer, and is illustrated in Figure 6.3 (a). In equilibrium there is a net built-in potential  $V_0$  across the depletion layer, which causes the energy bands to bend as illustrated in Figure 6.3 (b). The Fermi energy is constant across the junction in equilibrium.

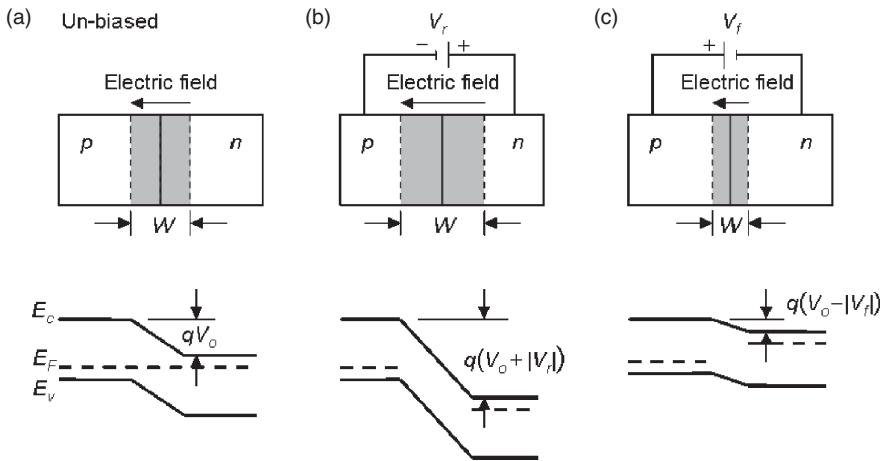
If both doping concentrations are known, the built-in potential of an abrupt  $p$ - $n$  junction can be calculated as follows:

$$V_0 = \frac{kT}{q} \ln \frac{N_A N_D}{n_i^2} \quad (6.6)$$

where  $N_A$  is the concentration of acceptors on the  $p$  side ( $\text{cm}^{-3}$ ),  $N_D$  is the concentration of donors on the  $n$ -side. The width of the depletion region  $W$  is given as follows:

$$W = \left[ \frac{2\epsilon V_0}{q} \left( \frac{1}{N_A} + \frac{1}{N_D} \right) \right]^{1/2} \quad (6.7)$$

where  $\epsilon$  is the permittivity of the medium ( $1.04 \times 10^{-12} \text{ F/cm}$  in silicon at 300 K). It is evident from Equation (6.7) that, if there is any significant difference in doping levels, the depletion layer width will be determined primarily by the doping concentration in the lower-doped region (the depletion layer will also extend mostly into this region).

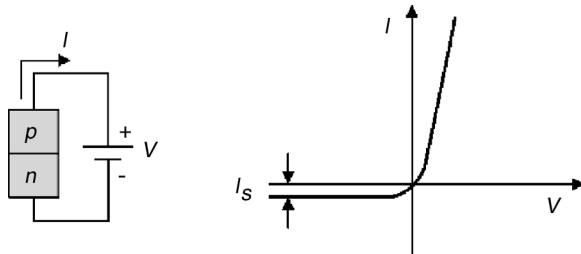


**Figure 6.4** The electric field, depletion width and energy band diagram of a  $p$ - $n$  junction under: (a) no applied bias ( $V = 0$ ) (b) reverse bias ( $V = V_r$ ); (c) forward bias ( $V = V_f$ )

The current through a  $p$ - $n$  junction can be divided into two components: *drift* and *diffusion* current. Drift current is the movement of charge carriers as a result of an electric field, such as the built-in field across the depletion layer. Diffusion current is caused by the difference in concentration of charge carriers between two regions. A summary of the diode characteristics in equilibrium, and under forward- and reverse-bias conditions is shown in Figure 6.4. Due to the statistical nature of the electron energies (described by the density of states and the Fermi-Dirac distribution function), some electrons in the conduction band on the  $n$ -side will have sufficient energy to diffuse across the potential barrier in the opposite direction of the drift current. In equilibrium, or at zero applied bias, drift and diffusion current cancel each other and there is no net current across the junction.

Under an applied bias the diode is no longer in equilibrium. In forward bias, with a positive voltage applied at the  $p$ -contact, an electric field results that opposes the built-in junction potential and the Fermi level is no longer constant. The result is a reduction in the potential barrier by the applied voltage  $V$  and a corresponding increase in the carrier diffusion current (more electrons now have sufficient energy to surmount the potential barrier). The current due to diffusion increases significantly as more minority carriers are injected across the barrier into both the  $p$  and  $n$  regions (for example electrons from the  $n$ -side into the  $p$ -side) where they recombine with the abundant majority carriers. The diffusion current quickly begins to dominate because the small negative drift current is relatively independent of the applied bias. Therefore, small increases in voltage result in large increases in current. Furthermore, the width of the depletion layer decreases because the applied voltage is positive and  $V_o$  in Equation (6.7) is now replaced by the term  $(V_o - V)$ .

In reverse bias, where a negative potential is now applied at the  $p$ -contact, the height of the electrostatic potential barrier is increased, which decreases minority carrier diffusion across it. The result is that the smaller drift current now dominates and there is a small and relatively constant reverse bias saturation current  $I_S$ . The reverse saturation current is relatively constant because the limiting factor for drift current is the generation time for minority carriers within



**Figure 6.5** An ideal p–n junction and its characteristic  $I$ – $V$  response

a diffusion length of the depletion layer, not the strength of the applied field. Once within the depletion layer they are rapidly swept through and the drift current depends solely on the small amount of minority carriers that become available. The applied voltage is also now negative so the depletion width is increased according to the nonequilibrium form of Equation (6.7).

In both forward and reverse bias the probability of diffusion across the junction is found to depend on the factor  $\exp(qV = kT)$  and in equilibrium the diffusion current is equal to the drift (or saturation) current  $I_S$ . The resulting current–voltage ( $I$ – $V$ ) equation for an ideal diode is then simply the diffusion current minus the drift current:

$$I = I_S \left( e^{\frac{qV}{kT}} - 1 \right) \quad (6.8)$$

The corresponding  $I$ – $V$  characteristic curve is illustrated in Figure 6.5. A silicon diode exhibits a sharp rise in current at a forward-bias voltage typically in the range 0.6–0.7 V. If the effects of majority carrier concentration changes on either side of the junction due to large forward-bias diffusion currents are considered, the applied voltage is generally limited to approximately  $V_0$ .

The drift current  $I_S$ , depends on carriers which are generated within a diffusion length of the depletion region. Therefore, it depends on the carrier concentrations and their associated diffusion characteristics as follows:

$$I_S = qA \left( \frac{D_p}{L_p} p_n + \frac{D_n}{L_n} n_p \right) \quad (6.9)$$

where  $p_n$  is the equilibrium concentration of holes (minority carriers) in the  $n$ -region ( $\text{cm}^{-3}$ ),  $n_p$  is the equilibrium concentration of electrons in the  $p$ -region,  $D_p$  and  $D_n$  are the diffusion coefficients of holes and electrons ( $\text{cm}^2\text{s}^{-1}$ ) and  $L_p$  and  $L_n$  are diffusion lengths for holes and electrons (cm). The diffusion length is defined as the average length a carrier diffuses before recombining and can be expressed by  $L = (D\tau)^{1/2}$ , where  $\tau$  is the recombination lifetime (s). The diffusion coefficient  $D$  is also directly related to the carrier mobility,  $\mu$  ( $\text{cm}^2/\text{V s}$ ) at a given temperature:

$$\frac{D}{\mu} = \frac{kT}{q} \quad (6.10)$$

where  $k$  is Boltzmann's constant ( $\text{J/K}$ ),  $T$  is the temperature (K),  $q$  is the magnitude of charge on a single electron ( $1.6022 \times 10^{-19} \text{ C}$ ).

A simplified model of diode operation has been presented. However, there are additional considerations which will cause the basic  $I$ - $V$  curve to deviate from the ideal model in Equation (6.8). The rest of this section briefly discusses some of these considerations, especially those most relevant to photodiode operation.

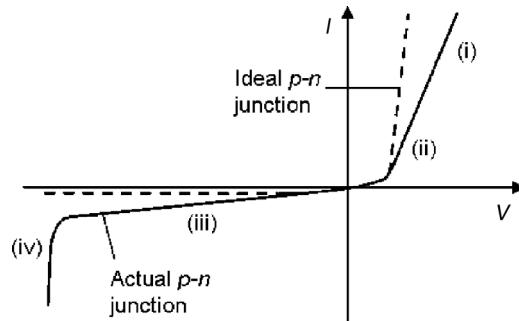
The diode model implies that the entire voltage appears across the junction and ohmic contacts exist at the  $p$ - and  $n$ -regions. However, there will always be some resistance (although small) in the  $p$ - and  $n$ -regions and therefore a small voltage drop across them. As injection increases in forward bias, a larger voltage drop appears across these regions, which must be subtracted from the applied bias to determine the voltage difference across the junction. This means that the current through the device at higher voltages will not increase as sharply as for the ideal case. However, the resistivity of the doped regions and contacts are generally very low as long as carrier injection remains low (such as in reverse bias where photodiodes are operated), so resistivity effects usually can be ignored.

It has also been assumed that the doping concentration variance across the junction is constant and there exists an abrupt transition from  $p$  to  $n$  at the interface. Often this is not the case, although if the more highly doped region is short (or shallow in the vertical case), it is sufficient to approximate an abrupt junction.

Most importantly perhaps, the ideal diode model ignores carrier generation in the depletion region. The ideal model implies that only carriers generated outside this region which are within a diffusion length are collected and swept through to contribute to the drift current. However, all carriers generated within the depletion layer can be assumed to be swept out, and will contribute to the drift current. This contribution becomes especially significant as the depletion width increases, since there is a greater volume available for generation. Therefore, because the depletion width increases with the square-root of applied voltage, the reverse bias current is also proportional to  $\sqrt{V}$ .

The current due to recombination in the depletion region is found to be proportional to  $n_i$  and increases according to  $\exp(qV/2kT)$ , while outside it is proportional to  $p_n$  and  $n_p$  and increases according to  $\exp(qV/kT)$ . The result is that the overall diode equation must include a factor  $x$  in the denominator of the exponential term (where  $x$  approaches 2). Recombination in the depletion region begins to dominate at low temperatures, at low applied voltages and in wide-bandgap materials (low  $n_i$ ). In addition, if recombination centers or traps are deliberately introduced into the material (as discussed in Section 6.2.2), the forward-bias current will be relatively decreased. This is because the mobile charge carriers have a much shorter recombination lifetime and fewer carriers are able to diffuse across the junction without recombining. Recombination via traps also increases the overall resistivity of the material.

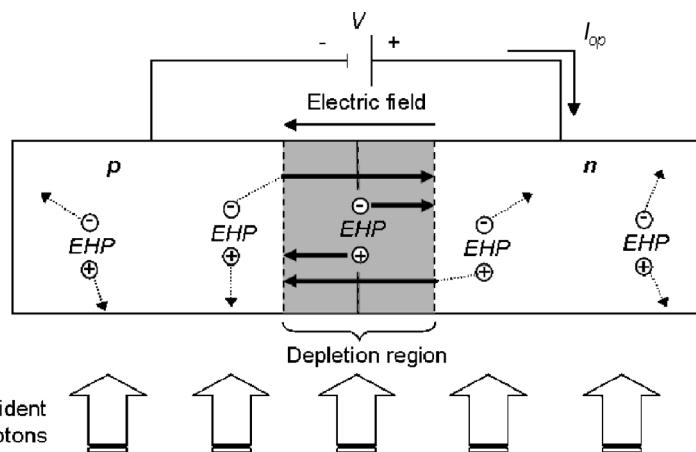
It has been shown that under reverse bias a small and relatively constant current exists. This description is valid only up to a critical voltage level, at which reverse-bias breakdown occurs. At this point, the current rises sharply with only a small increase in voltage, which is similar to the response in forward bias. Two breakdown processes exist: Zener breakdown and avalanche breakdown. Zener breakdown occurs as a result of the energy bands crossing on either side of the junction. If the depletion width is small, as is the case in heavily doped junctions, electron tunneling can occur. Avalanche breakdown is due to impact ionization resulting from highly energetic carriers being swept through the transition region. This effect is in fact used in a specific type of detector called an avalanche photodiode to generate large photocurrents. The deviations from simple diode theory discussed here and their effect on the  $I$ - $V$  response are summarized in Figure 6.6.



**Figure 6.6**  $I$ - $V$  response of an ideal and a nonideal  $p$ - $n$  junction, illustrating the effects of: (i) non-ohmic contacts and material resistivity; (ii) carrier recombination and (iii) carrier generation within the depletion region; (iv) reverse bias breakdown

#### 6.2.4 Photodiodes

In the presence of light of energy greater than the bandgap, additional electron–hole pairs are generated within a  $p$ - $n$  junction that change the  $I$ - $V$  response of the biased device. Figure 6.7 illustrates the charge carrier mechanisms in an illuminated reverse-biased  $p$ - $n$  junction. Extra photogenerated electron hole pairs in the depletion region are swept out by the electric field, after which the electrons recombine with holes on the  $n$ -side and holes recombine with electrons on the  $p$ -side. The additional hole necessary for recombination with a photogenerated electron on the  $n$ -side must be pulled across the terminals from the  $p$ -side. Similarly, electrons from the  $n$ -side are required for recombination with holes in the  $p$ -doped region. If the device



**Figure 6.7** Illumination of a reverse-biased  $p$ - $n$  junction and the resultant drift and diffusion of photo-excited electron–hole pairs. Carriers close to or within the depletion region are swept out by the faster drift process (indicated by thick arrows) and contribute to the photocurrent. Carriers generated far from the depletion region diffuse for some distance and then recombine without contributing to the photocurrent

is reverse biased, both processes enhance the magnitude of the reverse current. Electron hole pairs generated close to the depletion layer (i.e. within a diffusion length) are also likely to diffuse into this region where they undergo drift and contribute to the increased reverse current, or *photocurrent*. Electron hole pairs generated farther than a diffusion length from the depletion layer will diffuse randomly and recombine, without contributing to the photocurrent.

The  $I$ - $V$  relation of a diode can be modified to include the photocurrent  $I_{\text{op}}$  as follows:

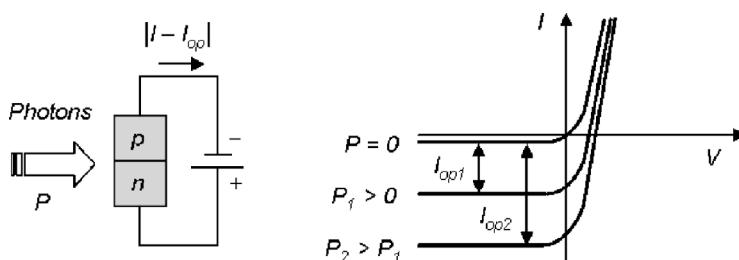
$$I = I_0 \left( e^{\frac{qV}{kT}} - 1 \right) + I_{\text{op}} \quad (6.11)$$

The magnitude of  $I_{\text{op}}$  is proportional to the incident optical power (below saturation) and depends on the generation rates and the separate contributions due to drift and diffusion of electron–hole pairs. Considering a uniformly illuminated photodiode, the electron–hole pair generation rate or  $G_{\text{op}}$  ( $\text{EHP}/\text{cm}^3 \text{ s}$ ) will be constant throughout the device. The number of holes generated per second within a diffusion length of the depletion region on the  $n$ -side is then  $AL_p G_{\text{op}}$  (where  $A$  is the area of the junction), while the generation rate of electrons on the  $p$ -side is  $AL_n G_{\text{op}}$ . Additionally,  $AWG_{\text{op}}$  carriers are generated within the depletion region. Assuming the drift time is much shorter than the recombination lifetime, all of these generated EHPs contribute to the photocurrent. Therefore,  $I_{\text{op}}$  can be expressed by:

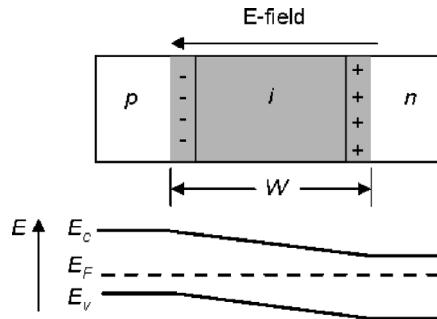
$$I_{\text{op}} = qAG_{\text{op}}(L_p + L_n + W) \quad (6.12)$$

The resulting change to the ideal  $I$ - $V$  diode response is illustrated in Figure 6.8. In open-circuit operation, or photovoltaic mode, the movement of mobile charge results in an increased electric field and generates a photovoltage across the device. In closed-circuit operation, the diode generates a photocurrent proportional to the incident optical power  $W$ . Photodiodes are usually operated under reverse bias because this increases the electric field in the depletion layer, which serves to increase the drift velocity of the carriers. The applied reverse bias also increases the width of the depletion layer, which means that more generated carriers are collected by drift, resulting in a larger signal. Furthermore, a greater proportion of the photocurrent is then due to the faster drift process (as opposed to diffusion), which lowers the transient response time of the photodiode.

A commonly used variation of the  $p$ - $n$  junction is the  $p$ - $i$ - $n$  junction diode, which is shown in Figure 6.9. It consists of an intrinsic (undoped) or lightly doped region between a  $p$ -type and  $n$ -type region. A  $p$ - $i$ - $n$  diode with lightly doped  $n$ -type (or  $p$ -type) as opposed to intrinsic



**Figure 6.8** An ideal photodiode and its corresponding  $I$ - $V$  relationship. The unbiased or reverse-biased diode experiences an increase in current  $I_{\text{op}}$  proportional to the incident optical power,  $P$



**Figure 6.9** A  $p$ - $i$ - $n$  diode and its corresponding electron energy band diagram

material is also referred to as a  $p$ - $\nu$ - $n$  diode (or  $p$ - $\pi$ - $n$  diode). In this structure two depletion regions form, one at the  $p$ - $i$  interface and the other at the  $i$ - $n$  interface. Both depletion layers extend mostly into the lightly doped  $i$ -region and can be made to overlap so that a single, large depletion layer and corresponding electric field extends entirely across the intrinsic region of the device. The size of the depletion layer can be effectively controlled in this manner. Such a structure is commonly used for semiconductor photodiodes instead of a regular  $p$ - $n$  junction because the larger depletion width means more photons can be captured, resulting in enhanced device efficiency. The larger depletion layer can also improve the response time of the device.

## 6.3 Photodetector Performance Characteristics

### 6.3.1 Quantum Efficiency

The quantum efficiency is defined as the probability that a single incident photon will generate an electron–hole pair that contributes to the detector current, or more simply, the number of electron–hole pairs generated per incident photon. Thus, the quantum efficiency  $\eta$  can be represented by the following equation:

$$\eta = \frac{I_{\text{op}}/q}{P/h\nu} \quad 0 < \eta < 1 \quad (6.13)$$

where  $P$  is the optical power (W) incident on the detector.

Another useful method for expressing  $\eta$  is as follows:

$$\eta = T_{\text{op}} F_{\text{EHP}} (1 - e^{-\alpha L}) \quad (6.14)$$

where  $T_{\text{op}}$  is the percentage of optical power transmitted at the detector interface,  $F$  is the fraction of electron hole pairs generated which contribute to the photocurrent,  $\alpha$  is the absorption coefficient ( $\text{cm}^{-1}$ ) and  $L$  is the detector length. The factor  $T_{\text{op}}$  represents loss due to reflection and in-coupling at the detector interface. For a detector integrated with an optical waveguide for example,  $T_{\text{op}}$  could represent reflection loss due to input coupling from a fiber to the waveguide. Equation (6.14) defines the external quantum efficiency of the detector, while the internal quantum efficiency is given as  $\eta/T_{\text{op}}$ .

Some generated electron–hole pairs will not contribute to the photodetector current due to recombination, often at the surface where recombination centers are abundant or in other regions with high defect concentrations. These, and electron–hole pairs generated outside the carrier collection region of a photodiode for example (see Figure 6.7), contribute to reducing the factor  $F_{\text{EHP}}$ . Furthermore, the photo-excitation process has a certain probability associated with it, which translates into the absorption coefficient  $\alpha$  (specific to the medium) in the last factor of Equation (6.14). The overall absorption increases with detector length. Quantum efficiency is also a function of the wavelength because the absorption coefficient is wavelength dependant. In intrinsic material the quantum efficiency is generally negligible at wavelengths  $\lambda_0 > \lambda_g = hc_0/E_g$ .

### 6.3.2 Responsivity

The responsivity,  $\mathfrak{R}$  (A/W), of a detector is the proportionality factor which relates the incident optical power to the amount of generated photocurrent. The incident optical power is given by  $P = h\nu N$  (W) where  $N$  is the photon flux ( $\text{s}^{-1}$ ) and the photocurrent is given as  $I_{\text{op}} = \eta q N$  (A). Taking the ratio of these expressions, the responsivity can be defined simply by the following equation:

$$\mathfrak{R} = \frac{\eta q}{h\nu} \quad (6.15)$$

In general, the photocurrent responds linearly with incident optical power, until a saturation level is reached. Like the quantum efficiency, the responsivity is wavelength dependent. Because silicon is an indirect bandgap material, the responsivity of a silicon photodetector will be significantly higher at wavelengths where the photon energy makes a direct transition possible.

### 6.3.3 Response Time

The impulse response function of a photodetector is limited by both the transit time of charge carriers and the  $RC$  time constant associated with the detector and its circuitry. In photodiodes, three factors determine the fundamental response time limit of the device: (1) the diffusion time of carriers generated outside the depletion region; (2) the drift time of carriers across the depletion region; and (3) the junction capacitance.

The diffusion process is relatively slow, and is limited by the recombination lifetime  $\tau$ . It is desirable to minimize the relative contribution of carrier diffusion to the overall photocurrent by ensuring that most electron–hole pairs are generated within the depletion region itself.

Drift of carriers through the depletion region is much more rapid due to the influence of the electric field. At low electric fields, the drift velocity is given by:

$$v_d = \mu E \quad (6.16)$$

where  $E$  is the electric field (V/cm). The drift velocity can be increased under reverse bias until a saturation drift velocity  $v_s$  is reached. The saturation velocity for silicon and germanium is found to be of the order of  $\sim 10^7$  cm/s [5]. If the depletion region is 1  $\mu\text{m}$  wide, this translates to a minimum drift time of picoseconds.

The junction capacitance  $C_j$  of an abrupt  $p-n$  junction, which arises due to the ionized donors and acceptors in the depletion layer, can be expressed as follows:

$$C_j = \varepsilon_A \left[ \frac{q}{2\varepsilon(V_o - V)} \frac{N_D N_A}{N_D + N_A} \right]^{1/2} = \frac{\varepsilon A}{W} \quad (6.17)$$

where  $A$  is the cross-sectional area of the diode. In analogy to the parallel-plate capacitor, where  $C = \varepsilon A/d$  (where  $d$  is the plate separation), the junction capacitance decreases with increasing width of the depletion layer. Therefore, it will decrease as the magnitude of the reverse-bias voltage is increased. The series resistance of a diode is inherently low, thus the junction capacitance, the small depletion layer resistance, and any load resistance applied to the device essentially determine the overall  $RC$  time constant of the detector.

To minimize the response time of a photodiode, the equilibrium depletion layer should be wide to increase the proportion of carriers collected by the rapid drift process and reduce the junction capacitance. Operating the diode in reverse bias further increases the depletion width and improves the drift speed, thus improving the response time. However, if the depletion width is too high, transit time across the region begins to limit the overall response speed of the device.

### 6.3.4 Detector Noise

There are two significant contributions to the noise current in a  $p-n$  junction photodiode: shot noise and thermal noise. The shot noise, or noise due to random fluctuations in the flow of charge carriers, is affected by the various contributions to the overall current, and can be defined as follows:

$$I_{s,RMS}^2 = 2qB(I_{op} + I_D + I_B) \quad (6.18)$$

where  $I_{s,RMS}$  is the root-mean-square (RMS) shot noise,  $B$  is the bandwidth,  $I_D$  is the dark current of the photodiode and  $I_B$  is the photocurrent due to background radiation. Under carefully controlled conditions,  $I_B$  should be small and can be neglected.  $I_D(V)$  is the reverse-bias current of the nonilluminated diode at a given voltage, equal to the constant value  $I_S$  in the ideal case. If the depletion layer is large (as in the case of a  $p-i-n$  photodiode),  $I_D$  is almost entirely due to thermal generation within this layer.

The thermal noise, or Johnson noise, is due to the equivalent resistance of the diode itself and any additional electronics (resistors, amplifiers) associated with the overall detection circuit. The diode has a small series resistance, which can normally be neglected, and a depletion layer resistance  $R_j$ . In the equivalent circuit, the two contributing resistances can be combined to give the equivalent resistance  $R_{eq}$ . The thermal noise is given by

$$I_{T,RMS}^2 = \frac{4kTB}{R_{eq}} \quad (6.19)$$

Taking the RMS value of the average photocurrent, the signal-to-noise ratio (SNR), which is simply the mean-squared photocurrent divided by the variance (due to thermal and shot noise),

is given by the following:

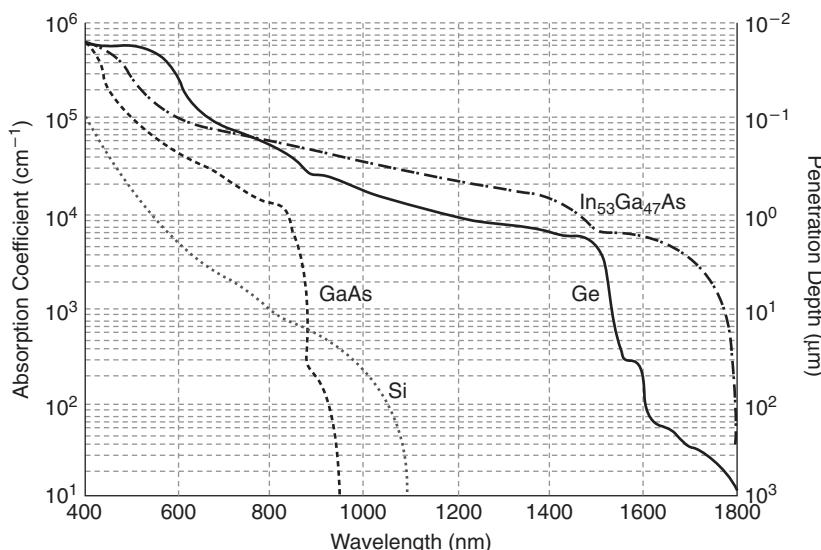
$$\text{SNR} = \frac{(\Re P / \sqrt{2})^2}{2qB(I_{\text{op}} + I_{\text{D}}) + 4kTB/R_{\text{eq}}} \quad (6.20)$$

The minimum optical power  $P$  required to produce a measurable signal for a given bandwidth and SNR, can be determined from this equation. When measuring a modulated optical signal for example, it becomes especially important to know the range of optical powers for which the detector is useful.

## 6.4 Sub-bandgap Detection

One of the most important properties of any semiconductor-based photodetector is its bandgap because this determines the wavelength range over which charge carriers may be generated by incident photons (as described in Section 6.2.1). Figure 6.10 shows the absorption coefficient as a function of wavelength for the elemental semiconductors Si and Ge, the compound semiconductor GaAs and the ternary semiconductor  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ . The detection cut-off wavelengths are clearly distinguishable [9].

While optical detectors, fabricated using silicon device technology, have been available since the concept of silicon integrated circuits was conceived, the silicon bandgap of  $\sim 1.1\text{ eV}$ , ensures that they are most commonly marketed with sensitivity around 700 nm but, with suitability for the short-haul telecommunications wavelength of 850 nm. They are, though, incompatible with long-haul telecommunication wavelengths in the infrared. This presents a challenge for those wishing to integrate detection with other photonic functionality on a single silicon chip. Clearly, it is desirable that the signal is carried at a wavelength in the infrared ( $> 1100\text{ nm}$ ) to



**Figure 6.10** Absorption coefficients for Ge, Si and (In)GaAs. Reproduced from [9] by permission of IEEE © (2004)

avoid significant, on-chip attenuation; however this would imply virtually zero responsivity for monolithically integrated detectors or optical monitors. Research currently in progress then attempts to reconcile this contradiction in performance specification.

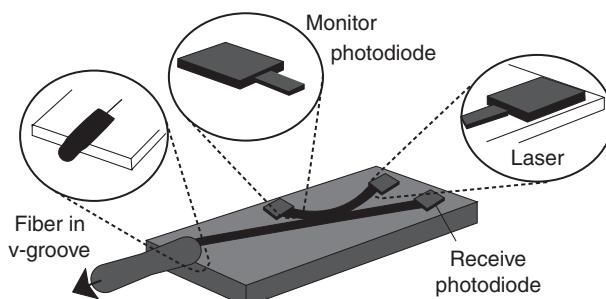
#### 6.4.1 Hybridization

The simplest approach to the integration of optical functionality is the hybridization of heterogeneous structures into a single multi-functional unit. Although a method for accurately integrated III–V components and silicon waveguides was outlined in 1992 [10] a viable high-volume fabrication technology was first pioneered by Bookham Technology in the late 1990s, work which culminated in their transceiver product capable of single or dual wavelength transmit and receive. Together with a semiconductor laser, two III–V photodetectors were mounted on a silicon-on-insulator (SOI) platform. SOI waveguides optically connected the receiver and signal monitor photodetectors and the laser to an adiabatic taper.

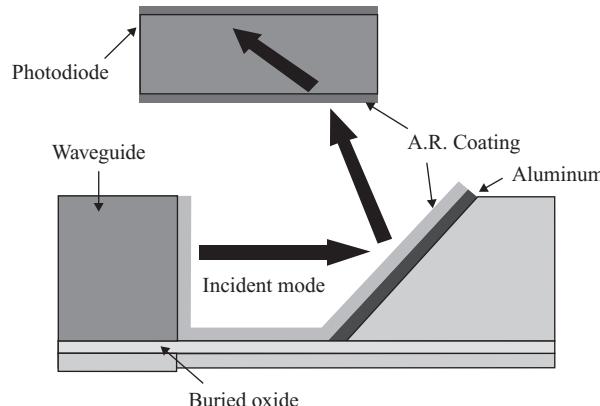
Exploiting the crystal plane etch selectivity of silicon to form deep V-grooves allowed the accurate (semi-passive) connection of a single mode fiber, thus providing external access to the optically active units. The layout of this device is shown in Figure 6.11. In 1998 Bookham reported the working specifications to include a launch power of 0.2 mW, a detection responsivity of 0.2 A/W and a data rate capability of 155 Mb/s [11].

The Bookham transceiver used reflection geometry for integrating the photodetector and the silicon waveguide. The interface was formed by producing a mirror between the two components, as shown schematically in Figure 6.12. The mirror was formed using a selective etch revealing a plane at  $54^\circ$  to the chip surface, virtually ideal for low-loss coupling of the reflected light into the photodetector.

The obvious benefits of hybridization are related to the use of discrete fabrication technologies to produce individual components with the best possible performance specifications. In the case of the transceiver, Bookham were able to combine the formation of the passive optical substrate in silicon (with associated advantages in fiber coupling and large scale manufacturing), with the optical performance of III–V-based emitters and detectors. The disadvantage of this approach is the requirement to position and reliably fix the III–V components onto the Si photonic chip. Ideally this process should be automated and compatible with CMOS fabrication, thus permitting the exploitation of integrating microelectronic circuitry in a monolithic fashion.



**Figure 6.11** A schematic layout of the Bookham Technology transceiver product. Reproduced from [11] by permission of IEEE © (1998)



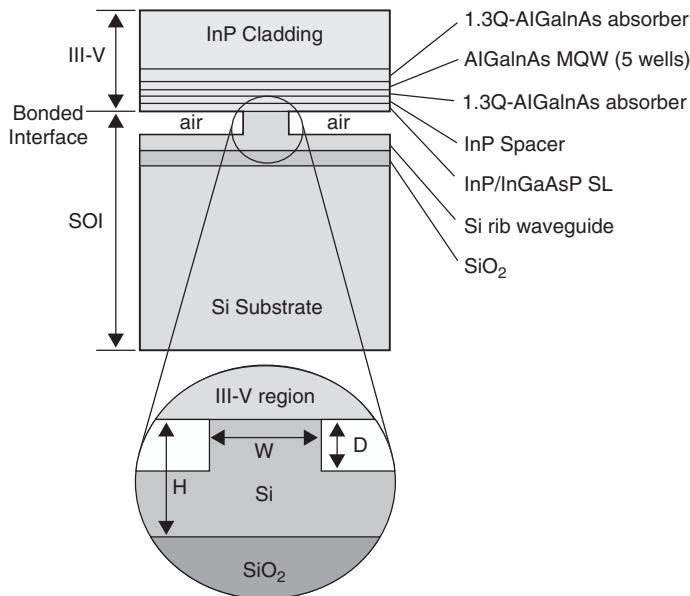
**Figure 6.12** The configuration of the photodetector and waveguide in the device layout shown in Figure 6.11. Reproduced from [11] by permission of IEEE © (1998)

Recently, work which emerged from UC Santa Barbara (in collaboration with the Intel Si photonics group) has considerably simplified the problems associated with hybridization [12]. Their unique approach utilized a silicon waveguide mode evanescently coupled to III–V semiconductor multiple quantum wells, thus combining the advantages of high-gain III–V materials and the integration capability of silicon. Moreover, the difficulty of coupling to silicon-based passive optical devices was overcome by confining most of the optical mode to the silicon. This approach restricted laser operation to the region defined by the silicon waveguide, relaxing the requirement for high precision *pick-and-place* of the III–V device on the silicon substrate. Concerns over processing compatibility of the disparate components were minimized because the bonding procedure used to attach the III–V device and the silicon is positioned at the back-end of the process flow. The fabrication thus consists of standard CMOS-compatible processing of the silicon waveguides and a low-temperature oxide-mediated wafer bonding process for heterogeneous integration. The authors reported the first demonstration of a silicon evanescently coupled laser operating at a wavelength of 1538 nm with an optically pumped threshold of 30 mW and a maximum power output of 1.4 mW. This remarkable device is shown schematically in Figure 6.13, while the calculated and observed optical mode is shown in Figure 6.14. Although designed as an emitter, it raises the possibility of a range of high-performance optical devices such as electrically pumped lasers, amplifiers, modulators, and detectors integrated on silicon.

During the final preparation stage of this chapter a significant development related to the UCSB/Intel project was announced. Whereas the device reported in [12] was optically pumped, the group had proceeded to fabricate an electrically pumped hybrid laser on a silicon waveguide [13]. This holds great promise as a method for the introduction of virtually any optical functionality (including high-performance detection) using a CMOS-compatible approach.

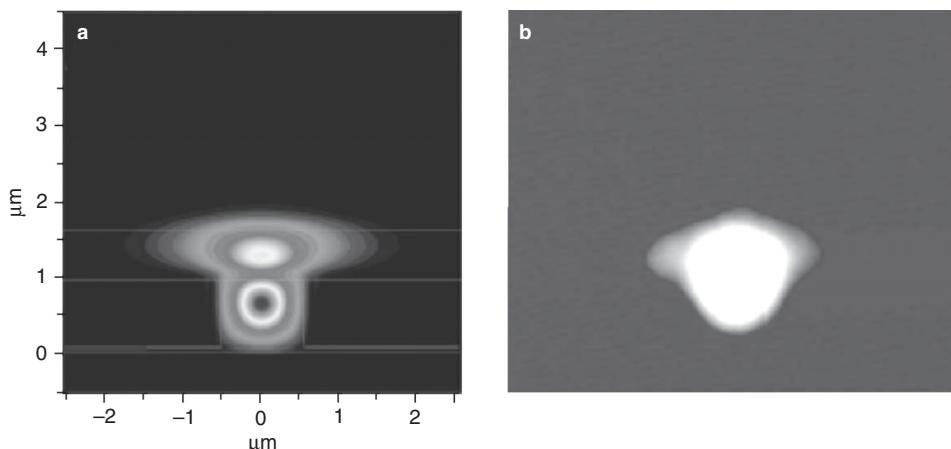
#### 6.4.2 SiGe-based Devices

As described in the introduction to Section 6.4, the transparency of silicon to wavelengths around 1300–1550 nm (the property that makes it an ideal substrate for infrared optical



**Figure 6.13** Cross-section of device structure showing an SOI rib bonded to a III–V active region grown epitaxially on InP. Reproduced from [12] by permission of the Optical Society of America

waveguiding) prevents the straightforward fabrication of monolithically integrated detectors. Absorption data summarizing this contradiction in functionality is shown in Figure 6.10. However, by adding Ge to the silicon matrix (i.e. the formation of  $Si_{1-x}Ge_x$  alloy) it is possible to shift the absorption edge from 1100 nm (for  $x = 0$ ), deeper into the infrared (coinciding with the absorption edge of Ge at 1880 nm (for  $x = 1$ ). For  $x > 0.3$ , absorption (and hence



**Figure 6.14** (a) Calculated fundamental TE mode; (b) observed lasing mode of the UCSB hybrid laser. Reproduced from [12] by permission of the Optical Society of America

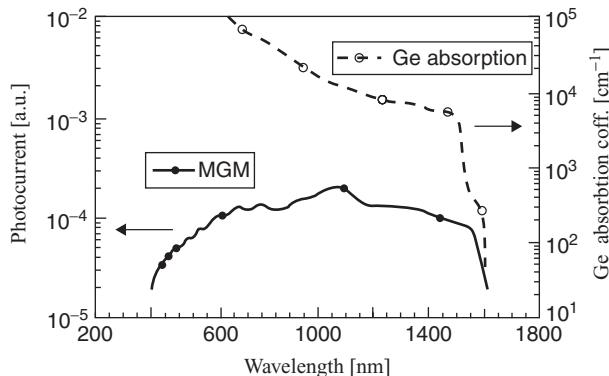
detection) of 1300 nm is possible; and for  $x > 0.85$  even 1550 nm wavelengths can no longer traverse a SiGe sample unattenuated.

The importance of Ge in the development of silicon photonics was outlined by Soref in his seminal review in 1992 [14]. There, Soref reproduced data from the *Handbook of Optical Constants of Solids* [15,16] showing the optical absorption coefficient as a function of wavelength for several compositions of unstrained SiGe. This has obvious ramifications for the design of detectors suitable for use at wavelengths in the infrared. Also reproduced was data describing the effect of Ge addition on the refractive index in the same wavelength range. Of some significance, there is a concomitant increase in refractive index with increasing Ge; a property which not only allows the fabrication of SiGe/Si passive waveguiding structures [17], but also the straightforward integration of SiGe-based optical detectors with silicon waveguides via evanescent coupling.

The heteroepitaxial growth of Ge (or Si/Ge) on a silicon substrate presents a number of issues related to the introduction of crystal defects. This is a result of the lattice mismatch of 4.2% between Si and Ge, which leads to significant strain introduction in the grown epilayer. In general, there exists a critical thickness specific to Ge concentration [18], beyond which the growth of the epilayer cannot proceed without the introduction of large concentrations of dislocations. The impact of such dislocations on the fabrication of detectors may manifest as unacceptable dark current, even for the lowest detector bias. The values of critical thickness given by [18] are too thin for efficient detection for devices fabricated in a planar geometry, providing insufficient absorption of wavelengths around 1550 nm. This has led to work which seeks to overcome the equilibrium constraints of epilayer relaxation and the subsequent introduction of defects, thus combining a large absorption coefficient with acceptable responsivity at infrared wavelengths. Here we examine several approaches to the integration of Ge and Si for detection of wavelengths in the infrared. The range of documented work in this area is vast, and hence we do not provide an exhaustive review. Instead, we have summarized work which demonstrates the clear goal of the incorporation of detector technologies for highly integrated silicon photonics.

Colace *et al.* [19] investigated the feasibility of the most obvious fabrication approach—the epitaxial growth of thick layers of pure Ge on silicon by chemical vapor deposition. To limit and confine lattice dislocations, Colace introduced a thin Ge buffer layer (50 nm thickness) grown at low temperature (350°C), where the surfactant action of hydrogen allows the deposition of flat films far thicker than the wetting layer in the normal Stranski–Krastanov growth mode. The remainder of the film was deposited at 600°C in a process demonstrating characteristics similar to homoepitaxial growth. In such a way they were able to realize flat films of thickness comparable to the penetration depth of the infrared light. Using a film of total thickness of 500 nm, the authors proceeded to fabricate photodetectors in a metal–semiconductor–metal configuration. The structure allowed straightforward processing and a subsequent fast response when closely spaced interdigitized electrodes were employed. The detector contacts were photolithographically defined silver, evaporated directly onto the Ge, with the contact via defined by a 1.5-μm-thick photoresist layer.

The spectral response of the Colace photodetector is shown in Figure 6.15, compared with the optical absorption of bulk germanium. Both show a distinctive wavelength cut-off above 1600 nm. The authors reported a responsivity at 1300 nm of 240 mA/W, with a collection efficiency of 89%. The response time of the detector was shown to be limited by the carrier drift time (2 ns for a 1 V bias). Although the paper included discussion on the impact of



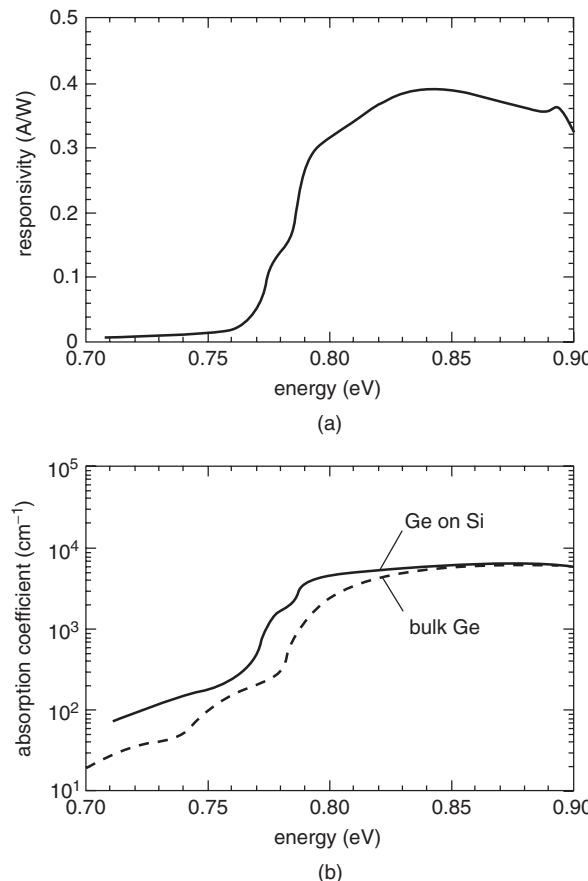
**Figure 6.15** Photocurrent spectrum of the photodetector reported in [19]. Reproduced by permission of American Institute of Physics

bulk defects on the recombination of charge carriers, there was no discussion of dislocation density or device dark current, although the high number of reported recombination centers ( $>10^{14} \text{ cm}^{-3}$ ), implies a significant concentration of extended lattice defects.

This initial and important work was extended by the same authors and the MIT group headed by Kimerling. In reference [20] a cyclic thermal annealing process in the temperature range 700–900°C after the deposition of Ge was shown to reduce the concentration of lattice dislocations from  $\sim 10^9 \text{ cm}^{-2}$  to  $2 \times 10^7 \text{ cm}^{-2}$  for large area growth, whereas selective area growth of small mesas ( $10 \times 10 \mu\text{m}$ ) could be achieved with no observable dislocation density. Furthermore, fabricated  $p-i-n$  photodiodes showed high responsivity ( $\sim 0.89$  and  $0.75 \text{ A/W}$  at 1.3 and  $1.55 \mu\text{m}$ , respectively) and a fast response time of close to 200 ps [21].

The tensile strain induced by the growth of Ge layers on silicon also affects the bandgap of these structures such that it shrinks in magnitude, extending the useful detection range well beyond that expected through the introduction of the Ge alone [22]. Generation of such a tensile strain can be explained by thermal expansion mismatch. Strain due to the 4% difference in the lattice constant between Ge and Si should be absent during the initial stage of the Ge growth. However, a decrease from the growth/annealing temperature to room temperature causes a strain due to the difference in thermal expansion coefficient between Ge and Si. Because Ge has a larger expansion coefficient, the decrease of the lattice constant of the Ge layer during cooling is suppressed by the thick Si substrate, resulting in a permanent tensile strain of around 0.2% in the Ge layer. Figure 6.16(a) shows the responsivity spectrum for a Ge photodiode fabricated on a  $p$ -type Si substrate. Above 0.79 eV, responsivity of the diode is greater than  $0.3 \text{ A/W}$ . In order to compare the result for the strained layer with that expected for bulk Ge, the absorption coefficient of the Ge layer was calculated and compared with that of a Ge crystalline sample. The result is shown in Figure 6.16(b). It is clear that the increase in absorption of the strained Ge layer starts at a lower energy than that of the bulk. This suggests a shift of absorption edge towards longer wavelengths, i.e. shrinkage of the bandgap.

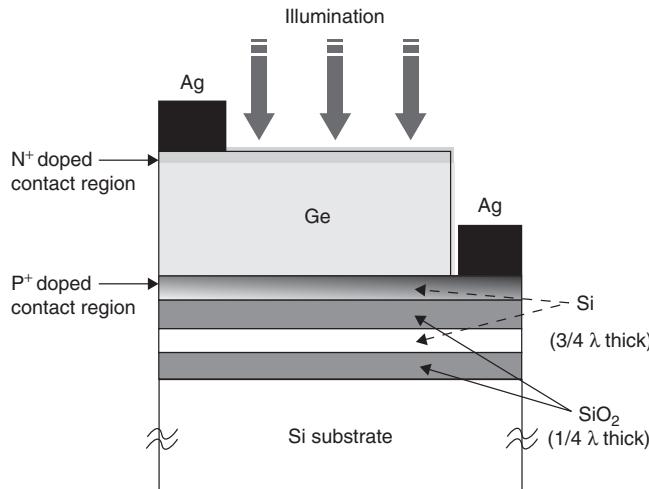
A significant increase in the effective absorption coefficient at 1550 nm for a thin-film Ge/Si heterostructure was demonstrated by Dosunmu *et al.* [9]. They exploited the high refractive index contrast of Si and  $\text{SiO}_2$  to create a resonant cavity-enhanced photodetector consisting of a Ge film of several hundred nanometers, grown on a double-layer SOI (DSOI) substrate. In a



**Figure 6.16** (a) Responsivity spectrum for Ge-on-Si photodiode; (b) absorption coefficients for Ge on Si (solid line) and bulk Ge (dashed). Reproduced from [22] by permission of American Institute of Physics

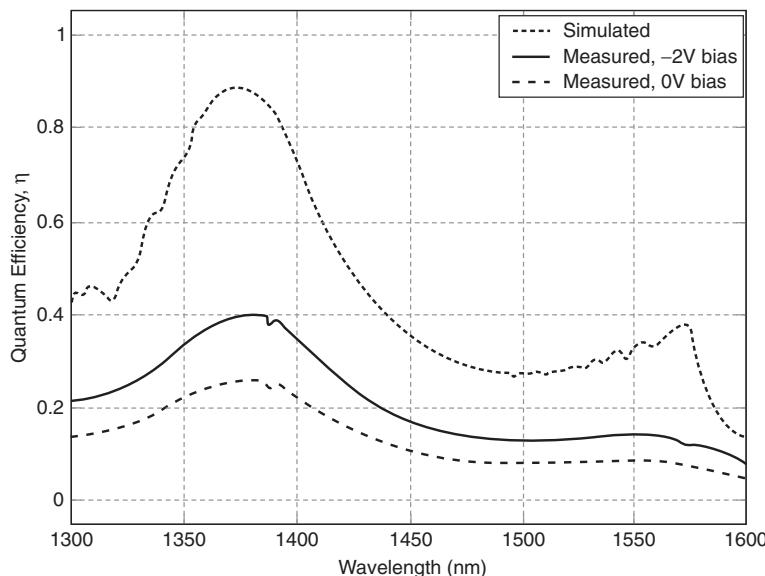
manner consistent with the work described above, a low-temperature Ge buffer was inserted between the DSOI and Ge detection region to minimize the dislocations usually introduced in heterolayer growth. Figure 6.17 shows a cross-sectional view of a Ge-DSOI detector design which is completed by the addition of a *p*-type doped region in the upper silicon layer and a *n*-type doped region in the Ge detection volume. The detector response to wavelengths ranging from 1300 to 1600 nm is shown in Figure 6.18.

For a detector area of  $140 \times 140 \mu\text{m}$  a relatively large dark current of  $54 \mu\text{A}$  was observed for a reverse bias of 1 V. This was attributed to the presence of dislocations in the Ge detection volume (despite the use of the low-temperature buffer layer). The quantum efficiency  $\eta$  at 1550 nm was increased four-fold compared with the expected response of a similarly designed single-pass detector. A significant increase in  $\eta$  was predicted for optimized layer thicknesses, while it was proposed that cyclic annealing during layer growth would suppress dark current. Simulations suggested a bandwidth approaching 25 GHz.



**Figure 6.17** Cross-sectional view of a top illuminated Ge-DSOI vertical  $p$ - $i$ - $n$  photodetector. Reproduced from [9] by permission of IEEE © (2004)

Despite the outstanding performance of detectors fabricated by epitaxial Ge, the high temperatures and aggressive substrate cleaning processes required to ensure a low density of defects somewhat hinders the seamless integration of these devices with standard silicon integrated circuits. Masini *et al.* [23] therefore developed a low-temperature ( $300^\circ\text{C}$ ) deposition technique to form a 120 nm polycrystalline Ge film on a silicon substrate. These films exhibited absorption



**Figure 6.18** Ge-SOI photodetector quantum efficiency as a function of wavelength and reverse bias. Reproduced from [9] by permission of IEEE © (2004)

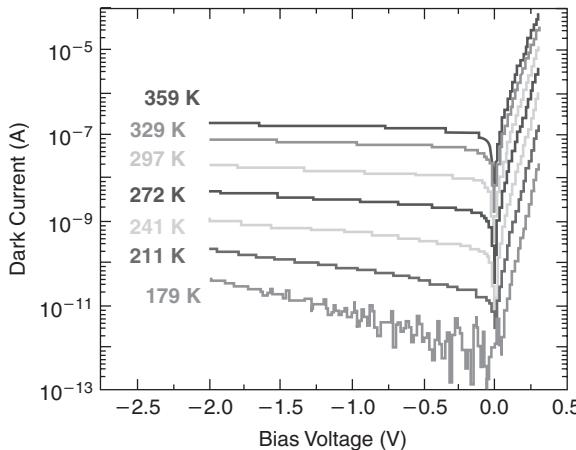
spectra similar to those of monocrystalline Ge, but, due to the lower material quality, carrier mobility and lifetime were significantly reduced. Despite the relatively poor-quality film, responsivities of 16 and 5 mA/W were measured in fabricated devices for optical wavelengths of 1300 and 1550 nm, respectively, with an associated dark current of  $1 \text{ mA cm}^{-2}$ . For a reverse heterojunction bias of 30 V, the device was shown to be capable of measuring optical source modulation at rates in excess of  $2.5 \text{ Gbits s}^{-1}$ .

Perhaps the strongest evidence for the potential of SiGe in silicon OICs is its reported adoption by two of the most influential industrial research groups for detection in silicon-based integrated optical circuits. At the third IEEE meeting on group IV photonics Morse *et al.* of Intel Corp. described the fabrication process for a Ge on Si photodetector [24]. Epitaxial Ge films were grown on a *p*-type silicon substrate in a commercial CVD reactor. The growth included an initial seed layer of  $0.1 \mu\text{m}$  Ge deposited at a temperature between 350 and  $400^\circ\text{C}$ , followed by a thicker Ge film grown between 670 and  $725^\circ\text{C}$ . Post-growth, circular mesas were etched through the Ge film down to the silicon substrate. The films were passivated with amorphous silicon and  $\text{Si}_3\text{N}_4$  and then annealed at  $900^\circ\text{C}$  for 100 min before contacts were added via ion implantation and aluminum metallization. The dislocation density in the devices was determined to be  $\sim 1 \times 10^7 \text{ cm}^{-2}$ . The reported optical characterization of the fabricated devices was performed at 850 nm, with emphasis being placed on the detector performance relative to commercially available GaAs structures. The leakage current for  $50 \mu\text{m}$  diameter mesa structures was  $\sim 1 \mu\text{A}$  for a reverse bias of 3 V, whereas the responsivity was found to saturate for a Ge film thickness of  $1.5 \mu\text{m}$  at  $0.6 \text{ A/W}$ . The bandwidth was determined to be  $\sim 9 \text{ GHz}$ . The authors assert that these characteristics position such detectors close to GaAs in commercial viability.

At the same meeting Koester *et al.* [25] reported an update on Ge/SOI detector technology under development at IBM. Previously, the IBM group had demonstrated the successful fabrication of lateral *p-i-n* Ge-on-SOI photodetectors with bandwidths as high as 29 GHz [26], while also showing that these device geometries, combined with a CMOS IC, could produce error-free operation at 19 Gb/s. In reference [25] they addressed one of the major issues of concern to those wishing to integrate Ge detectors on Si substrates—temperature sensitivity, particularly as it relates to the issue of dark current. This is a direct consequence of the relatively small bandgap of Ge and the potential for high numbers of defects when Ge is grown directly on Si, as outlined previously. Figure 6.19 shows the measured variation of dark current as a function of reverse bias for temperatures ranging from 179 to  $359 \text{ K}$ . Analysis of these results showed that the dark current generation mechanism has a distinctive activation energy close to half that of the bandgap, confirming the dominant role of trap-(defect)-assisted carrier generation. The authors acknowledged the consistency of this result with the relatively high concentration of defects ( $\sim 10^8 \text{ cm}^{-2}$ ) in the devices, but proceeded to show that this did not impact the  $10 \text{ Gbs}^{-1}$  performance of the detectors at an elevated temperature of  $85^\circ\text{C}$ .

#### 6.4.3 Integration of Ge Detectors with SOI Waveguide Geometries

The preceding sub-section described work which has contributed to the remarkable progress in the development of Ge on Si growth technology. It is clear that under the correct fabrication conditions, direct growth of Ge on Si is suitable for the production of highly responsive, high-bandwidth detectors which may be used for all perceivable applications. It is now worth considering how these detectors may be integrated in a waveguide geometry, particularly how

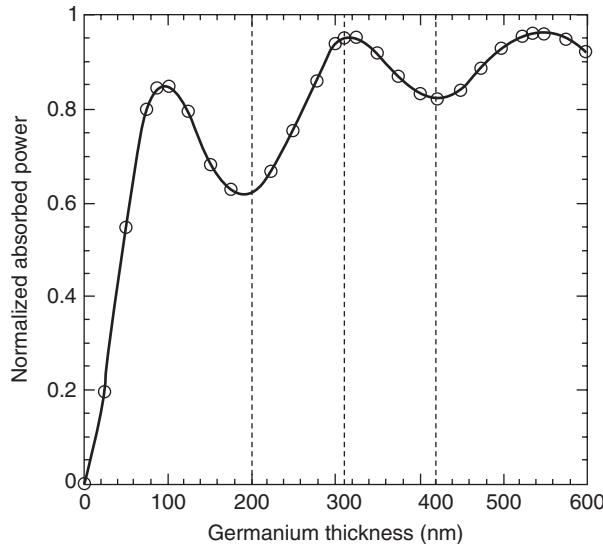


**Figure 6.19** Dark current versus reverse bias for Ge on SOI photodetector. Reproduced from [25] by permission of IEEE © (2006)

that might be achieved in waveguides based on silicon-on-insulator (SOI) and more specifically silicon-on-oxide.

In 2005, Rouvière *et al.* [27] outlined the challenge of Ge detector and waveguide integration for intra-chip optical interconnects, particularly with respect to the replacement of metal–dielectric stacks, as outlined by the International Technology Roadmap for Semiconductors. In addition to summarizing many of the strategies for hetero-epitaxy described in the preceding sections of this chapter, their paper investigated two potential approaches for integration: butt-coupled detectors inserted into trenches recess etched into an SOI waveguide, and vertically coupled detectors which rely on the relative increase in refractive index of Ge compared to silicon. They performed three-dimensional FDTD simulations to assess these approaches for use with a small-cross-section (sub-micrometer) SOI waveguide. In the case of butt-coupling, for a 1.3 μm optical signal almost all the light was shown to be transmitted from the Si waveguide to the Ge layer, with losses by reflection and diffraction at the waveguide discontinuity being negligible. The high absorption coefficient of the Ge required a detector length of only 4 μm for >95% absorption. The authors noted though that a significant disadvantage of this configuration was the requirement for accurate control of the etched depth of the silicon recess into which the Ge is deposited, allowing a sufficiently thin Si layer of around 30 nm for Ge growth above the buried silicon oxide.

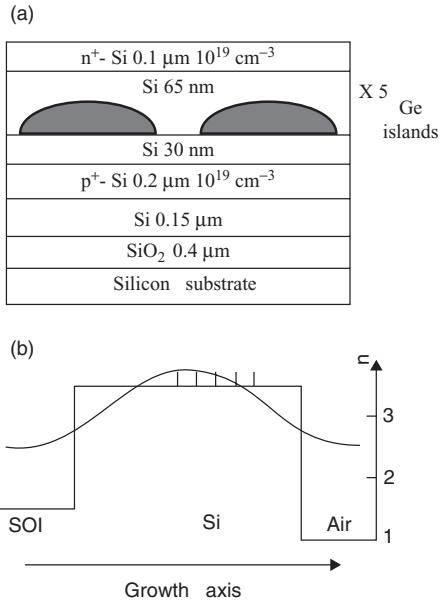
The authors conceded then that vertical coupling offers a simpler solution from a fabrication viewpoint (i.e. Ge growth can be performed without any additional etching step). In this case, the coupling efficiency is determined by the Ge layer thickness and the coupling length. Results reproduced from [27] are shown in Figure 6.20 for a Ge thickness up to 600 nm. Clearly observable are oscillations which were attributed to a Fabry–Perot effect in the silicon–germanium bilayer. Although there is a trend of increased absorption with an increase Ge thickness, this is not then monotonic. The local maximum in absorption at approximately 310 nm is well suited to photodetector integration with SOI submicron optical waveguides, which have comparable dimensions. For this Ge layer thickness, more than 95% of light absorption has occurred within the 7 μm device length.



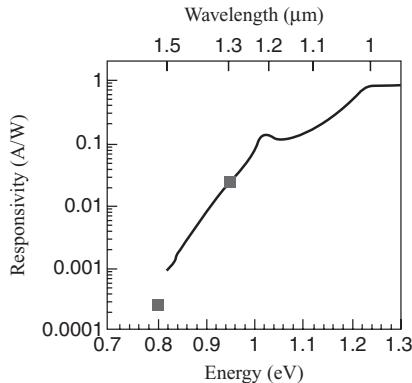
**Figure 6.20** Three-dimensional field-mode matching simulation of absorption of 1.3  $\mu\text{m}$  optical signal in a 7- $\mu\text{m}$ -long vertical coupled Ge/SOI photodetector. Reproduced from [27] by permission of SPIE

Preceding the paper by Rouvière *et al.*, practical realization of the integration of Ge and an SOI waveguide to form an in-line photodetector was reported by El kurd *et al.* in 2002 [28]. The photoresponsive element was grown on an SOI substrate and consisted of multiple layers of self-assembled Ge islands, terminated with a silicon cap. A large-cross-section waveguide was formed via dry etch, allowing the fabrication of a vertical  $p-i-n$  structure. Figure 6.21 shows a schematic diagram of the device structure and the index contrast with the lowest-order TE mode at 1.3  $\mu\text{m}$  superimposed, while Figure 6.22 shows the photoresponse of a typical device. The external quantum efficiency of this photodetector device for a 0 V applied bias was 2.35% at 1.3  $\mu\text{m}$  and 0.02% at 1.55  $\mu\text{m}$ . The responsivities plotted in Figure 6.21 are not corrected for the coupling efficiency into the waveguide. The authors noted that the efficiency would be increased via the use of waveguide tapers to improve coupling, and also by applying a reverse bias to the diode. For the device reported in [27] though, the dark current at a high-reverse applied bias induced an unacceptable dark current.

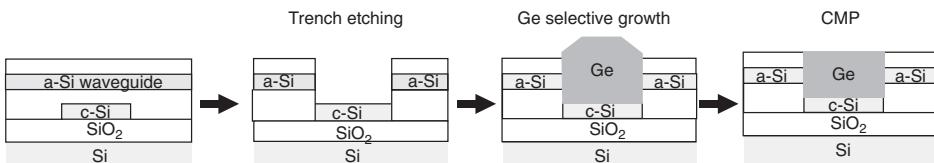
At the time of writing, the most recent report of Ge detector and waveguide integration originates from the MIT group, which also contributed a large volume of knowledge with regard to direct Ge growth on Si, outlined in Section 6.4.2. The MIT work was published in the proceedings of the third IEEE meeting on group IV photonics in Ottawa [29]. Of significance, the processing of the devices took place using a 180 nm industrial CMOS production line, thus demonstrating one of the advantages of silicon photonics—CMOS process compatibility. The waveguide section of the device was formed using low-loss amorphous silicon deposited on an SOI wafer. Optical butt-coupling from the amorphous silicon waveguides to Ge photodetectors was achieved by the selective growth of Ge in trenches opened through a  $\text{SiO}_2$  mask layer, followed by a chemical mechanical polishing process, required for planarization. This process is shown schematically in Figure 6.23.



**Figure 6.21** (a) Schematic diagram of the waveguide photodetector from El kurdi, showing the vertical stacking of the grown layers; (b) refractive index variation for the silicon-on-insulator structure. The profile of the first confined optical mode in TE polarization at  $1.3 \mu\text{m}$  is superimposed on the index variation. The island layers appear as five spikes. Reproduced from [28]



**Figure 6.22** Responsivity of the device described in Figure 6.21, measured at room temperature with a 0 V applied bias. The full curve corresponds to a calibrated measurement with a Fourier-transform infrared spectrometer. The squares correspond to measurements at  $1.3$  and  $1.55 \mu\text{m}$  with laser diodes. Reproduced from [28]



**Figure 6.23** Schematic diagram outlining the process flow used to form butt-coupled Ge photodetectors in a SOI waveguide. Reproduced from [29] by permission of IEEE © (2006)

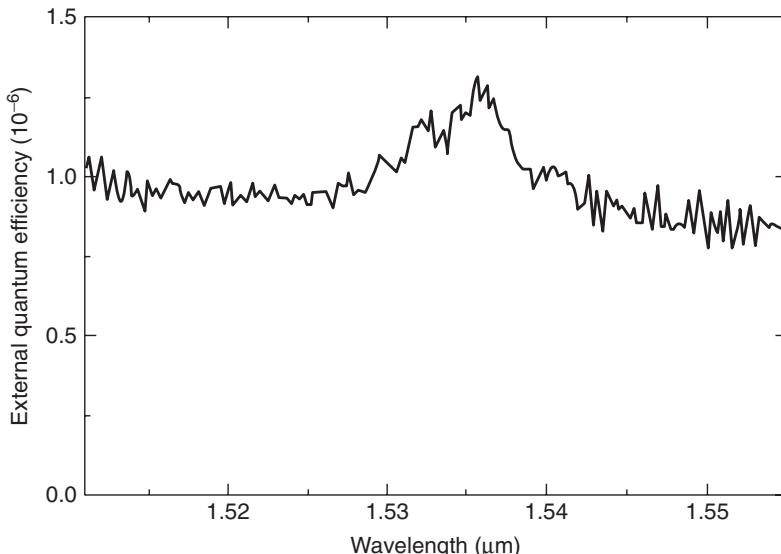
The trenches were 600 nm wide and 600 nm deep to facilitate the coupling from the Si waveguide to the Ge detector by maximizing the optical mode overlap between the detector and the waveguide. The coupling loss from the Si waveguide to the Ge detector was estimated to be 0.7 dB from simulation, and experimentally determined to be 0.8 dB. The underlying *c*-Si region was *p*-doped, and a *p*–*i*–*n* diode structure was formed by the deposition of an *n*<sup>+</sup> poly-Si top electrode. The responsivity of the detector was determined to be close to 1 A/W for an input wavelength of 1520 nm and a detector length of 50 μm. The theoretical detector bandwidth was 30 GHz, limited by carrier transit time. The experimental bandwidth was shown to be only 4.5 GHz, however the authors stated that this was limited by the external amplifier circuit and not the device itself.

The use of germanium as the optically responsive element for detectors compatible with silicon photonics constitutes the majority of work to date with regard to sub-bandgap detection. The obvious success of this approach is somewhat tempered by the material growth and waveguide integration issues, often manifesting as high dark current or sensitivity to the polarization of the incident light coupled from an SOI waveguide respectively. Further, while the high optical absorption provides for excellent terminal detection, there often exists the need for monitored detection where optical signals are sampled and the vast majority of the signal intensity proceeds in the optical circuit unperturbed. The following sub-sections describe the use of optical dopants (chemical and structural) for sub-bandgap detection. The advantages of these approaches are related to the ease of processing, usually at the cost of performance in either responsivity, bandwidth or both.

#### 6.4.4 Sub-bandgap Detection via Chemical Doping

Kik *et al.* [30] suggested and demonstrated a waveguide detector which used an Er-doped *p*–*n* structure, integrated with an SOI waveguide formed from a substrate with a 3-μm-thick silicon overlayer. It is well-known that Er<sup>3+</sup> can be electrically excited, subsequently emitting photons with an associated wavelength close to 1.5 μm, due to an electronic transition in its incompletely filled 4*f* shell. Indeed, silicon diodes doped with Er ions have proven to be one method for the fabrication of Si-based IR-LEDs [31]. Conversely, absorption in an Er-doped device may also occur whereby an Er-doped Si *p*–*n* junction can generate a photocurrent when illuminated with 1.5 μm light. Previous work by the same group using a planar diode geometry [32] determined that excited Er can generate electron–hole pairs with an efficiency as high as 70% at room temperature. However, due to the small optical absorption cross-section of the Er ions and the limited width of the depletion region when light is coupled normal to the device surface, the fraction of the light that is absorbed at normal incidence is small. To improve the absorption efficiency, a high Er concentration and a large interaction length are required, such as that provided by containing the dopant and optical signal in a waveguide configuration. The authors fabricated their Er-doped Si *p*–*n* junction waveguide detector using a 3 MeV Er implant to a fluence of 10<sup>13</sup> cm<sup>-2</sup>. A boron implant into the top of the rib waveguide, and a phosphorus implant on either side of the rib constituted the contacts for the *p*–*n* structure.

Photoreponse measurements were performed at room temperature by coupling the output of a tunable diode laser (1.51–1.57 μm) into the detector waveguide using a single-mode tapered fiber. The collected current was converted to external quantum efficiency  $\eta_{\text{ext}}$  by dividing the photocurrent by the 1.5 μm photon flux in the input fiber, measured using a calibrated detector. Figure 6.24 shows  $\eta_{\text{ext}}$  as a function of wavelength measured at a forward bias of 0.4 V. The



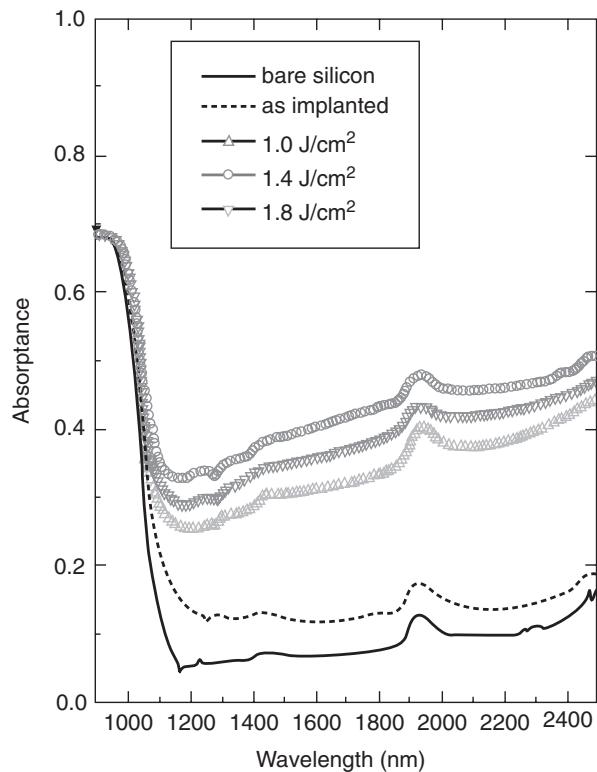
**Figure 6.24** External quantum efficiency of an Er-doped  $p$ - $n$  Si waveguide detector fabricated using standard SOI material and processing technology. Reproduced from [30] by permission of IEEE © (2002)

spectrum shows a peak centered around  $1.535\text{ }\mu\text{m}$ , attributed to an internal transition in the  $\text{Er}^{3+}$  ion. The absolute value of  $\sim 10^{-6}$  is 2–3 orders of magnitude smaller than that expected, a discrepancy attributed by the authors to the geometry of the device contacts. The reported device would thus likely yield photocurrents of the order of nanoamps, for a coupled power of milliwatts.

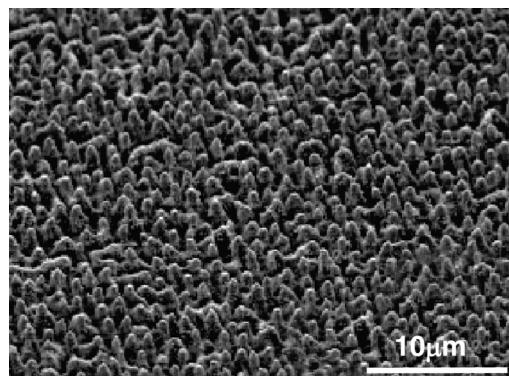
In addition to Er, recent work at Harvard University has shown that silicon, heavily doped ( $\sim 1\text{ atm\%}$ ) with sulfur, exhibits strong absorbance of sub-bandgap light. This is true for the case of silicon prepared by laser pulsing in a  $\text{SF}_6$  atmosphere, resulting in a microstructured surface [33], and also for samples ion implanted with  $\text{S}^+$  ions and recrystallized using a  $\text{XeCl}^+$  excimer laser [34]. In the latter case, strong absorption, summarized in Figure 6.25, was observed for samples with a planar geometry with no detectable extended defects.

In [33], Huang *et al.* explored the potential of sulfur-doped silicon as the basis of photodetector devices by fabricating planar devices with a diameter of up to  $500\text{ }\mu\text{m}$ , using standard processing technology. Ti:sapphire irradiation in a  $\text{SF}_6$  atmosphere created a microstructured surface with an  $n/n^+$  heterojunction (the surface being heavily  $n$ -doped with respect to the substrate). An electron micrograph of the surface is shown in Figure 6.26. Device metallization resulted in a Schottky barrier at the substrate, while an Ohmic contact was formed at the surface. The responsivity of a  $100\text{ }\mu\text{m}$  diameter device for  $1.55\text{ }\mu\text{m}$  light was determined to be  $20\text{ mA/W}$ .

The authors claimed that the high infrared absorption mechanism in the microstructured Si results directly from the creation of mid-gap energy states due to both the sulfur and microstructural defects (a description of the effects of crystal defects on optical absorption is expanded upon below).



**Figure 6.25** Relative optical absorption versus wavelength for silicon implanted with 200 keV S<sup>+</sup> ions to a dose of  $1 \times 10^{16} \text{ cm}^{-2}$ , subsequently irradiated by a pulsed XeCl<sup>+</sup> excimer laser. Reproduced from [34] by permission of American Institute of Physics



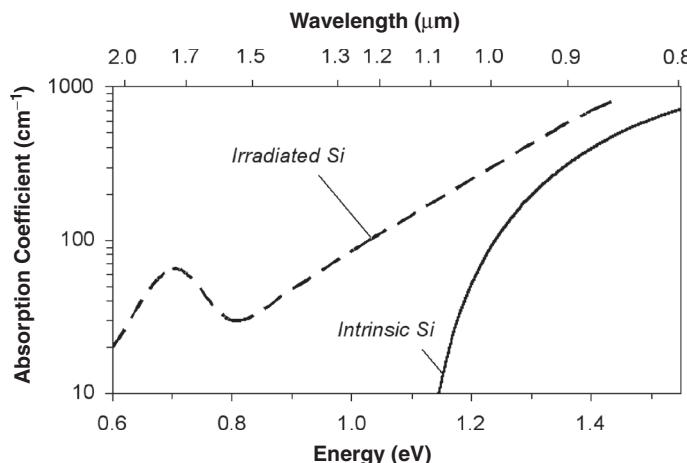
**Figure 6.26** Electron micrograph of the surface of microstructured silicon used in the fabrication of planar geometry IR detectors. Reproduced from [33] by permission of American Institute of Physics

#### 6.4.5 Defect-enhanced Infrared Response

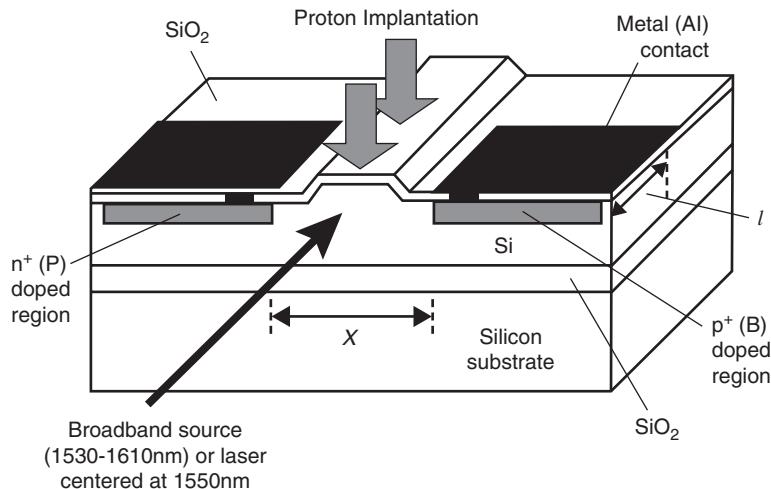
It has been known for five decades that the absorption edge of particle-irradiated silicon could be extended to wavelengths beyond that for intrinsic silicon [35]. For instance Figure 6.27 shows a typical plot of the absorption spectrum of silicon irradiated with a sub-amorphizing dose of energetic particles [36]. When crystalline silicon is irradiated with high-energy ions (available via standard ion implantation) mobile defects such as interstitials and vacancies are produced. At room temperature, most of these defects rapidly reincorporate into the crystal structure. However, some combine to form complex, stable defects such as the silicon divacancy—a pair of empty lattice sites [37].

A defect such as the divacancy disrupts the normal energy band structure of the material and introduces mid-bandgap levels. The photon energy required to excite an electron from such levels is lower than the energy normally required for valence-to-conduction-band excitation, allowing the absorption of lower energy photons. Since the  $1.8\ \mu\text{m}$  absorption band tail shown in Figure 6.27 passes through wavelengths around 1550 nm, the study of this excess optical absorption is of potential importance to the development of silicon photonics in general and integrated detectors in particular.

The first demonstration of an integrated detector using defect engineering in SOI was described by Bookham Technology in 2003 [38]. Further insights into the design parameters of this device were reported by Bradley *et al.* from McMaster University [39]. Their device structure is shown in Figure 6.28. It consists of a rib waveguide with integrated *p*-doped and *n*-doped contacts on either side of the rib, with the addition of a dilute concentration of vacancy-type defects into a volume consistent with that occupied by the optical mode. The defects were introduced via low-energy proton irradiation. Figure 6.29 shows the response of this in-line detector to 1550 nm light. The IR sensitivity is enhanced relative to a similar device without proton irradiation by approximately two orders of magnitude, with a measured responsivity of 3 mA/W. Of some note is the capability of the fabrication process to determine the fraction of

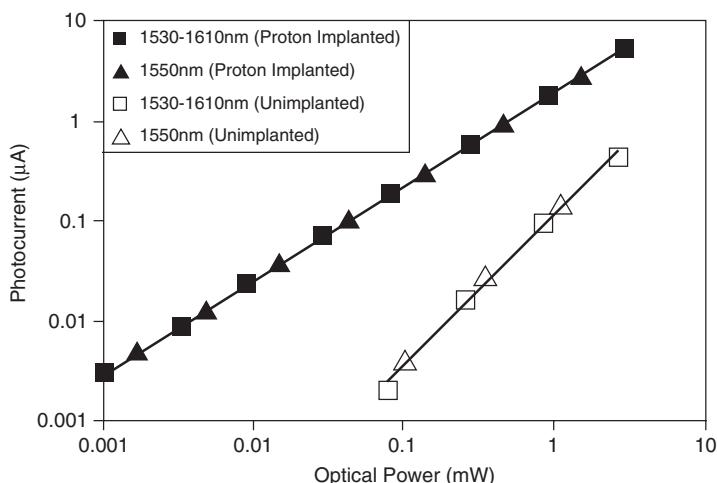


**Figure 6.27** Typical absorption versus wavelength for silicon irradiated with a sub-amorphising dose of energetic particles, and that for unirradiated silicon. Reproduced from [36] by permission of SPIE

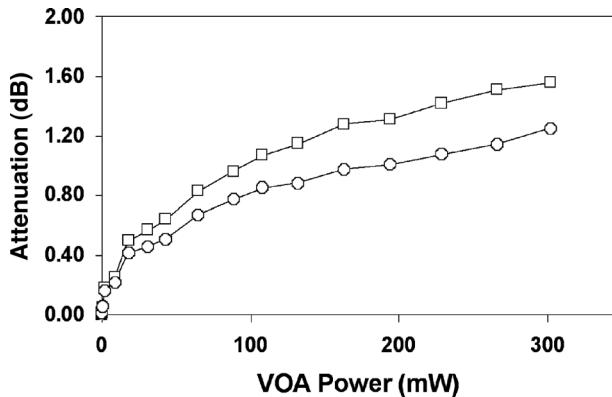


**Figure 6.28** Illustration of monolithically integrated in-line power monitor for sub-bandgap detection in an SOI waveguide. Reproduced from [39] by permission of American Institute of Physics

light which is removed from the waveguide and converted into an electrical signal. In [39], this was 19%, however, values from less than 1% could be absorbed simply by adjusting the proton irradiation dose. The authors suggested therefore that these detectors might find greatest application as in-line monitors. In a subsequent report the same group demonstrated monolithic integration of a carrier injection variable optical attenuator (VOA) and detector [36]. Results



**Figure 6.29** Photocurrent as a function of on-chip optical power for a proton implanted and an unimplanted device described in Figure 6.28. The photoresponse, shown for both a broadband source and laser centered at 1550 nm, is significantly enhanced in the proton implanted device. Reproduced from [39] by permission of American Institute of Physics



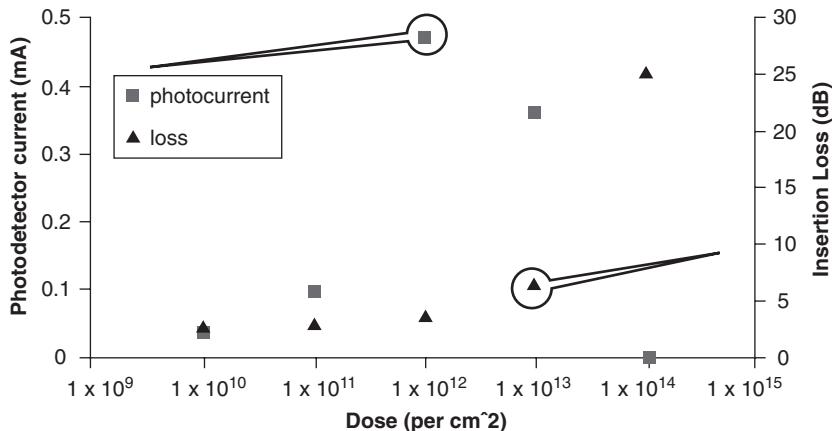
**Figure 6.30** Attenuation of on-chip power by integrated VOA measured by external detector (open squares) and by the integrated photodetector (open circles). Reproduced from [36] by permission of SPIE

from this first demonstration of monolithic integration of variable attenuation and monitoring are reproduced in Figure 6.30. The authors noted the poor efficiency of the VOA (optical attenuation >40 dB is common for carrier injection VOAs), which was of nonoptimal design and not the focus of their work.

One disadvantage of the McMaster defect-mediated detector is the relatively small bandwidth. For the devices described in [36] and [39], the photosensitive region of the waveguide is required to be >2 mm in length, while the waveguide cross-section is approximately 5  $\mu\text{m} \times$  4  $\mu\text{m}$ . This leads to a bandwidth of a few MHz, suitable for monitoring signals undergoing thermo-optic or carrier injection modulation only.

More recently, Liu *et al.* [40] showed that the performance of defect enhanced detectors could be improved by the application of a low temperature (<350°C) annealing step. They fabricated SOI waveguides 4  $\mu\text{m}$  in width, integrating *p*-doped and *n*-doped contacts in a manner similar to the device shown in Figure 6.28. However, in place of protons they used 800 keV He<sup>+</sup> ions to introduce mid-gap defect states below the rib. The authors proposed that helium would result in relatively stable complex defects in the silicon lattice, more amenable to post-implantation annealing. Figure 6.31 shows both the response of the waveguide detectors and the total waveguide loss, versus He<sup>+</sup> ion dose, for an input optical power of 25 mW at a wavelength of 1440 nm. The photocurrent is shown to increase initially with increasing ion dose, up to a maximum value of 0.47 mA for a dose of  $1 \times 10^{12} \text{ cm}^{-2}$ , and then decrease for doses  $>1 \times 10^{12} \text{ cm}^{-2}$ , as the concentration of defects begins to significantly reduce the generated carrier lifetime. The excess optical loss however demonstrates a monotonic increase with implanted ion dose, a phenomenon which has been well characterized by Foster *et al.* for SOI waveguides [41]. The authors reported a 300% increase in responsivity to 0.064 A/W for the samples implanted with  $1 \times 10^{12} \text{ cm}^{-2}$  He<sup>+</sup> ions following annealing at 200°C for 45 min. The implication for this result is the selective removal of ion implantation damage, causing optical loss, without the removal of the defects which contribute to the enhanced infrared responsivity.

The most recent report of defect-enhanced photodetection by Geis *et al.* [42] originates from the Lincoln laboratory located at MIT. This highly significant work addressed the issues of

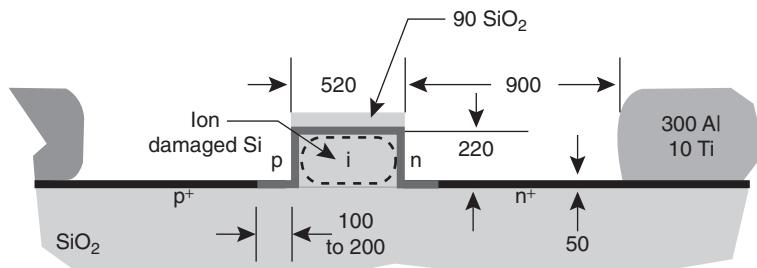


**Figure 6.31** Photoresponse and insertion loss versus  $\text{He}^+$  ion dose for SOI integrated waveguide photodetector. Data shown are for an optical input power of 25 mW at a wavelength of 1440 nm. Reproduced from [40] by permission of IEEE © (2006)

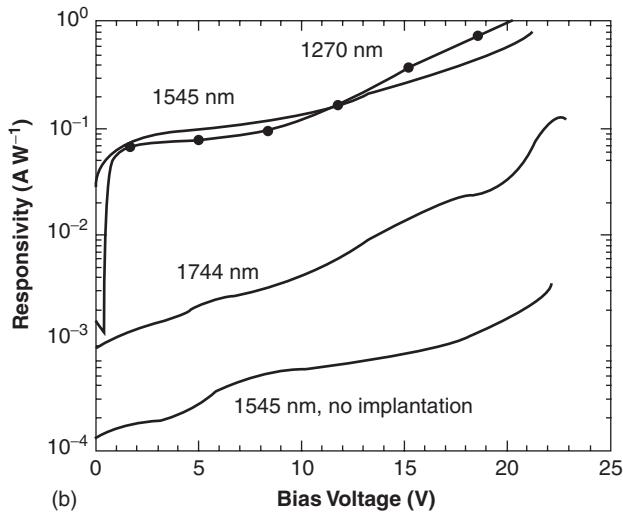
bandwidth and responsivity, previously shown to be inferior to other technologies. The basic structure of the detector was similar to that of the device in Figure 6.28, however the cross-section of the SOI rib waveguide was only  $520 \times 220$  nm, with a 50 nm slab region connecting the rib volume to electrical contacts 900 nm from the edge of the rib. The photosensitive region was introduced using  $\text{Si}^+$  ion implantation at an energy of 190 keV and dose of  $1 \times 10^{13} \text{ cm}^{-2}$ . The structure is shown in Figure 6.32.

A 3-mm-long device was reported to absorb 99% of the incident light at a wavelength of 1545 nm, coupled from a lensed fiber. Figure 6.33 shows the responsivity of the detector for a range of incident wavelengths and detector bias. At a reverse bias of 25V, the responsivity exceeds unity quantum efficiency, indicating carrier multiplication in the strong electric field.

As with [36] and [40], the authors demonstrated the thermal stability of these devices (tentatively associated with the presence of oxygen), showing that annealing at  $300^\circ\text{C}$  increased the detector responsivity. Perhaps the most important result from [42] is that associated with bandwidth. For a detector of length 250  $\mu\text{m}$  the frequency response was measured using a vector

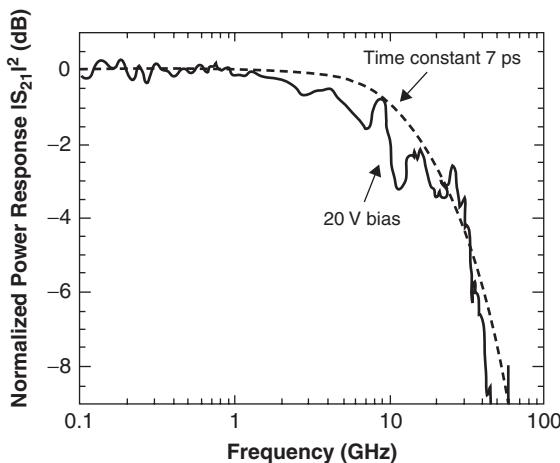


**Figure 6.32** The device structure of the Lincoln lab integrated photodetector. Reproduced from [42] by permission of IEEE © (2007)



**Figure 6.33** Responsivity for the detector structure shown in Figure 6.32, at 1270, 1545 and 1744 nm and a control device at 1545 nm. Reproduced from [42] by permission of IEEE © (2007)

network analyzer and an optical modulator capable of 50 GHz operation. Figure 6.34 shows the detector frequency response after correcting for the frequency response of the modulator. The half-power point is approximately 20 GHz. The larger 3-mm-long detector, with an increased capacitance, exhibits a 3 dB bandwidth of  $\sim 2$  GHz. The authors included a summary table comparing the defect-enhanced detector characteristics with those of a high-end InGaAs detector, and those of reported Ge-based detectors. The comparison is deemed favorable, particularly



**Figure 6.34** The frequency response of the detector shown in Figure 6.32; length 0.25 mm. Reproduced from [42] by permission of IEEE © (2007)

with Ge technologies, with the further advantage that the defect-enhanced detector is integrated in a straightforward manner, using standard CMOS fabrication processes.

## 6.5 Conclusions

In this chapter we have attempted to provide both background information relevant to the development of silicon-based optical detection technologies, and to outline recent research aimed particularly at the requirement for monolithically integrated detector functionality in silicon photonic circuits operating at a wavelength around 1550 nm. With regard to the latter, the majority of work to date has been based on the incorporation of Ge into the silicon matrix, with concomitant reduction in the host bandgap. This approach has the benefit of CMOS compatibility while providing high responsivity, bandwidth in excess of 10 GHz and an increase in refractive index, allowing evanescent coupling of the optical signal from the waveguide to the detector. Alternative approaches to detection use optical doping with impurities (such as sulfur or erbium) or with point defects, thus allowing the generation of carriers via sub-bandgap transitions. These methods are more straightforward than the fabrication of localized SiGe, but to date have resulted (in the main) in detectors with inferior characteristics. Relatively recent results do suggest though that defect-mediated waveguide detectors can be fabricated with responsivities to 1550 nm of 1 A/W, and with bandwidths in excess of 20 GHz.

Finally, it is worth recalling that hybridization of silicon and other semiconductor materials has matured to the point where the integration of electrically pumped lasing with a silicon waveguide is possible. Such an approach also allows optical detection at a wide range of wavelengths, normally inaccessible to silicon devices.

With such diverse technologies available to the design engineer it is clear that the future detector requirements of all applications of silicon photonics should be met with considerable success.

## References

- [1] H. Zimmermann, *Integrated silicon optoelectronics*. Springer, New York, 2000.
- [2] H. Zimmermann, *Silicon optoelectronic integrated circuits*. Springer, Berlin, 2004.
- [3] S. L. Chuang, *Physics of optoelectronic devices*. Wiley, New York, 1995.
- [4] B. G. Streetman and S. Banerjee, *Solid state electronic devices*, 5th edn, Prentice Hall, Upper Saddle River, NJ, 2000.
- [5] S. M. Sze, *Physics of semiconductor devices*, Wiley, New York, 1981.
- [6] B. E. A. Saleh and M. C. Teich, *Fundamentals of photonics*, Wiley, New York, 1991.
- [7] M. Dagnais, R. F. Leheny and J. Crow (eds.), *Integrated optoelectronics*. Academic Press, San Diego, 1995.
- [8] J. D. Plummer, M. Deal and P. B. Griffin, *Silicon VLSI technology: fundamentals, practice and modeling*. Prentice Hall, Uppder Saddle River, NJ, 2000.
- [9] O. I. Dosunmu, D. D. Cannon, M. K. Emsley, B. Ghyselen, J. Liu, L. C. Kimerling and M. Selim Ünlü, Resonant cavity enhanced Ge photodetectors for 1550 nm operation on reflecting Si substrates, *IEEE J. Selected Topics in Quantum Electronics*, **10**, 2004, 694.
- [10] E. E. L. Friedrich, M. G. Öberg, B. Broberg, S. Nilsson and S. Valette, Hybrid integration of semiconductor lasers with Si-based single mode ridge waveguides, *J. Lightwave Technology*, **10**, 1992, 336.
- [11] T. Bestwick, ASOC™ – A Silicon-based Integrated Optical Manufacturing Technology, *Proceedings of the IEEE Electronic Components and Technology Conference*, 1998, 566.
- [12] H. Park, A. W. Fang, S. Kodama and J. E. Bowers, Hybrid silicon evanescent laser fabricated with a silicon waveguide and III-V offset quantum wells, *Optics Express*, **13**, 2005, 9460.

- [13] A. W. Fang, H. Park, O. Cohen, R. Jones, M. J. Paniccia and J. E. Bowers, Electrically pumped hybrid AlGaInAs-silicon evanescent laser, *Optics Express*, **14**, 2006, 9203.
- [14] R. A. Soref, Silicon-based Optoelectronics, *Proceedings of the IEEE*, **81**, 1993, 1687.
- [15] E. D. Palik, *Handbook of Optical Constants of Solids*, New York, Academic Press, (1985).
- [16] E. D. Palik, *Handbook of Optical Constants of Solids II*, Boston, Academic Press/Harcourt Brace, (1991).
- [17] For example, see the work outlining the conditions for single mode propagation, R. A. Soref, J. Schmidtchen and K. Petermann, Large single-mode rib waveguides in GeSi-Si and Si-onSiO<sub>2</sub>, *IEEE J. Quantum Electron.*, **27**, 1991, 1971.
- [18] J. C. Bean, Silicon-based semiconductor heterostructures: Column IV bandgap engineering, *Proc. IEEE*, **80**, 1992, 571.
- [19] L. Colace, G. Masini, F. Galluzzi, G. Assanto, G. Capellini, L. Di Gaspare, E. Palange and F. Evangelisti, Metal-semiconductor-metal near-infrared light detector based on epitaxial Ge/Si, *Appl. Phys. Lett.*, **72**, 1998, 3175.
- [20] H.-C. Luan, D. R. Lim, K. K. Lee, K. M. Chen, J. S. Sandland, K. Wada and L. C. Kimerling, *Appl. Phys. Lett.*, **75**, 1999, 2909.
- [21] S. Fama, L. Colace, G. Masini, G. Assanto and H.-C. Luan, *Appl. Phys. Lett.*, **81**, 2002, 586.
- [22] Y. Ishikawa, K. Wada, D. D. Cannon, J. Liu, H.-C. Luan and L. C. Kimerling, Strain-induced band gap shrinkage in Ge grown on Si substrate, *Appl. Phys. Lett.*, **82**, 2003, 2044.
- [23] G. Masini, L. Colace and G. Assanto, 2.5 Gbit/s polycrystalline germanium-on-silicon photodetector operating from 1.3 to 1.55 μm, *Appl. Phys. Lett.*, **82**, 2003, 2524.
- [24] M. Morse, F. Dosunmu, E. Ginsburg, Y. Chetrit and G. Sarid, 850 nm Germanium Photodetector Performance, *Proceedings of IEEE Group IV Photonics Conference*, Ottawa, 2006 p. 170.
- [25] S. J. Koester, L. Schares, C. L. Schow, G. Dehlinger and R. A. John, Temperature-dependent Analysis of Ge-on-SOI Photodetectors and Receivers, *Proceedings of IEEE Group IV Photonics Conference*, Ottawa, 2006, p. 179.
- [26] G. Dehlinger, S. J. Koester, J. D. Schaub, J. O. Chu, Q. C. Ouyang and A. Grill, *IEEE Photon. Technol. Lett.*, **16**, 2004, 2547.
- [27] M. Rouvière, M. Halbwax, J.-L. Cercus, E. Cassan, L. Vivien, D. Pascal, M. Heitzmann, J.-M. Hartmann, S. Laval, Integration of germanium waveguide photodetectors for intrachip optical interconnects, *Optical Engineering* **44**, 2005, 075402.
- [28] M. El kurdi, P. Boucaud, S. Sauvage, G. Fishman, O. Kermarrec, Y. Campidelli, D. Besahel, G. Saitn-Girons, I. Sagnes and G. Petriarche, Silicon-on-insulator waveguide photodetector with Ge/Si self-assembled islands, *J. Appl. Phys.*, **92**, 2002, 1828.
- [29] J. F. Liu, D. Pan, S. Jonghammanurak, D. Ahn, C. Y. Hong, M. Beals, L. C. Kimerling, J. Michel, A. T. Pomerene, C. Hill, M. Jaso, K. Y. Tu, Y. K. Chen, S. Patel, M. Rasras, A. White and D.M. Gill, Waveguide-integrated Ge *p-i-n* photodetectors on SOI platform, *Proceedings of IEEE Group IV Photonics Conference*, Ottawa, 2006, p. 173.
- [30] P. G. Kik, A. Polman, S. Libertino and S. Coffa, Design and performance of an erbium-doped silicon waveguide detector operating at 15 μm, *J. Lightwave Technol.*, **20**, 2002, 862.
- [31] For example, see G. T. Reed and A. P. Knights, *Silicon Photonics—An Introduction*, Chapter 8, Wiley, Chichester, (2004).
- [32] N. Hamelin, P. G. Kik, J. F. Suyver, K. Kikoin, A. Polman, A. Schnecker and F. W. Saris, Energy backtransfer and infrared photoresponse in erbium-doped silicon *p-n* diodes, *J. Appl. Phys.*, **88**, 2000, 5381.
- [33] Z. Huang, J. E. Carey, M. Liu, X. Guo, E. Mazur and J. C. Campbell, Microstructured silicon photodetector, *Appl. Phys. Lett.*, **89**, 2006, 033506.
- [34] T. G. Kim, J. M. Warrender and M. J. Aziz, Strong sub-bandgap infrared absorption in silicon supersaturated with sulfur, *Appl. Phys. Lett.*, **88**, 2006, 241902.
- [35] For example of early work, see H. Y. Fan and A. K. Ramdas, Infrared absorption and photoconductivity in irradiated silicon, *J. Appl. Physics*, **30**, 1959, 1127.
- [36] A. P. Knights, J. D. B. Bradley, S. H. Gou and P. E. Jessop, Monolithically Integrated Photodetectors for Optical Signal Monitoring in Silicon Waveguides, *Proceedings of Photonics West*, SPIE, San Jose, USA (2006).
- [37] P. G. Coleman, C. P. Burrows and A. P. Knights, Simple expression for vacancy concentrations at half ion range following MeV ion implantation of silicon, *Appl. Phys. Lett.*, **80**, 2002, 947.
- [38] A. Knights, A. House, R. MacNaughton and F. Hopper, Optical poer monitoring function compatible with single chip integration on silicon-on-insulator, *Proceedings of the Optical Fiber Communications Conference* 2003 (OFC2003), 705.

- [39] J. D. B. Bradley, P. E. Jessop and A. P. Knights, Silicon-waveguide integrated optical power monitor with enhanced sensitivity at 1550 nm, *Appl. Phys. Lett.*, **86**, 2005, 241103.
- [40] Y. Liu, C. W. Chow, W. Y. Cheung and H. K. Tsang, In-line channel power monitor based on helium ion implantation in silicon-on-insulator waveguides, *IEEE Photonics Technology Letters*, **18**, 2006, 1882.
- [41] P. J. Foster, J. K. Doylend, P. Mascher, A. P. Knights and P. G. Coleman, Optical attenuation in defect engineered silicon rib waveguides, *Jnl. Appl. Phys.*, **99**, 2006, 073101.
- [42] M. W. Geis, S. J. Spector, M. E. Grein, R. T. Schulein, J. U. Yoon, D. M. Lennon, S. Deneault, F. Gan, F. X. Kaertner and T. M. Lyszczarz, CMOS-compatible, all-Si, high-speed, waveguide photodiodes with responsivity in near-infrared communication bands, *IEEE Photonics Technology Letters*, **19**, 2007, 192.

# 7

# Passive Silicon Photonic Devices

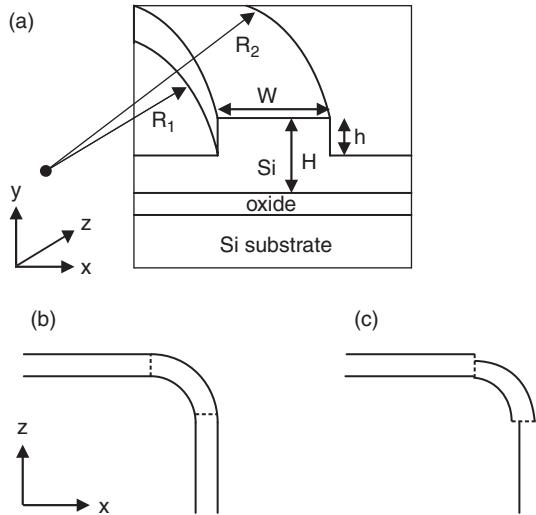
Ansheng Liu, Nahum Izhaky and Ling Liao

Complicated photonic integrated circuits are usually constructed from various fundamental photonic building blocks or components. In this chapter, we describe some selected devices such as waveguide bends, directional couplers, multi-mode interference couplers, Y-junctions, Mach–Zehnder interferometers, Bragg gratings, Fabry–Perot resonators, and ring resonators. The device operation principle and silicon waveguide-based applications are discussed.

## 7.1 The Waveguide Bend

In addition to straight waveguides, the waveguide bend is the most fundamental element in integrated optics. It is typically used for propagation redirection of the optical mode, which is an indispensable functionality in many optical circuits. Figure 7.1(a) shows schematically a bent waveguide in silicon-on-insulator (SOI) with a rib width of  $W$ , a rib height of  $H$ , and an etch depth of  $h$ . The transverse optical mode is confined in the  $xy$ -plane and propagates in the  $xz$ -plane. As any complicated bend structure can be constructed from simple circular bending sections, only the circular waveguide bend is discussed here. The inner bend radius of the waveguide is  $R_1$  and the outer radius  $R_2$  (the waveguide width is therefore  $W = R_2 - R_1$ ). For practical photonic circuits, the bent waveguide is usually combined with straight waveguides to achieve the device functionalities. As an example, Figure 7.1(b) and (c) show a top view of a 90° bend waveguide section placed in between two straight waveguides with (Figure 7.1c and without Figure 7.1b) a lateral offset.

Physically, bending of a waveguide leads to optical loss as the mode propagates around the bend because of radiation loss from its modal field in the cladding. This radiation loss is inherent to the waveguide bend, which is different from the scattering loss of a straight waveguide due to the waveguide roughness. In addition, there is modal transition loss between the bent waveguide and straight waveguide due to the modal profile mismatch. The magnitude of the bend loss depends on the bend radius and the waveguide confinement in the bend.



**Figure 7.1** (a) Schematic of a waveguide bend in SOI; (b) top view of a 90° bend without lateral offset; (c) top view of a 90° bend with lateral offset

direction. For SOI waveguide, the waveguide bend usually occurs in the plane parallel to the silicon-buried oxide interface ( $xz$ -plane in Figure 7.1) due to the device fabrication; therefore, the waveguide bend has most impact on the TE mode.

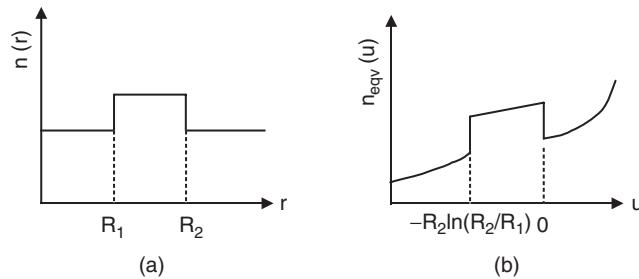
The first theoretical work on the waveguide bend was published by Marcatili in 1969 [1]. Since then, various models and methods [2–9] have been proposed and developed, including the conformal transformation method [2], the method of lines [3], the finite-difference method [4], the mode-matching method [8], and so on. Modeling of a bent waveguide has also been implemented in commercial software packages such as Beamprop [10] and Fimmwave [11]. Because the conformal transformation theory is physically intuitive, it will be discussed here to illustrate the waveguide bend effect. In the conformal transformation method [2], a curved waveguide is considered as a straight waveguide having an equivalent waveguide width of

$$W_{\text{equ}} = R_2 \ln \frac{R_2}{R_1} \quad (7.1)$$

with an equivalent index profile given by

$$n_{\text{equ}}(u) = n(u) \exp \left( \frac{u}{R_2} \right) \quad (7.2)$$

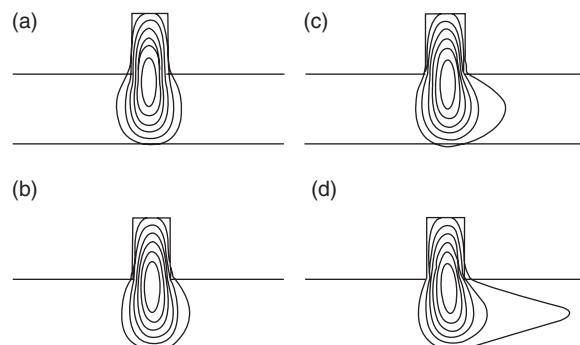
where  $n(u)$  is the refractive index profile without bending. Under appropriate boundary conditions such as adding a perfectly matched layer (PML) [9], the radiation loss of the bent waveguide could be calculated by solving the complex propagation constant of the mode in the equivalent waveguide. The transition loss between the straight and bent waveguides can be calculated by the modal overlap. Rigorously speaking, the conformal transformation method is accurate only for a curved slab waveguide (2D structure). However, it could still be applied



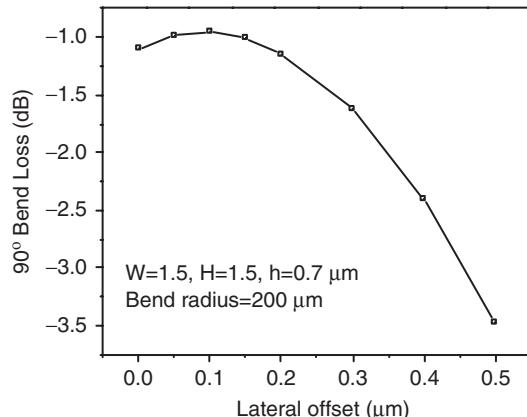
**Figure 7.2** (a) Effective index profile of a silicon rib waveguide; (b) the equivalent index profile of a bent waveguide in the conformal transformation theory

to a 3D bent waveguide, provided that the index profile  $n(u)$  is considered to be an effective lateral index profile obtained, for instance, by the effective-index method [12]. Figure 7.2 shows schematically the index distribution and its transformation of a vertically etched silicon rib waveguide that has a step-like effective index profile. It is clear from Figure 7.2(b) that all of the modes for the equivalent waveguide will radiate into the cladding region where the refractive index exceeds the maximum refractive index in the core. Namely, all the modes of a bent waveguide are leaky.

To give a quantitative analysis of a silicon waveguide bend, we take as an example a silicon rib waveguide having  $W = 1.5$ ,  $H = 1.5$ , and  $h = 0.7 \mu\text{m}$ . We calculate the TE mode of the waveguide for various bend radii using the commercial software Fimmwave [11]. Figure 7.3 shows the modeled  $E_x$  electric field distributions at a wavelength of  $1.55 \mu\text{m}$ . We see from Figure 7.3 that the mode of the waveguide bend tends to shift to the outer edge of the bend—the smaller the bend radius, the larger the modal shift. We also note that the mode distortion in a rib waveguide occurs mainly in the slab region. Because the bent waveguide mode is shifted toward the outer edge of the bend, one could reduce the mode mismatch induced transition loss by introducing a lateral shift between straight waveguide and bent waveguide [5], as



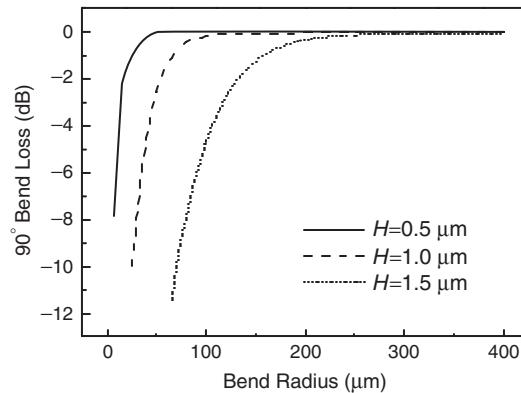
**Figure 7.3** Contour plots of the  $E_x$  component of the electric field of TE mode for a silicon waveguide having  $W = 1.5$ ,  $H = 1.5$ , and  $h = 0.7 \mu\text{m}$  for different bend radii: (a) infinite; (b) 400; (c) 200; (d) 160  $\mu\text{m}$



**Figure 7.4**  $90^\circ$  bend loss of a silicon waveguide as a function of the lateral offset. The waveguide dimensions are  $W = 1.5$ ,  $H = 1.5$ , and  $h = 0.7\text{ }\mu\text{m}$ . The bend radius is 200 μm

shown in Figure 7.1(c). Figure 7.4 shows the modeled  $90^\circ$  bend loss (both radiation loss and transition loss) as a function of the lateral offset for a silicon waveguide having the same dimensions as in Figure 7.3 with a bend radius of 200 μm. We see that at the optimal lateral offset of  $\sim 0.1\text{ }\mu\text{m}$  the total bend loss is reduced by  $\sim 0.14\text{ dB}$ . We note that, although the lateral offset can somewhat reduce the transition loss, the dominating bend loss in this case is the radiation loss.

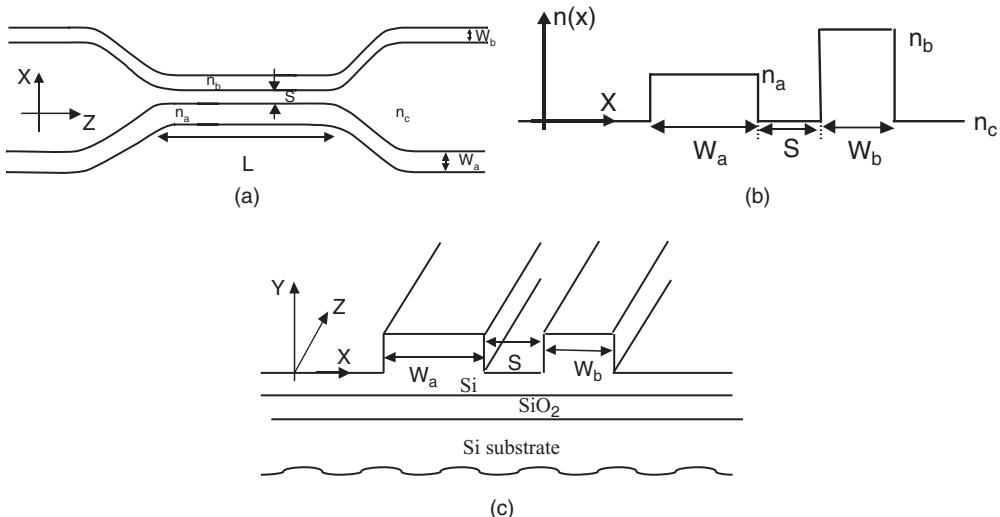
The minimum bend radius that results in negligible bend-related loss, including radiation loss and straight-to-bend waveguide transition loss, is an important design parameter in integrated optics. The size of the waveguide bend ultimately determines the maximum density with which photonic circuits can be integrated on a single chip. It is well known that the silicon strip (channel) waveguide has a very small bend radius (a few micrometers) because of the high refractive index contrast and waveguide geometry. However, the single mode operation requires a subwavelength waveguide size, which leads to a relatively high transmission loss ( $\sim 3\text{ dB/cm}$ ) due to fabrication induced waveguide imperfections [13]. In contrast, silicon rib waveguide exhibits single-mode behavior for a large range of waveguide sizes. The waveguide transmission loss can be much smaller. Because the radiation loss depends on the lateral confinement strength of the waveguide, for silicon rib waveguide, the minimum bend radius is strongly dependent on the waveguide geometry. In Figure 7.5 we show the modeled loss of a  $90^\circ$  bend waveguide section in between two straight waveguides as a function of bend radii for three different waveguide rib heights, namely  $H = 0.5, 1.0$ , and  $1.5\text{ }\mu\text{m}$ . The corresponding waveguide rib widths and etch depths are  $W = H$  and  $h = H/2$ . These waveguide dimensions are so chosen so the resulting waveguides are single-mode devices at a wavelength of  $1.55\text{ }\mu\text{m}$ . As expected, the bend radius is smaller for smaller waveguide cross-sections. The bend radius can be as small as  $\sim 50\text{ }\mu\text{m}$  for a rib waveguide with rib height of  $0.5\text{ }\mu\text{m}$ . This makes the small-cross-section rib waveguide very attractive for both active and passive devices. We note that one can also minimize the bend loss by using a deeper rib etching on the outer edge of the bend. For example, low loss of an asymmetric rib waveguide bend has been reported [14].



**Figure 7.5**  $90^\circ$  bend loss of a silicon waveguide as a bend radius for various waveguide dimensions

## 7.2 The Directional Coupler

The directional coupler (DC) is one of the fundamental and extensively used elements in integrated optics and optical communications. In its most basic configuration, a DC consists of two single-mode waveguides (WGs) in close proximity with input/output waveguide fanout, as illustrated in Figure 7.6(a). Its main function is to exchange optical power between the two adjacent WGs due to the modal interaction. A general refractive index profile of the two coupled WGs and its realization using SOI rib waveguides are shown in Figure 7.6(b) and (c).



**Figure 7.6** (a) Top view of a directional coupler; (b) refractive index distribution; (c) realization in SOI rib waveguides

When the WG separation  $s$  is small enough, the evanescent parts of the guided modes overlap and coupling between the two WGs occurs. In the case where the coupling is weak, the proximity of one WG to the other can be considered as a perturbation. Thus a combination of the evanescent tail of the mode guided in one WG and the neighboring WG induces an electric perturbing polarization generating a wave in the neighboring WG which in turn may couple back to the first WG. This coupling mechanism is ubiquitous in nature and can be found in many analogous systems, including coupled mechanical springs, coupled pendulums, coupled LC circuits, coupled molecules, predator-prey populations, coupled quantum potential wells, and many more. In fact, the mathematical equations, see Eqation (7.4) below, modeling these numerous phenomena are very similar. One expects stronger coupling for weaker mode confinement and better overlap; therefore, narrower WGs, longer wavelengths, or smaller WG separation (gap) provides stronger coupling and hence more compact elements. In addition, a pair of symmetric WGs is necessary for full power transfer, whereas asymmetric WGs (e.g.  $W_a \neq W_b$ , or  $n_a \neq n_b$ ) cannot achieve full coupling. Overall, as in other coupled systems, the lightwave is codirectionally coupled back and forth between the two WGs in a periodic manner in accordance with appropriate phase conditions between the coupled modes.

There are various ways to model and simulate the behavior of DCs, including, for example, the beam propagation method (BPM), finite difference time domain (FDTD), and coupled mode theory (CMT). While the BPM and FDTD methods usually give accurate results and are particularly useful for real device design, the CMT provides analytical solutions, hence strong physical intuition. The CMT for optical WGs has a long history of development starting in the early 1970s [15–18]. These models have been successfully applied to the analysis of guided wave optical devices; however, their results are accurate only in special cases of near identical WGs and weak coupling. Hardy *et al.* [19, 20] have improved the CMT for the case of forward coupling. Their theory, enabling the analysis of strong interactions and asymmetric couplers, is derived directly from Maxwell's equations and employs the mode expansion conjecture [21]. Since then several equivalent formulations have been developed [22, 23]. In the following, we use the CMT to describe the operation of the DC.

Based on the perturbation theory of coupled modes [18, 21, 24], the coupling between two guided modes a and b (while neglecting other guided and radiation modes) can be described by an approximate solution to the wave (or Maxwell's) equation written as

$$E(r, t) = a(z)\varepsilon_a(x, y)\exp[i(\omega t - \beta_a z)] + b(z)\varepsilon_b(x, y)\exp[i(\omega t - \beta_b z)] \quad (7.3)$$

where  $\varepsilon_a(x, y)$  and  $\varepsilon_b(x, y)$  are the transverse spatial electric field distributions of particular unperturbed guided modes of WGs a and b, respectively, and their propagation constants are denoted by  $\beta_a$  and  $\beta_b$  accordingly. The amplitudes  $a(z)$  and  $b(z)$  vary along the DC length due to the evanescent coupling between the two modes. The governing set of coupled mode equations derived from perturbation theory is

$$\begin{aligned} \frac{dR(z)}{dz} &= i\delta R(z) - i\kappa S(z) \\ \frac{dS(z)}{dz} &= -i\delta S(z) - i\kappa R(z) \end{aligned} \quad (7.4)$$

where  $a(z) \equiv R(z) \exp(i\delta z)$ ,  $b(z) \equiv S(z) \exp(-i\delta z)$ ,  $2\delta \equiv \beta_a - \beta_b$ , and  $\kappa$  is the coupling coefficient which determines the coupling efficiency and is usually defined as

$$\kappa \equiv \frac{\omega}{4} \langle a | \Delta \varepsilon | b \rangle = \frac{\omega \varepsilon_0}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varepsilon_a^*(x, y) \Delta n^2(x, y) \varepsilon_b(x, y) dx dy \quad (7.5)$$

where  $\Delta n^2(x, y)$  represents the perturbation in the dielectric coefficient ( $\Delta \varepsilon = \varepsilon_0 \Delta n^2$ ). Equation (7.5) indicates that  $\kappa$  is dependent on  $\omega$ ,  $\Delta n^2$ ,  $\varepsilon_a$ , and  $\varepsilon_b$ . The unit of  $\beta_a$ ,  $\beta_b$ ,  $\delta$ , and  $\kappa$  is 1/length.

For an arbitrary inputs (at  $z = 0$ )  $R_0$  and  $S_0$  the solution to Equation (7.4) is

$$\begin{bmatrix} R(z) \\ S(z) \end{bmatrix} = \begin{pmatrix} A & -iB \\ -iB^* & A^* \end{pmatrix} \begin{pmatrix} R_0 \\ S_0 \end{pmatrix} \quad (7.6a)$$

where  $A$  and  $B$  are defined as

$$\begin{cases} A \equiv \cos(z\sqrt{\kappa^2 + \delta^2}) + i \frac{\delta}{\sqrt{\kappa^2 + \delta^2}} \sin(z\sqrt{\kappa^2 + \delta^2}) \\ B \equiv \frac{\kappa}{\sqrt{\kappa^2 + \delta^2}} \sin(z\sqrt{\kappa^2 + \delta^2}) \end{cases} \quad (7.6b)$$

For power relations with only one input  $S_0 \neq 0$  and  $R_0 = 0$ , we obtain from Equation (7.6)

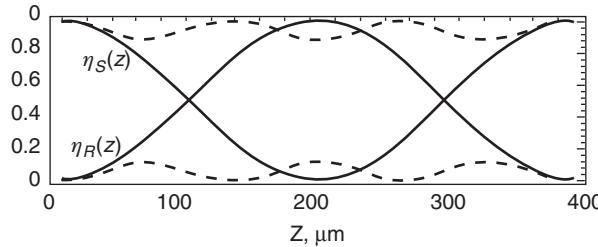
$$\begin{cases} \eta_R(z) \equiv \frac{P_R(z)}{P_0} = \frac{\kappa^2}{\kappa^2 + \delta^2} \sin^2(z\sqrt{\kappa^2 + \delta^2}) \\ \eta_S(z) \equiv P_S(z)/P_0 = 1 - \eta_R(z) \end{cases} \quad (7.7)$$

where  $P_R$  and  $P_S$  are the powers related to  $R$  and  $S$ , respectively, and  $P_0$  is the injected power (proportional to  $|S_0|^2$ ). From Equation (7.7) one can see that maximum power coupling from one WG to the other is obtained at  $z = L_c \equiv (2m+1)\pi/(2\sqrt{\kappa^2 + \delta^2})$ , where  $m = 0, 1, 2, \dots$  and the maximum power fraction of coupling is  $\kappa^2/(\kappa^2 + \delta^2)$ . As expected, for stronger asymmetry (larger  $\delta$ ) one obtains less coupling. This is similar to two coupled pendulums with masses. The more the mass difference is, the less complete the coupling is. A special case is the 3 dB coupler for which the power is split equally between the two output WGs of the DC. It is obtained at  $z = L_c/2$ .

For the case of symmetric DC ( $\delta = 0$ , i.e.  $\beta \equiv \beta_a = \beta_b$ ), Equation (7.6) is reduced to the following transfer matrix form

$$\begin{bmatrix} R(z) \\ S(z) \end{bmatrix} = \mathbf{M} \begin{pmatrix} R_0 \\ S_0 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} \cos(\kappa z) & -i \sin(\kappa z) \\ -i \sin(\kappa z) & \cos(\kappa z) \end{pmatrix} \quad (7.8)$$

Figure 7.7 depicts the behavior of the DC in relative power as a function of length for symmetric (full coupling, solid line) and asymmetric (partial coupling, dashed line) cases. As can be expected from lossless modeling, the  $\mathbf{M}$  matrix in Equation (7.8) is unitary. In addition, it is symmetric, as required from reciprocity—the coupler behaves the same when lightwave is incident from either directions—input or output. Also as indicated by the factor of  $-i = \exp(-i\pi/2)$ , the DC introduces a relative phase shift of  $\pi/2$  between the two output ports. For the symmetrical case (and  $m = 0$ ),  $L_c = \pi/(2\kappa)$ .



**Figure 7.7** DC relative power as function of length for symmetric DC (solid line) and asymmetric DC (dashed line). CMT modeling of Si slab and  $\text{SiO}_2$  cladding for DC with  $W_a = 0.25 \mu\text{m}$ ,  $W_b = 0.25 \mu\text{m}$  (symmetric) or  $0.248 \mu\text{m}$  (asymmetric), and  $s = 0.5 \mu\text{m}$  at  $\lambda = 1.55 \mu\text{m}$

The DC can also be analyzed in terms of supermodes or normal modes of coupled waveguides. In this technique, the whole structure of two coupled WGs is regarded as one waveguide and its eigenmodes (called supermodes or normal modes) are usually solved numerically by modal solvers. For two weakly coupled identical waveguides, there are two orthogonal eigenmodes, i.e. symmetrical (fundamental) mode  $\varepsilon_s$  and the antisymmetrical (first) mode  $\varepsilon_a$ . The two eigenmodes have different propagation constants  $\beta_s$  and  $\beta_a$ . Note that the supermode of the coupled waveguides is different from the mode of the individual waveguide even when the coupling is weak. Hence the different propagation velocities yield a phase difference of  $\Delta\phi = \ell\Delta\beta = \ell(\beta_s - \beta_a) = (2\pi/\lambda)\ell(n_s - n_a)$ , where  $n_s$  and  $n_a$  are the relevant mode effective indices,  $\ell$  is the length of the coupler, and  $\lambda$  is the vacuum wavelength. At  $\Delta\phi = \pi$  one gets complete coupling, namely  $L_c \equiv \pi/(\beta_s - \beta_a) = \lambda/[2(n_s - n_a)]$ . Comparing the CMT and the supermode analysis, we obtain  $\beta_s - \beta_a = 2\kappa$ .

Until now we have assumed a constant coupling coefficient in our analysis of the DC. However the DC usually includes input/output waveguide fanout for practical implementation, see Figure 7.6(a). Thus, one should take into account a varying  $\kappa$ , i.e.  $\kappa = \kappa(z)$ . For numerical modeling, using for example the BPM method, such an effect is automatically included. When the CMT or supermode analysis is employed, one could consider the contribution of the bends in the waveguide fanout region as an effective additional length ( $\Delta\ell$ ) of the same  $\kappa$  of the straight part ( $\ell$ ). The additional length depends on the waveguide bend design. For a given waveguide fanout design, one can determine  $L_c$  and  $\Delta\ell$  empirically from the power coupling ratio dependence on the straight interaction length  $\ell$  by using a least-squares approximation to the experimental results.

In real devices, the DC always has an additional loss in reference to a straight WG with the same cross-section and length. This excess loss comes from the bends and the interacting WGs and gap inhomogeneity that can be considered phenomenologically and equally distributed along the coupler as a factor of  $\exp(-\alpha z)$  where  $\alpha$  (in  $1/\text{length}$ ) is the power loss coefficient, and in  $\text{dB}/\text{length}$  units it is  $4.343\alpha$ . This parameter simulates power decay along the DC length. In order to minimize DC loss one should design and fabricate its bends properly, using bends of large enough radius to reduce bend radiation loss, utilizing offsets if necessary to reduce bend transition loss, and optimizing processing steps to obtain smooth interaction region with appropriate cross-section.

Directional couplers are useful in numerous optical components. For example, optical switches in a Mach-Zehnder interferometer (MZI) configuration and ring-resonator-based optical filters utilize DCs. The wavelength dependence of the DC allows for spectral MUX/DEMUX of different wavelength signals. DCs are used also in grating-assisted devices for efficient WG couplers [25]. In fact, one can find optical DCs since the advent of integrated optics [26], and they have been demonstrated and tested in every available material system—LiNbO<sub>3</sub>, silica-on-Si, polymers, glass, GaAs, InP, and in the last decade also on SOI. In general DCs using silica buried (channel) WGs provide the lowest loss (0.05–0.1 dB/coupler), but their length is relatively large (on the order of millimeters) due to the low core-cladding index contrast ( $\sim 10^{-2}$ , —a relative index contrast of  $\sim 0.75\%$ ), which requires bend radii of more than 5 mm for low-loss behavior. In addition, their polarization-dependent loss (PDL) is very low and negligible.

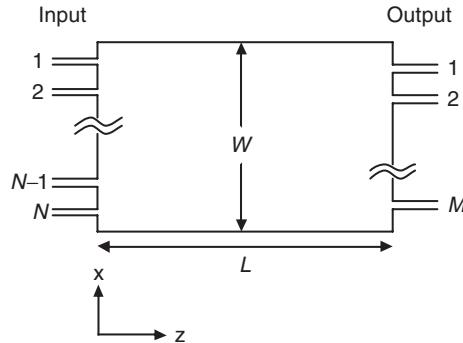
Silicon DCs are formed in two different configurations—one is the rib WG (partially etched epilayer) and the other is the strip (or Si wire) WG (fully etched epilayer). The rib WG can be used with bend radii of several hundreds of micrometers, whereas the strip WG has better confinement and can be bent to several micrometer radii. Recent lithography capabilities like DUV or e-beam allow also a sub-micrometer separation (gap) (e.g. 0.1, 0.2  $\mu\text{m}$ ). Therefore DCs in Si photonics can be very short (tens of micrometers). However, very small WG dimensions of single-mode strip are very sensitive to any imperfection in fabrication, such as side wall roughness, and process tolerances are tighter. Hence, higher DC loss is expected, but the shorter length compensates for that, and DC loss is in the range of several tenths of dB.

The first reported fabricated DC in SOI technology was published in 1995 by Trinn *et al.* [27]. They have demonstrated a 3 dB DC based on a rib WG designed by a simple BPM technique, and obtained excess loss of 1.9 dB for several hundreds of micrometers of the straight interaction length, with a gap of  $\sim 2.5 \mu\text{m}$ . A significant reduction of coupling length ( $\sim 50\%$ ) was demonstrated both numerically and experimentally by Cau *et al.* [28], through asymmetric rib configuration of etching the outer side of the epilayer of the rib WGs in the coupling region, hence increasing the mode overlap and so shortening the required 3 dB length. Lately, a compact SOI polarization splitter based on DC was suggested by Kyat *et al.* [29]. It is based on geometrically induced birefringence ( $B = |n_e - n_m|$ ) by selecting small width  $W$  and height  $H$  with  $(H-h)/H = 0.6$ , yielding  $B \approx 0.018$ . Then by selecting a coupling length of  $L_c(TM)/L_c(TE) = 2$  for a gap of  $g = 0.7 \mu\text{m}$  they showed a cross state for the TM polarization and bar state for the TE polarization. Yet the ER (between the two output ports) was  $\sim 18$  dB for TE and only  $\sim 8$  dB for TM.

Si wire (strip) DC was demonstrated by Yamada *et al.* [30]. They have fabricated strips of  $0.3 \times 0.3 \mu\text{m}^2$  and gaps of 0.3 and 0.2  $\mu\text{m}$  with e-beam lithography; by that they allowed bends of 10  $\mu\text{m}$  radius and straight coupling lengths of  $L_c = 10$  and 5  $\mu\text{m}$ . Design of a long coupler (800  $\mu\text{m}$ ) provided a wavelength-dependent transmittance of 2.5 nm spacing over the C + L bands, with ER of  $\sim 20$  dB.

### 7.3 The Multimode Interference Coupler

The multimode interference (MMI) coupler [31] is a photonic device that employs self-imaging [32, 33] in multimode waveguides and has recently become a key component for photonic integrated circuits. The central structure of a MMI coupler is the multimode waveguide that



**Figure 7.8** Schematic of a  $N \times M$  MMI coupler with  $N$  input waveguides and  $M$  output waveguides. The multimode waveguide has a width of  $W$  and a length of  $L$  and supports a large number of modes

can support a large number of guided modes (typically more than three modes). Figure 7.8 shows schematically a  $N \times M$  MMI coupler with  $N$  input waveguides and  $M$  output waveguides connected to the multimode waveguide having a width of  $W$  and a length of  $L$ . The MMI coupler can be used for beam coupling, splitting, and combining.

Following the work of Soldano and Pennings [31], we use the modal propagation method to analyze the operation of a MMI device. Because the multimode waveguide usually has a much larger lateral dimension (width) compared with the waveguide height, we adopt a 2D approximation to the analysis of the propagation constants  $\beta_v$  ( $v = 0, 1, 2, \dots$ ) of the modes. We only consider the mode propagation in the  $xz$ -plane (see Figure 7.8) and assume all the modes have the same behavior due to the vertical confinement. Then, the propagation constant spacing between fundamental mode and higher-order modes is approximately given by [34]

$$\beta_0 - \beta_v = \frac{v(v + 2)\pi}{3L_\pi} \quad (7.9)$$

where  $L_\pi$  is the beat length between the fundamental and the first excited modes of the waveguide. In a 2D approximation, the beat length is related to the effective MMI width  $W_{\text{eff}}$  by [31]

$$L_\pi = \frac{4n_r W_{\text{eff}}^2}{3\lambda} \quad (7.10)$$

where  $n_r$  is the effective refractive index of the core, and  $W_{\text{eff}}$  is given by [31]

$$W_{\text{eff}} = W + \frac{\lambda}{\pi} \left( \frac{n_c}{n_r} \right)^{2\sigma} (n_r^2 - n_c^2)^{-1/2} \quad (7.11)$$

with  $\sigma = 0$  for TE and  $\sigma = 1$  for TM polarizations, respectively, and  $n_c$  is the effective refractive index of the cladding. Note that the effective MMI width is wider than the physical width  $W$  because of the modal penetration into the cladding.

We now discuss the field distribution in the multimode waveguide. Because we are interested in the modal propagation in the  $xz$ -plane, we will ignore the  $y$ -dependence of the field in the

following. Assume an input field profile  $\Psi(x, 0)$  is incident into the MMI at  $z = 0$ . The input field can be expressed in terms of a linear superposition of a set of waveguide modes  $\varphi_v(x)$  as

$$\Psi(x, 0) = \sum_v c_v \varphi_v(x) \quad (7.12)$$

where  $c_v$  is the field excitation coefficient. Rigorously speaking, the summation in Equation (7.12) should run over all the modes, including both guided and radiative modes. In practice, however, it is enough to consider only the guided modes. Using Equation (7.9), it can be shown that the field distribution at  $z = L$  is given by

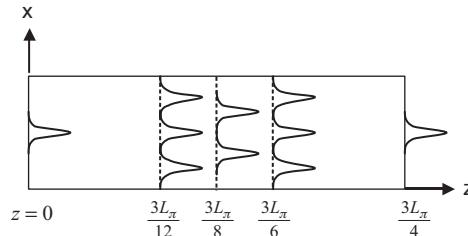
$$\Psi(x, L) = \sum_v c_v \varphi_v(x) \exp\left[i \frac{v(v+2)\pi}{3L_\pi} L\right] \quad (7.13)$$

By comparing Equations (7.12) and (7.13), it is clear that the image of the input field is determined by the phase-related factor  $\exp[i v(v+2)\pi L/(3L_\pi)]$ . In the case where the phase factor is 1 or  $(-1)^v$ , we obtain a single image of the input field at  $z = L$ . In general, the MMI field profile depends on the input field profile, the MMI device width and length as well as the mode confinement.

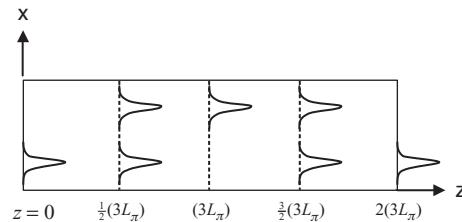
Two commonly used MMI couplers in photonic circuits are  $1 \times N$  and  $2N \times 2N$  couplers.  $1 \times 2$  and  $2 \times 2$  couplers, 3 dB splitters are typical examples. For  $1 \times N$  MMI couplers, the one input waveguide is placed at the middle of the multimode waveguide. Symmetric guided modes  $v = 0, 2, 4, \dots$  are excited only due to the waveguide symmetry and a self-imaging distance is  $3L_\pi/4$ . The  $M$ -fold image is found at a length of  $3L_\pi/4M$ . Figure 7.9 schematically shows the input field, self-image, two-fold, and three-fold image. For  $2N \times 2N$  couplers, all guided modes are excited by the input field that is placed off the waveguide center. For each input field, the self-imaging (at the same location in the lateral direction) or mirrored self-imaging (at the opposite location in the lateral direction) are found at

$$L = p(3L_\pi) \quad \text{with } p = 0, 1, 2, \dots \quad (7.14)$$

When  $p$  is even integer ( $0, 2, 4, \dots$ ), we obtain the self-imaging. When  $p$  is odd integer ( $1, 3, 5, \dots$ ), we obtain the mirrored self-imaging. The  $M$ -fold images can be found at  $L = p(3L_\pi)/M$ , where  $p \geq 0$  and  $M \geq 1$  are integers with no common divisor. Figure 7.10 shows the input field, self-image, mirrored self-image, and two-fold images of MMIs with off-center excitation.



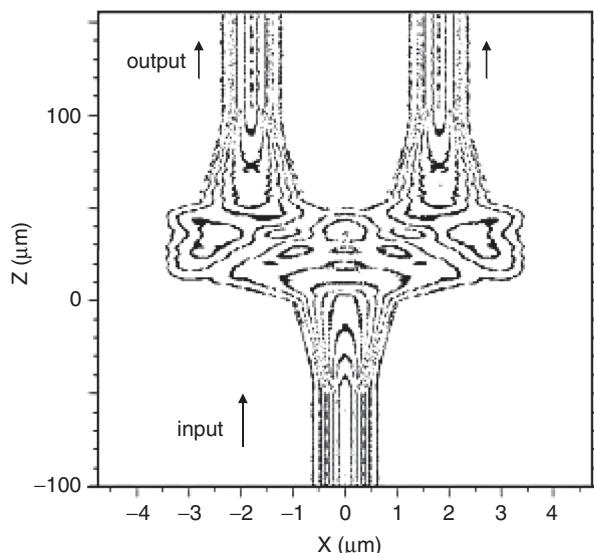
**Figure 7.9** Symmetric excitation of a multimode waveguide by the input field ( $z = 0$ ) and a self-image ( $3L_\pi/4$ ), two-fold ( $3L_\pi/8$ ), and three-fold images ( $3L_\pi/12, 3L_\pi/6$ )



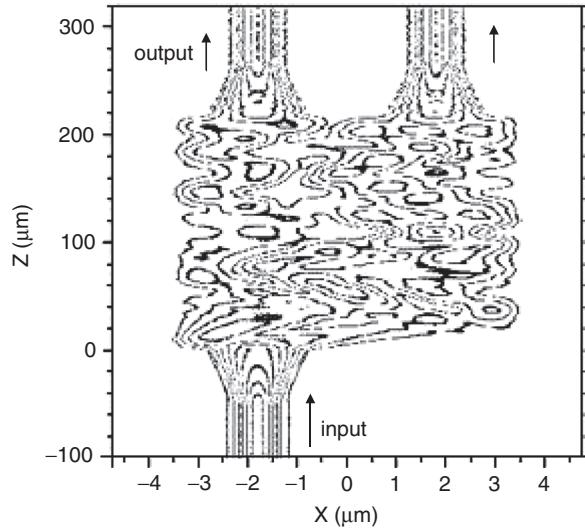
**Figure 7.10** Off center excitation of a multimode waveguide by the input field ( $z = 0$ ) and a self-image ( $2 \times 3L_\pi$ ), a mirrored self-image ( $3L_\pi$ ), two-fold images ( $3L_\pi/2, 3 \times 3L_\pi/2$ )

The separation between input waveguide and output waveguides in a MMI coupler are usually required to be small in order to minimize the MMI device length. To avoid the modal interaction between the input waveguides or output waveguides, waveguide bends are used for practical implementation. Because the MMI modal behavior also depends on the input mode profile, a tapered input/output waveguide is also used to optimize the device performance, for example, to reduce the MMI transmission loss.

The above modal propagation analysis provides a basic understanding of the MMI device principle; however, it is essentially a 2D analysis. For 3D structures, we need 3D vectorial numerical simulations to analyze and design the MMI coupler. Fortunately, a number of commercial softwares based on BPM [11] and local mode expansion methods [12] are available for the modeling. Figures 7.11 and 7.12 show, respectively, the modeled (by BPM) modal



**Figure 7.11**  $xz$ -plane contour plot of the optical field propagating in a  $1 \times 2$  3 dB MMI splitter with tapered input and output waveguides at the waveguide center in the  $y$ -direction. The MMI width is 7  $\mu\text{m}$ , MMI length is 54.5  $\mu\text{m}$ . The waveguide height is 0.5  $\mu\text{m}$ , and the etch depth is 0.25  $\mu\text{m}$



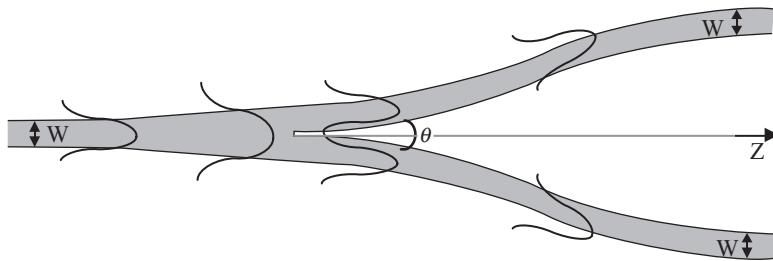
**Figure 7.12**  $xz$ -plane contour plot of the optical field propagating in a  $2 \times 2$  3 dB MMI splitter with tapered input and output waveguides at the waveguide center in the  $y$ -direction. The MMI width is 7  $\mu\text{m}$ , MMI length is 219.4  $\mu\text{m}$ . The waveguide height is 0.5  $\mu\text{m}$ , and the etch depth is 0.25  $\mu\text{m}$

propagations of  $1 \times 2$  and  $2 \times 2$  3 dB silicon waveguide splitters at 1.55  $\mu\text{m}$ . In both calculations, the silicon waveguide height is 0.5  $\mu\text{m}$  and the etching depth is 0.25  $\mu\text{m}$ . Tapered input/output waveguides are used (the waveguide fanout is not included in the simulation). For the  $1 \times 2$  MMI coupler, the MMI width is 7  $\mu\text{m}$  and the length is  $L = 54.5 \mu\text{m}$ . For the  $2 \times 2$  MMI coupler, the MMI width is 7  $\mu\text{m}$  and the length is 219.4  $\mu\text{m}$ . TE polarization was used in both Figures 7.11 and 7.12.

We note that most of the MMI couplers are based on a straight multimode waveguide. In this case, the MMI device length is essentially determined by the MMI device width, which is in turn determined by input/output waveguide width and number. While the input/output MMI device width can be a fixed design parameter, the MMI device width in the center part is flexible. A so-called tapered MMI device has been proposed and analyzed [35–37]; significant reduction in the MMI device length was achieved [36].

Silicon-waveguide-based MMI couplers can be used in many optical components such as optical modulators, optical switches, and ring resonators. For these applications, the MMI device works as a 50–50% splitter/combiner, just like a 3 dB directional coupler. The first SOI waveguide based  $1 \times 2$  and  $2 \times 2$  MMI couplers were reported by Fischer, Zinke, and Petermann in 1995 [7,38]. These devices were used to form  $1 \times 2$  and  $2 \times 2$  optical switches. Silicon waveguide  $1 \times 4$  and  $1 \times 8$  MMI couplers were also reported [39]. Good power splitting and low on-chip loss were obtained. The MMI coupler has also been used in an arrayed waveguide grating to broaden the spectral response [40] and in a variable optical attenuator [41].

Owing to the multimode interference nature, the MMI coupler has low polarization sensitivity, large wavelength bandwidth, and large fabrication tolerance [42]. Recently, a MMI-based polarization insensitive ring resonator was designed [43]. Also, because there is no small feature required in the design, the MMI coupler is insensitive to the photolithography



**Figure 7.13** Top view of an optical Y-junction

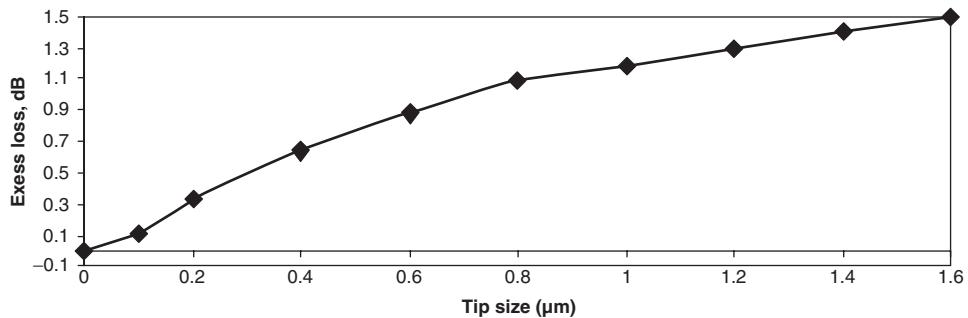
resolution. However, because of the high index contrast of the SOI system, the reflection in the MMI device might be a concern for certain applications. Such an effect was observed in a MMI-based ring laser [44]. However, the reflectivity from the interface between the MMI region and output waveguide was shown to be relatively small [45].

## 7.4 The Y-junction

The Y-junction is a simple three-port device in which one single-mode (SM) WG branches into two SM WGs, as shown in Figure 7.13. A tapered waveguide region bridges the single WG and the two WGs. In order to have small excess loss, the tapered waveguide must be able to support more than one guided mode, typically, both the symmetric and the antisymmetric eigenmodes.

The Y-junction can operate as either a splitter (propagating from the single WG to the two WGs) or a combiner (propagating from the two WGs to the single WG). When the junction is designed symmetrically (as in Figure 7.13), it performs as a 3 dB splitter. When light mode is injected through the one WG input, due to geometrical symmetry of the structure and the adiabatic design of the taper, only the symmetric mode is excited in the tapered region. Hence an equal power split occurs into the two branches with little excess loss. In practice, due to fabrication limitations in litho and etch resolutions, a blunt tip occurs below a certain distance between the two WGs and violates the adiabatic requirement, resulting in extra loss. The loss depends on the tip size and on the opening angle. The smaller the tip and the angle, the smaller is the excess loss. From a practical point of view, it is better to design the Y-junction with the minimal achievable tip size limited by litho resolution and etching. In silica WGs, an excess loss of about 0.2 dB per Y-junction was usually reported [46]. The higher the refractive index contrast, the higher the loss, yet a higher branching angle can be tolerated for the same excess loss of lower index contrast systems. Figure 7.14 shows the simulated (by BPM) excess loss in a SOI waveguide splitter as a function of the tip size. As the figure shows, a tip size beyond 0.3  $\mu\text{m}$  can degrade the splitter loss performance significantly ( $>0.5$  dB).

Y-junction splitters can also be designed to realize weighted splitting or even optical taps. For instance, one can modify the width of the two output WGs to differ from each other or to change their angle with respect to the  $z$ -axis to obtain broadband and polarization-independent asymmetric splitters.



**Figure 7.14** Simulated (by BPM) excess loss in SOI splitter as function of the tip size. Rib cross-section is  $W = 1.5 \mu\text{m}$ ,  $H = 1.5 \mu\text{m}$ ,  $h = 0.7 \mu\text{m}$  at  $\lambda = 1.55 \mu\text{m}$

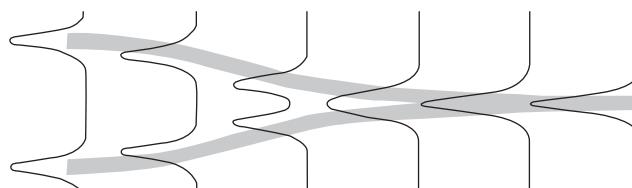
When the Y-junction operates as a combiner, its principles are less obvious. Based on its eigenmodes (symmetric and antisymmetric) and the reciprocity principle, if we inject one mode into each of the two WG ports with equal power and in phase with the same wavelength and polarization (see Figure 7.15), one would justifiably expect a coherent power combining with negligible loss; this is exactly the reciprocal case of the splitter operation described above.

On the other hand, in the case where the two injected modes are out of phase ( $\pi$  phase shift, Figure 7.16), only the antisymmetric mode is excited in the coupled waveguide region. Because that mode cannot be guided in the SM output WG, it radiates out of the WG and all the power is lost.

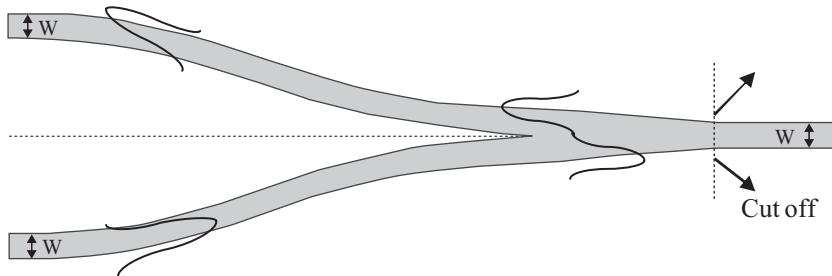
In the case of a single input from one of the two WGs (Figure 7.17), the mode in the coupled waveguide region can be considered as a superposition of a symmetrical mode and an antisymmetrical mode with equal amplitudes. Because only the symmetrical mode can be guided in the output SM WG port (Figures 7.15 and 7.16), half of the input power is therefore lost after passing the Y-junction. In this case, one obtains an inherent loss of 3 dB!

From this explanation it is clear that, if the one port side can contain multimodes, a lossless operation is expected in all the above cases. The description can be confirmed easily by the BPM technique.

Y-junctions can find numerous applications. Working as a 3 dB splitter/combiner, the Y-junction is frequently used to construct MZIs. Because the MZI requires very accurate 50–50% splitting with wavelength insensitivity and polarization independence, the Y-junction-based MZI typically has a robust performance. The Y-junction can also be used for constructing  $1 \times N$  splitter. A multiplication of concatenated 3 dB splitters (as a binary tree) of  $N$  levels



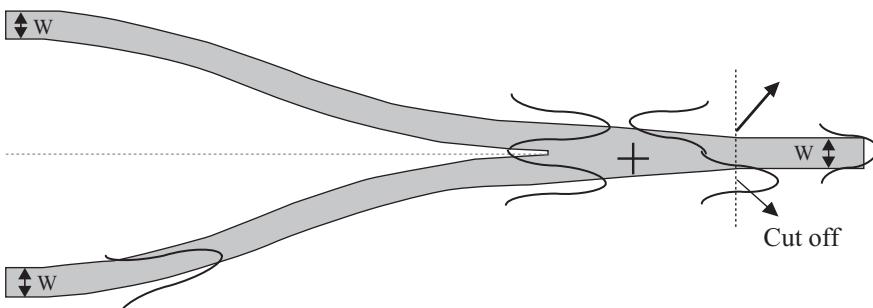
**Figure 7.15** Two modes are injected in-phase, excite the symmetrical mode and are fully combined



**Figure 7.16** Two out-of-phase modes are injected, excite the asymmetrical mode and are lost into radiation modes

can split a single input into  $2^N$  equal outputs. Lastly, we want to mention that the Y-junction is useful for digital optical switch (DOS). Optical switches employing an interferometric mechanism possess strong wavelength and polarization dependence, which requires specific and quite accurate voltages for each element due to unavoidable fluctuations in fabrication. The DOS switch [47–49], however, has a step-like nonperiodic switch response and is much less wavelength and polarization dependent. This digital (binary) characteristic can provide a unique and very attractive and tolerant element (wide window of operational parameters and process).

As a very fundamental building block, a silicon-waveguide-based Y-junction was designed and fabricated in the early stage of silicon photonics research. For example, Rickman and Reed and Tang *et al.* [50, 51] have designed and fabricated Y-junctions based on Si rib WGs, and obtained junction excess loss of 1.92 dB for TE and 2.235 dB for TM at  $\lambda = 1.15 \mu\text{m}$  and 1.82 dB for TE and 3.51 dB for TM at  $\lambda = 1.523 \mu\text{m}$ . The higher TM loss was related to lower confinement at this polarization. Dumon *et al.* [52] have designed and fabricated Y-junctions based on Si wire WGs. The reported excess loss is  $\sim 1.5$  dB with a slight unbalanced response of 0.13 dB due to small fabrication variation or mask digitalization error. They employed the Y-splitter within  $1 \times 1$  MZIs and demonstrated an ER of more than 25 dB. A Si  $1 \times 2$  DOS (based on the plasma dispersion effect) was demonstrated at  $\lambda = 1.3 \mu\text{m}$  by Liu *et al.*



**Figure 7.17** Single input mode is injected exciting both symmetrical and the asymmetrical modes. The asymmetrical is lost yielding 3 dB loss

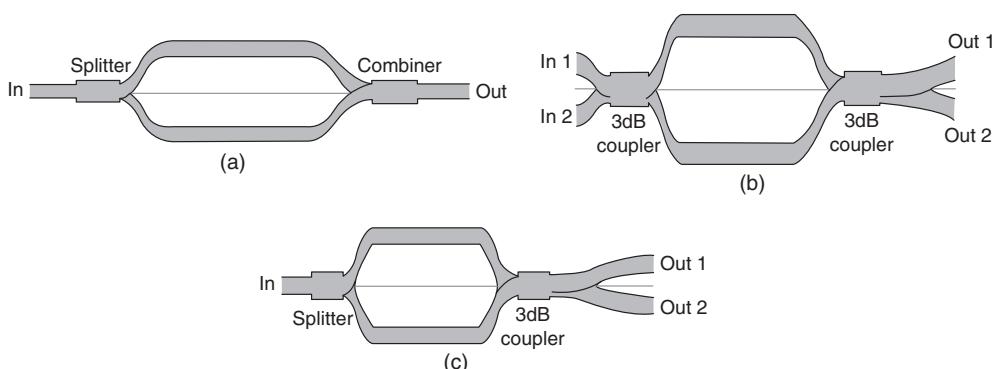
[53]. A Y-junction in a photonic crystal waveguide has also been demonstrated. A very short splitter can be formed via Si photonic crystals with low loss. For example, Borel *et al.* [54] demonstrated TE polarized 3 dB splitting via a Y-junction with 60° bend angle and an average excess loss of  $0.44 \pm 0.29$  dB for a 100 nm bandwidth on a SOI substrate. The splitter size was less than  $10 \times 15 \mu\text{m}^2$ .

## 7.5 The Mach–Zehnder Interferometer

The Mach–Zehnder interferometer is a most versatile and widely used device in which a light beam is split into two parts, and then each part undergoes a different physical path which may create a phase difference, and finally the two parts are combined to provide the end result where the phase difference is converted into a change in output light intensity. The MZI was invented by Mach and Zehnder about 115 years ago and since then has been extensively utilized in numerous forms and materials and for countless applications. In integrated optics there are mainly three basic MZI configurations, as schematically illustrated in Figure 7.18.

The  $1 \times 1$  MZI (Figure 7.18a) makes use of a  $1 \times 2$  splitter and a  $2 \times 1$  combiner usually based on Y-junctions or MMIs. The  $2 \times 2$  MZI (Figure 7.18b) consists of two  $2 \times 2$  3 dB couplers based on DCs or MMIs. The third type of MZI (Figure 7.18c) can be defined as  $1 \times 2$  or  $2 \times 1$  MZIs, and is a combination of the first two configurations.

The operation principle of the  $1 \times 1$  MZI is relatively simple. The input light is split into two equal parts; each propagates in its own WG arm. The two WG arms may have the same (symmetric) or different (asymmetric) optical path length (OPL) due to possible effective refractive index change or waveguide length difference. If light in the two arms acquires no phase difference between them, they arrive at the combiner as a symmetrical mode, i.e. they recombine in phase (constructive interference) and the lightwave fully emerges from the output SM WG – *on state*. When the two beams acquire a phase difference of  $\pi$ , only an antisymmetrical eigenmode is produced and is totally lost in the combiner output SM WG – *off state*. In the case where the two WG arms differ in their propagation losses, a non-equal arm loss ingredient would degrade the extinction ration (ER) defined as  $10 \log_{10}(P_{\text{on}}/P_{\text{off}})$ . When the two arms of the MZI have a waveguide length difference (unbalanced MZI)  $\Delta L$ , the



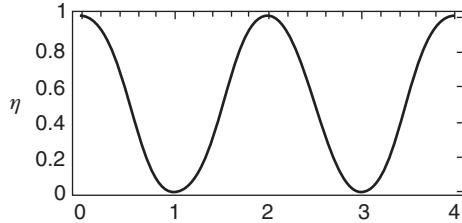
**Figure 7.18** Three main MZI configurations: (a)  $1 \times 1$  MZI; (b)  $2 \times 2$  MZI; (c)  $1 \times 2$  MZI

optical phase difference is  $\Delta\phi = 2\pi n_{\text{eff}} \Delta L / \lambda$ , where  $n_{\text{eff}}$  is the effective index. Thus, for the unbalanced MZI configuration, one expects a wavelength dependence of the output intensity.

In the case of  $2 \times 2$  MZI, after the first 3 dB coupler, light from one of the input ports is equally split into two beams, and the beam in the cross port lags behind the beam in the bar port by a phase of  $\pi/2$ . When the arms have no phase shift, the two equal power beams arrive at the second coupler with  $\pi/2$  phase difference. This is exactly the same conditions the light would ‘see’ in the middle of a full coupler (of length  $L_c$ ); thus, the second 3 dB coupler can be considered as a direct continuation of the first one and the two work together as a single coupler with full coupling length. All the light therefore propagates through the MZI out of its cross output (cross state). If the difference in OPL between the two arms is  $\lambda/2n_{\text{eff}}$ , i.e. a  $\pi$  phase shift, then the phase conditions between the two arms at the input to the second 3 dB coupler are exchanged with regard to the previous symmetrical case and the second coupler acts in reverse to the first one. Therefore, the light wave goes back to its original WG – the MZI is in bar state. From the above description, it is clear that the cross state requires that the combination of the two couplers (continuation in the coupling direction) will provide an exact full coupling ( $X + Y = 1$ , where  $X$  and  $Y$  are the two power coupling ratios in the first and second couplers) whereas the bar state (reverse operation) requires only that the two couplers will be identical in coupling ratios ( $Y - X = 0 \Rightarrow X = Y$ ). Therefore, in order to allow both states one requires that the coupling ratio will be identical, but with a specific ratio of 3 dB couplers ( $X = Y$  and  $X + Y = 1 \Rightarrow X = Y = 1/2$ ). The identical requirement can be easily obtained with very good process tolerance since the two couplers are in relatively close proximity on the wafer. However, the exact 3 dB coupler requirement (or any other specific value) is much harder to achieve with standard process capabilities. As a result, the bar state can be achieved more easily than the cross state and its cross-talk (CT) is usually much better than in cross state (approximately  $-30$  dB vs  $-15$  dB). The wavelength dependence of the  $2 \times 2$  MZI comes from the 3 dB coupler wavelength dependence and also from the OPL difference which depends on  $\lambda$ . We expect a wider spectral response at bar state (requires only identical couplers) than at cross state (requires exact 3 dB couplers). The cross state wavelength-insensitivity can be improved by using wavelength insensitive couplers, but these are usually more complicated—increase loss and length than with the conventional 3 dB couplers and should be used only when necessary.

The operational principle of the third type—combined configuration (Figure 7.18c) splitter and 3 dB coupler—is described as follows. First, light is equally split by the splitter with no phase difference (in phase). In this case the input to the 3 dB coupler is only a symmetrical eigenmode hence the total output is also symmetrical. Therefore, the  $1 \times 2$  MZI with no OPL difference in the arms act as a 3 dB coupler in its passive state. When the two arms have a phase difference of  $\pi/2$  all the light will emerge from one output WG, and when the phase difference is  $-\pi/2$  all the light will emerge from the second output WG similarly to the cross state of the  $2 \times 2$  MZI. The main difference between the  $1 \times 2$  MZI and the  $1 \times 1$  or  $2 \times 2$  MZIs is in its smaller required maximum power values (half) due to the half phase shift required ( $\pi/2$  vs  $\pi$ ). Yet the average power in time is similar since the  $1 \times 2$  requires a constant activation and the others consume power only in about half of their life. In addition, the  $1 \times 2$  configuration lacks the process immunity of the bar state in the  $2 \times 2$  MZI.

A very important feature can be achieved with an MZI—in addition to the full cross or bar states (binary) – one can tune the MZI to get any intermediate state to allow an analog output conditions. This can be obtained by inducing intermediate phase conditions and enable various



**Figure 7.19**  $1 \times 1$  MZI relative power transmission  $\eta$  as function of phase difference  $\Delta\phi/\pi$  for  $\tau_i = 1$

applications, such as weighted multicasting and weighted equalization at the output ports of an optical circuit [55].

The MZI can be modeled quantitatively by various techniques, such as the BPM, FDTD, finite element method, CMT, and others. Here we demonstrate the calculation in a CMT based transfer matrix formalism for the  $1 \times 1$  and the  $2 \times 2$  MZIs.

For the  $1 \times 1$  MZI the output power can be derived by adding the two incoming waves from the two arms and using  $P_o \propto |\sum E_i|^2$  (see for example [56]) to obtain

$$\eta \equiv \frac{P_o}{P_{in}} = \frac{1}{2}[1 + \cos(\Delta\phi)] = \cos^2(\Delta\phi/2) \quad (7.15a)$$

In the case where optical loss (absorption or scattering) occurs in the two arms, this can be generalized to

$$\eta \equiv \frac{P_o}{P_{in}} = \frac{1}{4} [\tau_1^2 + \tau_2^2 + 2\tau_1\tau_2 \cos(\Delta\phi)] \quad (7.15b)$$

where  $\Delta\phi \equiv \phi_2 - \phi_1 = \beta_2 L_2 - \beta_1 L_1$  is the phase difference between the two WG arms and  $\tau_i^2 = \exp(-\alpha_i L_i)$  is the power transmission in arm  $i$  ( $i = 1, 2$ ) due to loss and  $\tau_i$  are for field amplitude transmission. In Figure 7.19 the relative power dependence is depicted as a function of  $\Delta\phi/\pi$  for  $\tau_i = 1$ .

As expected for  $\Delta\phi = 0$ , we obtain an on state ( $\eta = 1$ ) and for  $\Delta\phi = \pi$  we obtain an off state ( $\eta = 0$ ). To illustrate the influence of the WG loss on ER, we consider a symmetrical MZI with arm length of 1 mm, with phase modulation based on the free-carrier plasma dispersion effect (PDE). In order to obtain an off state, a  $\Delta n_{eff}$  of  $7.75 \times 10^{-4}$  is required at  $\lambda_0 = 1.55 \mu m$ . This can be done by the plasma dispersion effect via a free carrier concentration change of  $\Delta N_e = \Delta N_h \cong 2 \times 10^{17} cm^{-3}$  [56]. However, the phase modulation also introduces an optical absorption of  $\Delta\alpha = 2.9(1/cm)$ , i.e.,  $\tau^2 = 0.75$ . Thus, from Equation (7.15b) we obtain  $\eta(\Delta\phi = \pi) \cong 4.5 \times 10^{-3}$ . Therefore, the ER is limited to be  $ER_{max} = 10\log(1/4.5 \times 10^{-3}) \cong 23.5 dB$ .

For the  $2 \times 2$  MZI configuration based on symmetric 3 dB DCs ( $\delta = 0$ ), the transfer matrix of the combination of DC1 + phase shifter + DC2 (see Figure 7.18b) is given by

$$\begin{pmatrix} E_{o1} \\ E_{o2} \end{pmatrix} = \mathbf{M} \begin{pmatrix} E_{i1} \\ E_{i2} \end{pmatrix} \quad (7.16)$$

where

$$\mathbf{M} \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \\ = \begin{pmatrix} \cos(\kappa_2 z_2) & -i \sin(\kappa_2 z_2) \\ -i \sin(\kappa_2 z_2) & \cos(\kappa_2 z_2) \end{pmatrix} \begin{pmatrix} \exp(i \Delta\phi/2) & 0 \\ 0 & \exp(-i \Delta\phi/2) \end{pmatrix} \begin{pmatrix} \cos(\kappa_1 z_1) & -i \sin(\kappa_1 z_1) \\ -i \sin(\kappa_1 z_1) & \cos(\kappa_1 z_1) \end{pmatrix}$$

Assuming  $\ell \equiv z_1 = z_2$ , for the simple case of single input (in powers,  $P_{in} \equiv P_{i1}$ ,  $P_{i2} = 0$ ) and two exact 3 dB DCs<sub>[ $\ell=\pi/(4\kappa)$ ]</sub>

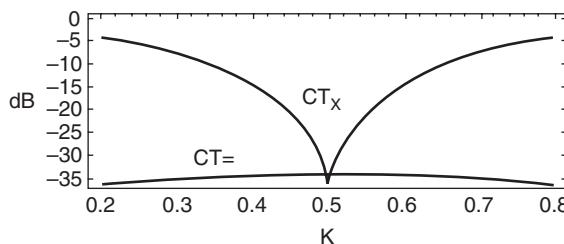
$$\eta_x \equiv \frac{P_{o2}}{P_{in}} = \cos^2(\Delta\phi/2) \\ \eta_- \equiv \frac{P_{o1}}{P_{in}} = \sin^2(\Delta\phi/2) \quad (7.18)$$

In the case of equal DCs, but not necessarily 3 dB couplers, if we define the power coupling ratio of the DCs as  $K \equiv \sin^2(\kappa\ell)$ , we obtain for single input ( $P_{in} \equiv P_{i1}$ ,  $P_{i2} = 0$ )

$$\eta_x \equiv \frac{P_{o2}}{P_{in}} = 4K(1-K)\cos^2(\Delta\phi/2) \\ \eta_- \equiv \frac{P_{o1}}{P_{in}} = 1 - \eta_x = (1-2K)^2\cos^2(\Delta\phi/2) + \sin^2(\Delta\phi/2) \quad (7.19)$$

where Equation (7.18) is a special case of Equation (7.19) for which  $K = 1/2$ . For the bar state, cross-talk (CT) is defined as the relative power emerging at the cross output and in dB it is defined as  $CT \equiv 10\log(\eta_x)$ , and at cross state the CT is emerging in the bar output and defined in dB as  $CT_x \equiv 10\log(\eta_-)$ . Figure 7.20 describes the simulated CT near both the cross state and bar state as function of the power coupling ratio  $K$ .

As expected, the bar state is much more stable and does not depend on the coupling ratio (CR) of the couplers, whereas the cross state is very sensitive to the CR. Therefore, the bar state is much less wavelength dependent and much more tolerant to process fluctuations in comparison to the cross state. Usually in silica on Si switches the immunity to CT at the bar state of a single  $2 \times 2$  MZI provides a CT of only about  $-30$  dB leakage to the cross output for polarization-independent broadband ( $\sim 50$  nm) operation with wide process tolerances. On the other hand a relatively high CT of about  $-15$  dB is obtained in the cross state at the bar output. For a switching matrix ( $M \times N$ ,  $M$  inputs and  $N$  outputs, e.g.  $8 \times 8$  or  $16 \times 16$ ), one can use a double-stage MZI to improve CT, and the fabric design should use the tolerance merit of the



**Figure 7.20** Simulated CT close to cross state ( $\Delta\phi = 0.01 \times \pi$ ) and close to bar state ( $\Delta\phi = 0.988 \times \pi$ ) as function of the power-coupling ratio  $K$

bar state in  $2 \times 2$  MZI and avoid, as far as possible, cross states [55, 57], these provide better CT, loss and PDL at broadband operation.

An MZI can act also as an optical filter with free spectral range of (Equation 7.19)  $FSR = \Delta\nu = c/\Delta$  (*OPL*) where  $\Delta$  (*OPL*) =  $n_2 L_2 - n_1 L_1$ . The simplest way to create an OPL difference is to employ different arms lengths, as mentioned previously. In this case, the phase shift between the two arms is given by  $\Delta\phi = 2\pi n_{\text{eff}} \Delta L / \lambda$ . When the arm length difference satisfies the relation  $\Delta L = \lambda_1 \lambda_2 / (2n_{\text{eff}} |\lambda_2 - \lambda_1|)$ , one input beam with wavelengths  $\lambda_1$  and  $\lambda_2$  is split into two beams with  $\lambda_1$  at one output and  $\lambda_2$  at another output. In general, many useful filter functions can be obtained through cascading of MZIs, each designed with different delay and different coupling ratio. A chain of  $n$  cascaded MZIs, i.e.  $n$  delays and  $n + 1$  couplers ( $2n + 1$  parameters) allows synthesizing a general optical filter, called an optical Fourier filter. The transfer function of such a filter can be obtained very elegantly, as suggested by Li and Henry [46] the cascaded MZIs structure can be easily obtained by the principle of the sum of all  $2^n$  possible optical paths from the input to the output.

Telecom applications with optical fibers require polarization-independent (PI) photonic devices; hence additional special care is needed in designing PI MZIs. If for instance the WG has birefringence due to material, geometry, or stress reasons, a different  $\Delta\phi$  may be obtained for TE and TM polarizations. This in turn creates polarization dependence (PD) of the MZI, which may ruin its ER and CT performance or even its excess loss via PDL. If we define the PD of an MZI as the separation in  $\Delta\phi$  between the TE and TM characteristics of the MZI, the birefringence influence can be included as follows.  $\Delta\phi$  is equal to  $\Delta(\beta L)$  hence a variance in  $\Delta\phi$  is  $\delta[\Delta(\beta L)] = \delta(\beta \Delta L + L \Delta \beta) = \Delta L \delta \beta + L \delta(\Delta \beta)$ , therefore

$$PD \equiv \delta(\Delta\phi) = k_0(B \Delta L + L \Delta B) \quad (7.20)$$

where  $B = |n_{\text{TE}} - n_{\text{TM}}|$  is the birefringence and  $\Delta B$  is the difference in  $B$  between the two WG arms. From Equation (7.20) we see that  $B$  can create PD in an MZI response which is proportional only to  $\Delta L$ , whereas if a different  $B$  occurs between the two WG arms (due to design or process stress effects) it can create significant PD in the MZI response since it is proportional to the whole arm length  $L$ . To generalize the PD discussion of MZIs, the PD within the couplers should be considered also. In general, Y-junctions and MMIs are much more PI than DCs, but may contribute higher losses (by  $\sim 0.5$  dB for MZI).

As expected, the MZI is also one of the most investigated devices in silicon photonics. The most widespread and well-developed element is the Si modulator based on  $1 \times 1$  MZI, but with various different modulation mechanisms employed on one or both of the WG arms. For a modulator the thermo-optic (TO) effect is too slow ( $\sim$ KHz) whereas the free carrier plasma dispersion effect can be much faster (MHz and GHz) and can be realized in numerous ways. The first is a forward-bias Pin diode [56, 58, 59] in which an operation of 10–20 MHz was achieved. The GHz boundary was first crossed by Intel [60] and later on was extended to 10 GHz [61]. Both were based on forward-biased MOS capacitors operating in the very fast accumulation state. Recently, a high-speed modulator based on a reverse-biased *pn* diode was designed and fabricated [62]. Using such a modulator, data transmission with bit rate of 30 Gb/s was demonstrated.

Other  $1 \times 1$  MZIs were also suggested based on the TO effect, see for example in [52] and [63] on Si wire WGs. The TO effect in Si WGs is based on the very high polarizability of silicon and bandgap dependence on temperature which create a TO coefficient of

$dn/dT \cong 1.86 \times 10^{-4}(\text{K}^{-1})$ , which is more than 15 times higher than in  $\text{SiO}_2$  or  $\text{SiON}$  materials. In addition, the thermal conductivity in silicon is much higher than in other materials; about 100 times higher than in  $\text{SiO}_2$ ,  $\text{SiON}$ , or polymers. Both the thermal conductivity and the TO coefficient provide a much faster and more efficient TO effect in silicon than in other materials. As an example, in [63] the reported switching time is  $3.5 \mu\text{s}$  with power consumption of  $50 \text{ mW}$ , whereas in silica-on-silicon [55, 57] switching time is  $\sim 1 \text{ ms}$  and power consumption is  $300\text{--}400 \text{ mW}$ . Therefore, a Si-based  $2 \times 2$  MZI TO switch may be a viable and very attractive. For example, Harjanne *et al.* [64] proposed a  $2 \times 2$  DC-MZI TO rib WG switch with switching time of  $\sim 0.75 \mu\text{s}$ , ER of  $15 \text{ dB}$ , and power consumption of  $\sim 300 \text{ mW}$ . The reason for the fast switching time is the differential control method which allows faster response on the account of higher powers.

The spectral filtering capability of MZIs was also demonstrated on silicon WG technology. In [65] Jalali *et al.* have demonstrated a wavelength splitter based on Si  $2 \times 2$  DC-MZI that splits two wavelengths within the C-band with FSR of  $8 \text{ nm}$  and CT of  $-18 \text{ dB}$ . Liu *et al.* [66] have demonstrated a MZ wavelength combiner based on  $2 \times 2$  MMI-MZI device for Raman amplification ( $\lambda_1 = 1440 \text{ nm}$ ,  $\lambda_2 = 1556 \text{ nm}$ ) with wavelength selectivity of  $20 \text{ dB}$ , insertion loss (including fiber-WG couplings) of  $9 \text{ dB}$ , and PDL of  $0.8 \text{ dB}$ .

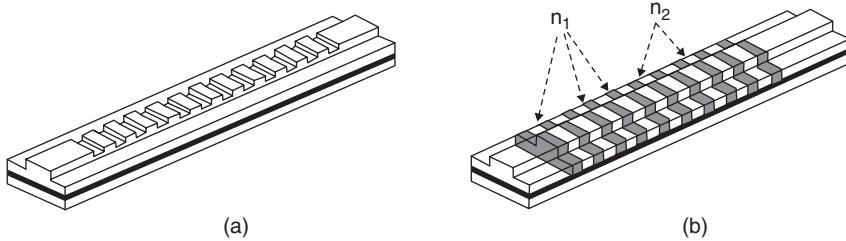
The MMI structure possesses no gap between WGs (as opposed to DCs) and much less polarization dependence, thus it is attractively employed within a  $2 \times 2$  MZI. In [67] House *et al.* have demonstrated a  $2 \times 2$  MMI-MZI TO switch showing a switching speed of  $\sim 10 \mu\text{s}$  and an ER of  $23.5 \text{ dB}$ . The on-chip MZI loss was  $\sim 1 \text{ dB}$ , mainly due to the MMI loss of  $0.4 \text{ dB/MMI}$  and MMI PDL of less than  $0.2 \text{ dB}$ . In [68] Liang and Tsang have demonstrated a  $2 \times 2$  MMI-MZI acting as a polarization beam splitter (PBS), based on geometric birefringence of rib WGs for the WG arms to produce different phase difference for TE and TM polarizations by  $\pi$ . They show an ER of  $\sim 13 \text{ dB}$  within the C-band with excess on-chip loss of  $1.1 \text{ dB}$ .

A  $4 \times 4$  switching matrix in SOI-based on a fabric of five TO  $2 \times 2$ -MMI-MZI elements was considered by Wang *et al.* [69] and by Li *et al.* [70], demonstrating at  $\lambda = 1550 \text{ nm}$  a CT of  $-12$  to  $-20 \text{ dB}$  with on-chip path loss of  $6.6\text{--}10 \text{ dB}$  and switching time of less than  $30 \mu\text{s}$ . The power consumption per switch element was  $330 \text{ mW}$ .

The  $1 \times 1$  MZI was also designed and fabricated in photonic crystals (PC). Silicon is a promising platform for compact 2D optical integrated circuits (OIC) PC to provide additional properties to the SOI material system. When PC is incorporated into an MZI structure it leads to a significant enhancement in the phase efficiency due to the slow group velocity effect at PCs. Hence, it may further reduce the device size by several orders of magnitude. Gu *et al.* [71] have demonstrated this while reducing power consumption of the TO effect at  $\lambda = 1550 \text{ nm}$ ; they obtained heating power of  $70 \text{ mW}$  in a switching time of less than  $20 \mu\text{s}$  and MZI WG arm length of only  $100 \mu\text{m}$ .

## 7.6 Bragg Gratings

Bragg gratings [72] are spectral filters that typically reflect light in a narrow band of wavelengths and allow light transmission in all other wavelengths. They are based on the principle of Bragg reflection and are formed by creating a periodic structural corrugation or material refractive index modulation along the length of a waveguide.



**Figure 7.21** Schematics of two waveguide grating designs: (a) based on surface corrugations; (b) based on refractive index modulation

Figure 7.21 shows schematics of two waveguide gratings. One has shallow corrugations etched into the top surface of the rib waveguide, and the other has alternating regions of higher and lower refractive indices built into the rib waveguide. The surface corrugation results in a periodic refractive index modulation along the grating structure, because the effective index of the waveguide depends on the waveguide geometry. In general, any periodic variation in a waveguide dimension can create a Bragg grating. For example, etching a silicon slab instead of the rib is another way to fabricate a silicon waveguide grating.

Each unit or period of the grating, due to its structural or index discontinuity, acts as a weak reflector where light would experience both reflection and transmission. If the length of each period  $\Lambda$  is such that all the partial reflections add up in phase, which occurs when the round trip of the light between two reflections is an integer multiple of the wavelength, the Bragg condition is satisfied and the total reflection can sum up to be nearly 100%. The Bragg phase-matching condition can therefore be expressed simply as

$$2\beta_{\text{Bragg}}\Lambda = 2m\pi \quad (7.21)$$

where  $m$  is an integer and  $\beta_{\text{Bragg}}$  is the propagation constant of the optical wave meeting the Bragg condition, defined as

$$\beta_{\text{Bragg}} = \frac{2\pi n_{\text{ave}}}{\lambda_{\text{Bragg}}} \quad (7.22)$$

where  $n_{\text{ave}}$  is the averaged effective refractive index of the entire grating structure and  $\lambda_{\text{Bragg}}$  is the reflected wavelength or Bragg wavelength. By substituting Equation (7.21) into Equation (7.22), the Bragg condition simplifies to

$$m\lambda_{\text{Bragg}} = 2n_{\text{ave}}\Lambda \quad (7.23)$$

where  $m$  is the order of the grating.

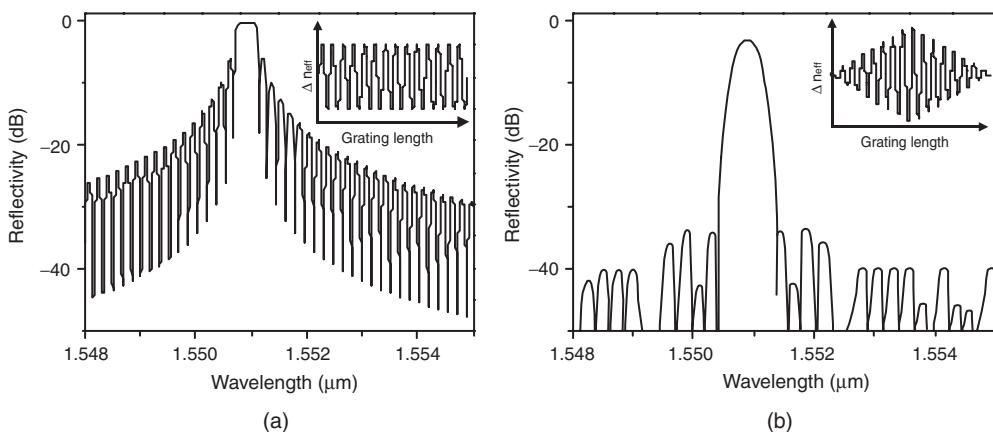
A first-order ( $m = 1$ ) Bragg grating is usually analyzed using the CMT that describes the coupling between the forward and backward propagating waves at a given wavelength. For a uniform grating in which the refractive index varies along the grating longitudinal axis  $z$  as  $n(z) = n_{\text{ave}} + \Delta n \cos(2\pi z/\Lambda)$ , the grating amplitude reflection coefficient is given by

$$r_g = \frac{i\xi \sin(qL)}{q \cos(qL) - i\Delta\beta \sin(qL)} \quad (7.24)$$

where  $\Delta\beta = \beta - \beta_{\text{Bragg}}$  is the propagation constant detuning away from the Bragg wavelength,  $\zeta = 2\pi \Delta n/\lambda$  is the coupling coefficient that is related to the index modulation depth,  $q = \sqrt{\Delta\beta^2 - \zeta^2}$ , and  $L$  is the grating length. The grating power reflectivity is simply given by  $|r_g|^2$ . Note that the waveguide loss is not included in the above analysis.

As it stands, Equation (7.24) suggests that the grating reflection spectrum strongly depends on the grating coupling coefficient or the refractive index modulation depth and grating length. The grating spectral bandwidth, centered on the Bragg wavelength, as well as the grating reflectivity can therefore be tailored to meet the intended application requirement. To achieve a narrow bandwidth, the grating coupling strength needs to be weak, such that the reflection by each grating period is small. This can be designed by reducing the cross-section of the structural corrugation, reducing the periodic material refractive index contrast, or reducing the overlap of the optical mode with the grating. For such weak gratings, bandwidth is also inversely proportional to the grating length, so the longer the grating, the narrower the bandwidth. Increasing the grating length has the added benefit of increasing the grating reflectivity. Nonetheless, reflectivity and bandwidth saturation will eventually occur, either when the incident beam no longer interacts with additional grating periods due to complete reflection or transmission loss or when the reflected light no longer makes it out of the grating due to loss. When the grating strength is increased, grating bandwidth and reflectivity can both increase; furthermore, saturation can be achieved with a shorter device.

Uniform gratings are simple to design and fabricate, but their spectra generally exhibit large power fluctuations, known as side-lobes, at wavelengths around the grating bandwidth (see Figure 7.22a). To reduce these side-lobes, the grating strength needs to gradually increase and then decrease, for example, with a Gaussian-shape envelope, along the length of the waveguide. Structures based on this design are known as apodized gratings. For illustration, Figure 7.22 shows the effective refractive index modulations as well as the modeled reflection spectra of a uniform and an apodized grating with duty cycle variations. Note the suppression of the side-lobes with apodization. Another common Bragg grating design is to linearly vary (or chirp) the grating period or effective index modulation along the structure. In this case, the grating



**Figure 7.22** Modeled reflection spectra of: (a) a uniform grating; (b) an apodized grating with duty cycle variations. Representative index contrast profiles are shown as inserts

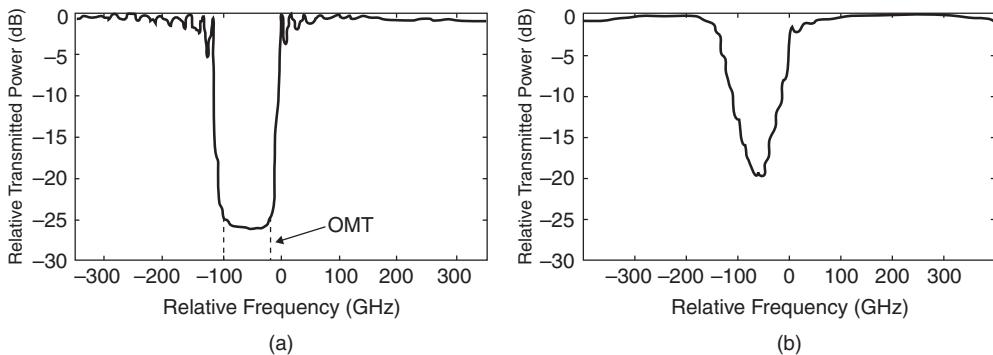
reflects a much wider band of wavelengths; furthermore, because of the varying period or index modulation, different wavelengths within the grating's spectral bandwidth are reflected at different depths of the grating. For instance, if the variation is decreasing along the direction of the forward propagating wave, light with longer wavelengths will be reflected earlier than those with shorter wavelengths. This spectrally dependent delay is very useful for chromatic dispersion compensation to extend the reach of optical links.

Some notable applications of waveguide Bragg gratings are optical filtering, optical add-drop multiplexing (OADM), and the enabling of narrow linewidth semiconductor lasers. For an ideal device, light that is not transmitted is reflected, so the transmission spectrum can be used to predict the reflection spectrum and *vice versa*. For many gratings, however, this is not the case. For instance, if the grating-induced excess loss is not negligible, then the reflectivity of the Bragg wavelength can not be accurately calculated simply by examining the transmission spectrum. So, depending on the application, different sets of grating performance parameters must be characterized by evaluating the grating's transmission and/or reflection spectra. For instance, for optical filtering where a band of wavelengths needs to be blocked while all others need to transmit unaffected, the transmission spectrum is more important to understand transmission loss, stop-band spectral bandwidth and extinction ratio (ER), and side-lobe suppressing. On the other hand, if the grating is used as a reflector, say for an external cavity laser, then the reflection spectrum needs to be analyzed to quantify reflection spectral bandwidth, reflectivity, and side-lobe suppression.

In recent years, researchers in the field of Si photonics have proposed and demonstrated numerous waveguide-based Bragg gratings in Si. Because most Si waveguides have effective refractive indices greater than 3, the first-order Bragg grating period must be  $< 0.25 \mu\text{m}$  to achieve a Bragg wavelength of  $1.55 \mu\text{m}$ . This small size makes device processing very challenging because the lithographic resolution needed is  $< 0.13 \mu\text{m}$ . While  $0.13 \mu\text{m}$  lithography is common in today's IC fabs, it is not readily available to most optics researchers. Many groups have therefore turned to e-beam lithography to pattern their gratings [73–75]. A group has also recently experimented with focus ion beam milling to define their devices [76]. All of these workers formed their gratings by etching structural corrugations either into the top surface or side-walls of the waveguides.

A rib waveguide grating filter, with  $4 \times 3 \mu\text{m}$  cross-section and surface corrugations of 150 nm depth and 223 nm period, demonstrated a transmission spectrum with multiple distinct dips, one of which corresponded well with the expected Bragg wavelength of  $\sim 1545 \text{ nm}$  [74]. The stop-band showed a maximum ER close to 10 dB and a 3 dB bandwidth (spectral width at half maximum) of 0.12 nm, or 15 GHz. All other dips were shown, through simulation, to be light coupled into higher-order leaky modes. To increase the spectral separation of these leaky modes so they lie outside the wavelength range of interest, the researchers proposed shrinking the waveguide dimensions.

Waveguide gratings with side-wall corrugations also demonstrated promising performance. A rib device with  $1.6 \times 2.2 \mu\text{m}$  cross-section demonstrated a distinct stop-band with 0.8 nm, or 100 GHz, 3-dB bandwidth centered around the Bragg wavelength of 1553 nm [77]. It also achieved an ER of  $\sim 25 \text{ dB}$  across the entire 100 GHz stop-band. Because the grating is etched into the side walls of the grating, the researchers also experimented with an apodized design by varying the waveguide width and corrugation width along the waveguide's length such that the grating strength followed a raised cosine apodization function. Unfortunately, the fabricated device did not exactly duplicate the design, so the filter response was suboptimal with

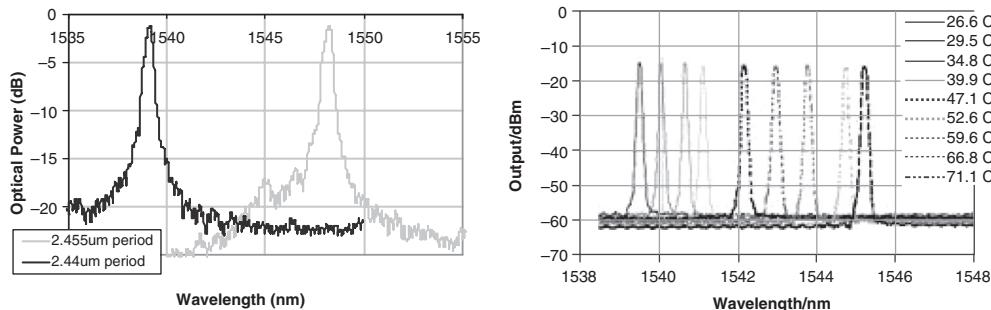


**Figure 7.23** (a) TE transmission spectra for a uniform side wall grating. The dashed line indicates the response predicted by coupled-mode theory; (b) TE transmission spectrum for an apodized sidewall grating. The center wavelength for both is 1553 nm. Reproduced from [77] by permission of IEEE © (2004)

nonuniform extinction across the stop-band; maximum ER also dropped to  $\sim 20$  dB. Nonetheless, this apodized grating did demonstrate drastic side-lobe suppression. The transmission spectra of both the uniform and apodized gratings are shown in Figure 7.23 (a) and (b).

The side wall corrugation design was also demonstrated using  $0.3 \times 0.3 \mu\text{m}$  strip waveguides, actually as a part of an integrated OADM [75]. The gratings have a corrugation width of 30 nm and a period of 370 nm. The grating period here is larger than the rib-waveguide-based devices because the effective refractive index of this small strip waveguide is much smaller, at  $\sim 2$ . The spectral response of both the through and drop ports of the multiplexer was characterized. Though the transmission and reflection spectra of the gratings did not exactly correlate, the basic function of the multiplexer was clearly demonstrated. The dropped Bragg wavelength was 1551.4 nm, the 3 dB bandwidth was  $\sim 1$  nm, and the ER of the stop-band was 18 dB. Furthermore, because the refractive index of silicon is sensitive to temperature, the Bragg wavelength of the grating can be dynamically tuned using heaters. The researchers demonstrated this tuning capability using metal thin-film heaters placed on top of the Bragg gratings. The metal is separated from the waveguide core by 1  $\mu\text{m}$  of  $\text{SiO}_2$  top cladding to minimize optical loss. Measurements showed a dropping wavelength shift of 6.6 nm with 0.8 W of heating power, which resulted in a temperature increase of 48 K of the Bragg grating waveguides.

While some Si photonics researchers experimented with structural corrugations as the basis for forming Bragg gratings, others investigated techniques to create material refractive index modulation. Starting with amorphous Si ( $\alpha$ -Si) films, one group studied the use of femtosecond laser annealing to crystallize the  $\alpha$ -Si [78]. Material analyses revealed that a pattern of alternating  $\alpha$ -Si and microcrystalline Si lines with 2  $\mu\text{m}$  period was created. Because Si with different crystalline structures have different refractive indices, it is possible to use the laser annealing technique to create Si Bragg gratings. Another group also recognized the potential of altering Si's material structure to tailor its refractive index; they demonstrated a grating with  $3.5 \times 4 \mu\text{m}$  rib waveguide cross-section and alternating regions of Si and polycrystalline Si (poly-Si) [79]. The poly-Si regions were formed by first etching multiple trenches into the Si along the length of the waveguide and then filling the trenches with



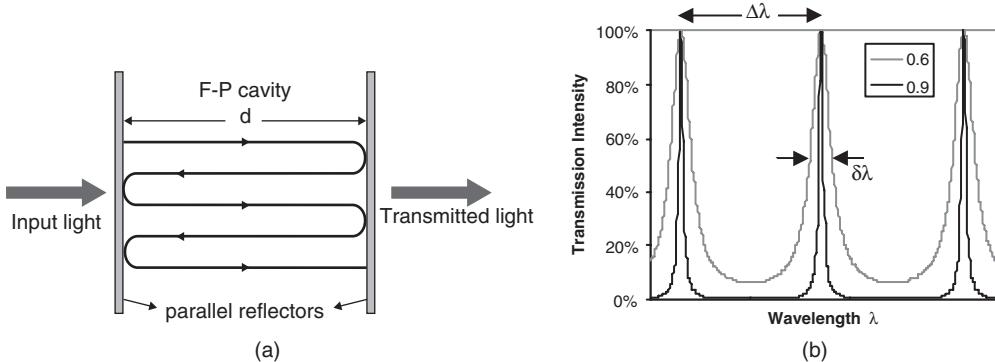
**Figure 7.24** (a) Reflection spectra of two poly-Si Bragg gratings with different grating periods; (b) ECL optical output, showing spectral shifting with temperature tuning (a Bragg grating with  $2.44\text{ }\mu\text{m}$  period was used). Reproduced from [80] by permission of the Optical Society of America

$\alpha$ -Si. This amorphous material was crystallized into poly-Si using a high-temperature anneal. Poly-Si was used to create the grating, as opposed to  $\alpha$ -Si, because it has refractive index better matched to Si and lower transmission loss. The grating created was an 11th-order grating, which allowed the grating period to be significantly increased to  $2.455\text{ }\mu\text{m}$  for a Bragg wavelength of  $1548.2\text{ nm}$ . The reflection spectra of two gratings are shown in Figure 7.24(a); both have  $3\text{ dB}$  bandwidth of  $0.5\text{ nm}$  and  $\text{ER} > 12\text{ dB}$ . The peak reflectivity is measured to be  $70\%$ , which is limited partly because of the high order of the grating. Thermal tunability was demonstrated by placing this device on a thermoelectric cooler (TEC) module. The observed tuning is very linear, with a Bragg wavelength shift of  $1.3\text{ nm}$  per  $10^\circ\text{C}$  change in temperature.

As a proof of concept demonstration of the applicability of the poly-Si grating, it was butt-coupled to a gain chip to form an external cavity laser (ECL) [80]. One facet of the gain chip had a reflectivity of  $90\%$ , and the other was angle-polished and coated with anti-reflective coating so its reflectivity was  $\sim 10^{-5}$ . The Bragg grating is brought within close proximity of the low-reflectivity facet of the gain chip. It, together with the high-reflectivity facet, formed the Fabry–Perot resonant cavity for the laser (Fabry–Perot resonators will be discussed in detail in the following section.) The optical output of the ECL was collected from the high-reflectivity side of the gain chip using a lensed single-mode fiber. The ECL ran single-mode and had a measured line-width of  $118\text{ MHz}$  and a maximum output power of  $0.45\text{ mW}$ . The laser output could be further increased with improved optical coupling between the gain chip and grating as well as coupling between the gain chip and fiber. Again using the TEC, thermal tuning of the ECL was demonstrated. Figure 7.24(b) shows the spectral shift of the laser output as a function of temperature. Note that the laser output had side-mode suppression in excess of  $40\text{ dB}$ , and it maintained its peak intensity and shape with tuning.

## 7.7 Fabry–Perot Resonators

A Fabry–Perot (F–P) interferometer or resonator typically consists of a transparent medium bound by two parallel reflecting surfaces facing each other; a schematic representation is shown in Figure 7.25(a). For simplicity, the following discussion will only address devices



**Figure 7.25** (a) Schematic representation of an F-P resonator; (b) the calculated F-P transmission spectra for two  $R$  values: 0.6 and 0.9

where the reflecting surfaces are perpendicular to the direction of light propagation. The transmission spectrum of an F-P exhibits a series of transmission peaks characteristic of the resonances of the device. This distinct spectral dependence, together with the F-P's design simplicity, has made the F-P resonator a valuable part of many applications. The most notable is the resonant cavity of a laser, and the ECL discussed in the previous section is of course one example. Some other applications include optical modulation, WDM channel monitoring, spectral filtering to reduce the dispersion penalty of optical links, and chemical and gas sensing.

Light inside an F-P experiences multiple reflections at its two reflecting surfaces. The interference of these reflected light results in a periodic variation of the transmitted optical intensity as a function of wavelength. Maximum transmission occurs when the reflections are in-phase and the light interferes constructively; minimum transmission occurs when the light is out-of-phase and interferes destructively. The optical phase difference  $\Delta\phi$  between two successive reflections can be expressed as

$$\Delta\phi = \frac{4\pi n_{\text{eff}} d}{\lambda} \quad (7.25)$$

where  $n_{\text{eff}}$  is the effective refractive index and  $d$  is the length of the transparent medium, or cavity, of the resonator;  $n_{\text{eff}}$  is used generically here so the equation can describe all F-P resonators, including those based on waveguides. In the simple F-P resonator example given above,  $n_{\text{eff}}$  is simply the material refractive index of the cavity. Assuming that the two parallel surfaces have the same intensity reflection coefficient  $R$ , the intensity of the transmitted light is given by

$$I = \frac{(1 - R)^2}{1 + R^2 - 2R \cos(\Delta\phi)} \quad (7.26)$$

The transmission spectra of two F-P resonators, one with  $R$  of 0.6 and the other with  $R$  of 0.9, are shown in Figure 7.25(b). The transmission is maximum whenever  $\Delta\phi$  is an integer multiple of  $2\pi$ , which is the condition for resonance or constructive interference. A simple manipulation

of Equation (7.25) will yield the following expression for the resonant wavelengths,  $\lambda_r$ :

$$\lambda_r = \frac{2n_{\text{eff}}d}{m} \quad (7.27)$$

The wavelength separation  $\Delta\lambda$  between adjacent transmission peaks, or resonances, is known as the free spectral range (FSR) of the F–P and is given by

$$\Delta\lambda = \frac{\lambda_r^2}{2n_g d} \quad (7.28)$$

where  $\lambda_r$  is the nearest resonant wavelength and  $n_g = n_{\text{eff}} - \lambda(dn_{\text{eff}}/d\lambda)$  is the group index of refraction. A parameter that is commonly used to compare the spectral selectivity of different resonators is the finesse  $F$ , which is defined as

$$F = \frac{\Delta\lambda}{\delta\lambda} = \frac{\pi\sqrt{R}}{1-R} \quad (7.29)$$

where  $\delta\lambda$  is the 3 dB bandwidth of any transmission peak. For the two resonators depicted in Figure 7.25 (b),  $F$  is  $\sim 6$  for  $R$  of 0.6 and  $\sim 30$  for  $R$  of 0.9.

F–P-based devices have been demonstrated in Si for a variety of applications. For optical communications, they have been studied as a low-cost solution for wavelength monitoring and spectral filtering of WDM channels in the C and L bands [81–83]. Most of these experiments used bulk Si wafers (tens to hundreds of micrometers thick) as the F–P cavity and multilayer dielectrics as the reflecting surfaces. Cocorullo *et al.* [81] started with a double-side polished Si wafer  $\sim 105$   $\mu\text{m}$  thick and, using e-beam evaporation, deposited alternating  $\lambda/4$  layers of  $\text{SiO}_2$  and  $\alpha\text{-Si}$  on its two surfaces. With reflectors comprised of two pairs of  $\text{SiO}_2$  and  $\alpha\text{-Si}$  layers, they obtained a resonator with 3.27 nm FSR and 0.24 nm 3 dB bandwidth, for a finesse of 13.7. Again, by taking advantage of Si's strong thermo-optic effect, the researchers heated the resonator by 32°C and demonstrated 3 nm tuning of the resonant wavelength. For a DWDM system with 50 GHz channel spacing, this F–P-based tunable optical filter can monitor seven wavelength channels with  $\sim 10$  dB isolation. Taking the same approach of sandwiching a Si wafer between multilayer dielectric mirrors, Niemi *et al.* [82] demonstrated a F–P based optical filter with even narrower 3 dB bandwidth of 0.11 nm. They achieved this by using three pairs of alternating  $\text{Si}_3\text{N}_4$  and  $\text{SiO}_2$  for a reflectivity of 0.66. Using this narrow-band device, the researchers filtered the optical output of a distributed feedback (DFB) laser that was directly modulated at a bit rate of 2.5 Gb/s. The spectral filtering improved the signal ER from 9.2 to 13.3 dB. Bit error rate (BER) performance was tested with 350 km of standard single-mode fiber. The dispersion penalty improved from 2.4 to 0.9 dB at a BER level of  $10^{-9}$ .

Using wafer bonding, Zuo *et al.* [83] demonstrated flat-top F–P optical filters in Si. Filters with flat-top passbands are desirable for channel detection in optical networks because their signal fidelity and tolerance of signal drift are important. To create the device, the researchers first e-beam sputtered a triple-layer of  $\text{SiO}_2/\alpha\text{-Si}/\text{SiO}_2$  on the carrier wafer as the bottom reflector. On top of this, they bonded a second wafer, which was then thinned by chemical mechanical polish to 30–50 nm to eventually act as the resonator cavity. Multiple 10  $\mu\text{m}$  square holes were etched 8 nm deep and 10  $\mu\text{m}$  apart into this top Si layer. Finally, a single pair of  $\text{SiO}_2/\alpha\text{-Si}$  films was deposited to form the top reflector. Areas of the Si cavity that were etched effectively had cavity lengths 8 nm shorter than areas that were not etched. The final

transmission peaks of the resonator were therefore each a superposition of the two transmission peaks that corresponded to the two different cavity lengths. Measurements showed flat-top response with 1 dB bandwidth of 1 nm, 3 dB bandwidth of 3 nm, and FSR of 8 nm. The obvious broadening of the transmission peaks is of course undesirable; it is likely related to roughness of the mirrors and should be improved with process optimization.

For today's dense wavelength division multiplexing (DWDM) systems, where channel spacings of 100 and 50 GHz are common, optical filter bandwidths of less than 0.8 nm are desirable. This requirement is or can be met by the Si-based F-P filters discussed above, but one common issue all these devices share is that their FSR are all less than 10 nm such that multiple channels are passed at the same time. To increase FSR, the cavity length must be decreased. For example, to achieve an FSR of 40 nm, the cavity length must be < 9 nm. This should be achievable with wafer bonding, Si micromachining, and waveguide-based devices.

For key active components of an optical link, such as lasers, modulators, and detectors, Si-based F-P devices have also proven valuable. In the previous section, an ECL was discussed where a poly-Si Bragg grating was used as a wavelength selective reflector for the laser's external cavity. Heikkinen *et al.* [84] made a similar demonstration using a Si-based F-P interferometer. It was formed through micromachining where the movable top and stationary bottom mirrors were separated by  $\sim 2.5\text{ }\mu\text{m}$  of free space, formed by chemical etching of Si. The device was aligned to the low-reflectivity facet of an edge-emitting F-P laser diode, and a fiber was butt-coupled to the opposite, high-reflectivity facet. By optimizing the cavity length of the F-P interferometer, the researchers achieved 0.1 mW single-mode lasing at  $\sim 1543\text{ nm}$  with side-mode suppression greater than 25 dB. Spectral width at 20 dB below the peak power was 0.18 nm, significantly improved compared with the 8 nm without the interferometer. Furthermore, by adjusting the distance between the F-P top mirror and the laser facet, they demonstrated 13 nm tuning of the laser output. These demonstrations of ECLs show how the combination of simple Si filters and low-cost gain chips can enable low-cost lasers with narrow linewidth and tunability.

Photodetectors have also seen performance improvement with the incorporation of F-P resonators. While Si photodetectors typically have much lower dark current compared to III-V devices, they suffer from poor optical absorption efficiency due to its indirect bandgap. To achieve acceptable responsivity, they require long absorption path lengths, which can adversely affect device bandwidth because photo-generated carriers need to traverse a longer distance to reach the contacts. One way to circumvent this problem is to fabricate the photodetector inside an F-P cavity to use the resonant effect to enhance absorption for a given detector thickness. Unlu *et al.* [85] used this approach and made Si *p-i-n* detector arrays in  $\sim 2\text{ }\mu\text{m}$  epitaxial device layers of wafers formed by a repeat SOI process. The bottom reflector is a two-pair Si/SiO<sub>2</sub> structure with reflectivity in excess of 90%, and the top reflector is simply the epitaxial Si/air interface. The detectors achieved responsivity above 0.3 A/W at 822 nm and bandwidth in excess of 10 GHz, which are performances that could soon compete with existing compound semiconductor photodetectors, but at a fraction of the cost. The benefit of this type of resonant cavity device is of course the increase in absorption efficiency without compromise of bandwidth, but its drawback is the wavelength selectivity of the photodetector efficiency.

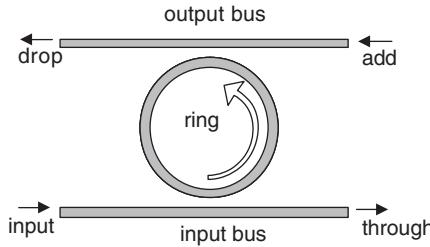
The high spectral selectivity of the F-P resonator, together with its sensitivity to even small changes in the cavity refractive index, is the basis for F-P-based optical modulators. Barrios *et al.* [86] demonstrated a compact, tunable F-P resonator integrated with an SOI waveguide. The cavity was formed using 20  $\mu\text{m}$  of the waveguide and reflectors consisting of three pairs of

Si/SiO<sub>2</sub> that were etched down to the buried oxide layer. Doped areas,  $p^+$  and  $n^+$ , were defined in the cavity region on the sides of the rib to form a  $p-i-n$  diode to electrically drive the device. The device transmission showed resonant peaks with 1.54 nm 3 dB bandwidth and 11.2 finesse. To achieve modulation, the measurement was made using a wavelength that corresponded to one of the resonances, 1552.89 nm. When no voltage was applied to the diode, transmission is maximum. When a voltage was applied, both carrier injection and heating occur. Unfortunately, the two effects result in opposite changes of the refractive index. The experiment observed a red-shift of the F-P spectrum, indicating that the thermo-optic effect was dominating. As a result of the shift in resonance, optical transmission of the probe beam dropped. Modulation depth, defined as the change in optical power (due to applied voltage) divided by the maximum power, is 3 dB for an electrical drive of 20 mW. Calculations show that if the device's high transmission loss can be reduced, modulation depth can improved to as much as 7 dB. Despite the suboptimal performance of the demonstrated modulator, it does show the feasibility of using F-P resonators to realize compact, efficient devices.

Besides devices for optical communications, Si F-P resonators have also demonstrated their utility in chemical sensing and spectroscopy for chemical analysis. For sensing, the resonant cavity needs to be, or contain, free space so it can be filled with the chemical gases or liquids to be tested. To create such cavities, the two common approaches are porous Si and Si micromachining to create suspended reflectors [87–89]. Under test, the substitution of air in the cavity with chemicals will increase its refractive index, leading to a red-shift of the F-P resonant wavelength or peak. Because different chemicals have different characteristic refractive indices, they can be identified by quantifying the resulting resonant shift. Spectrometers for chemical analysis operate similarly to the wavelength monitoring filters for DWDM. They are used to examine the spectral characteristics of light that has passed through gases or liquids under test. Depending on the chemicals present, parts of the spectrum is absorbed. The F-P spectrometer is tuned so its resonant peak is scanned across the wavelength range of interest. Together with a detector, the optical spectrum can be captured for analysis. For this application, wide wavelength tuning is required. One demonstration used an array of 16 tunable Si F-P resonators, each with a different cavity length. Together, the resonator array covered the entire visible spectrum [90]. F-P-based sensors are of course not limited to chemical sensing and analysis; they can be used to monitor many different parameters that result in a change in refractive index. Furthermore, the F-P resonant effect has also been used extensively to measure the propagation loss of optical waveguides. An optical waveguide with polished facets can be regarded as a resonant cavity because the facets are in essence reflectors. The facet of a Si waveguide, for instance, has a reflection coefficient of  $\sim 0.3$ . Light coupled into the waveguide therefore undergoes multiple reflections as it propagates along the waveguide and back. The loss coefficient can therefore be determined by measuring the output spectrum of the waveguide and comparing the ratio of its maximum intensity to its minimum intensity [56].

## 7.8 Ring Resonators

Another resonator design that has gained widespread interest is the ring resonator (RR), which is functionally the same as the F-P resonator. The difference is that it consists of a waveguide in a closed loop, commonly in the shape of a ring or race-track. Also, the RR provides traveling wave operation, as opposed to the standing wave operation characteristic of F-P resonators.



**Figure 7.26** Top-down schematic view of a ring resonator

As Figure 7.26 illustrates, light can be coupled into the ring via evanescent field coupling by placing the input waveguide, also known as the input bus, within close proximity of the ring. Like the F-P, the ring behaves as an interferometer and will be resonant for light whose phase change after each full trip around the ring is an integer multiple of  $2\pi$ , which is of course the condition for the light in the ring to be in-phase with the incoming light and constructive interference. Light that does not meet this resonant condition is transmitted through the bus waveguide. The expression for the resonant wavelengths of the ring is very similar to that of the F-P and is given by:

$$\lambda_r = \frac{2\pi R n_{\text{eff}}}{m} \quad (7.30)$$

where  $R$  is the ring radius with circular waveguide ring, and  $m$  is an integer.

When the RR is coupled to a single waveguide, the transmission response of the resonant wavelengths through the bus strongly depends on the optical loss of the ring and the coupling efficiency between the ring and bus waveguide. The device can act as a phase filter where all wavelengths are transmitted and the resonant wavelengths, having also traversed the ring, acquire a phase change. Or, it can exhibit notch filter behavior where all the light on-resonance is coupled to the ring, giving high extinction in the transmission spectrum. To capture or separate the resonant wavelengths from the rest, an additional waveguide, an output bus, can be placed on the opposite side of the ring, as shown in Figure 7.26. If the optical loss in the ring is negligible, all the on-resonance light that is coupled into the ring can be coupled into the output bus towards the drop port. This design configuration is actually even more versatile; it can have all the functionality of an optical add-drop multiplexer (OADM) because wavelengths can also be added via the add port, get coupled into the RR, and exit the through port via the input bus waveguide.

Similar to the F-P resonators, key performance parameters of the RR include the FSR, the ER, and the finesse. Actually, for historical reasons, the quality (Q) factor is more often quoted than the finesse as a measure of the spectral selectivity of the RR. While the finesse is defined as the FSR divided by the 3 dB bandwidth of the resonance, the Q factor is defined as the resonant wavelength divided by the 3 dB bandwidth. As expected, the expression for the FSR of a RR is very similar to that of the F-P resonator, it is given by

$$\Delta\lambda = \frac{\lambda_r^2}{2\pi R n_g} \quad (7.31)$$

To calculate the finesse or Q factor, one must of course first define the 3 dB bandwidth. Assuming the coupling between the ring and the two bus waveguides are weak and can both be represented by the coupling coefficient  $k$ , the 3 dB bandwidth can be approximated as

$$\delta\lambda \approx \frac{\kappa^2 \lambda_r^2}{2\pi^2 R n_g} \quad (7.32)$$

When the RR is on resonance, light coupled into the ring constructively interferes with the input light; as a result, optical intensity in the ring can build up and be significantly higher than that in the bus waveguide. This field enhancement is an important property of RRs and can be measured by its finesse or Q factor. As light makes multiple round trips in the ring, it will also experience loss—transmission loss of the ring waveguide and loss due to coupling to the bus waveguides. If  $N$  is the number of round trips required to reduce the optical energy to  $1/e$  of its initial value, finesse is given by

$$F = 2\pi N \quad (7.33)$$

and the Q factor is given by

$$Q = \omega_r T N \quad (7.34)$$

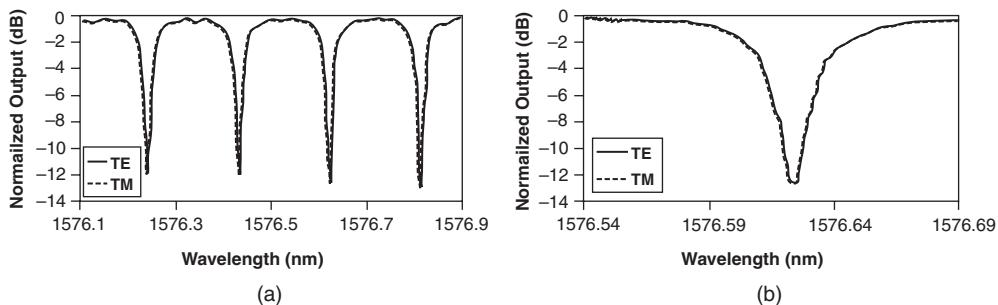
where  $\omega_r$  is the resonant frequency and  $T$  is the time it takes to make one trip around the ring. It is therefore easy to see that to achieve high field enhancement, high finesse or Q, the transmission loss of the ring must be minimized and the coupling efficiency must be optimized.

SOI-based RRs have been fabricated and successfully demonstrated by numerous research teams. Many are based on very compact, single-mode strip waveguides whose core cross-sections are approximately  $200 \times 500$  nm. In SOI systems, where optical confinement is strong due to the large refractive index contrast between the Si core and SiO<sub>2</sub> cladding, such small core dimensions ensure that very compact rings with bending radii on the order of a few micrometers can be realized without significant bend loss. Numerous benefits can be associated with such small devices. The most obvious ones are that they allow high integration density and the realization of very high optical power density for nonlinear optical effects. Furthermore, because a RR's FSR is inversely proportional to the ring diameter, small rings have large FSRs, which is a key requirement for WDM spectral filters because usually only one wavelength or channel is dropped at a given time. There are numerous challenges, however, associated with the successful demonstration of such compact RRs. One of the most critical design parameters that must be carefully controlled is the coupling efficiency between the ring and the input/output waveguide. Since the device operates through evanescent coupling that exponentially depends on the gap separation, lithographic resolution and critical dimension (CD) control are very important. For the sub-micrometer waveguides mentioned above, the gap between the ring and bus waveguide is generally less than 400 nm. Thus, in order to reliably process RR devices according to design, control of a few nanometers demands CD control only achieved by modern deep UV lithography processes used for IC production or e-beam lithography used in research environments. Another key challenge is to keep the transmission loss of the ring low. As Equations (7.33) and (7.34) clearly illustrate, high loss results in small  $N$  which ultimately leads to low finesse or Q-factor. For small waveguides, whose optical mode interacts strongly with the core's sidewalls and surfaces, optical scattering caused by side wall

roughness is usually the primary source of transmission loss. The origin of side wall roughness can be traced to the lithography and etch processes, so these process steps are generally very carefully developed and scrutinized. Any residual roughness that is present after etch can be further minimized with oxidation. This oxidation smoothing can be performed with oxygen gas and/or steam at elevated temperatures or with multi-step wet chemical oxidation [91, 92].

Very promising filter performance has been demonstrated using RRs based on single-mode SOI strip waveguides. In one experiment, an OADM was formed using waveguides with  $0.4 \times 0.4 \mu\text{m}$  cross-section, ring with  $3 \mu\text{m}$  radius, and ring-to-bus separations of  $0.15 \mu\text{m}$  [93]. The total footprint of the device is less than  $60 \mu\text{m}^2$ . To achieve good dimension control and smooth waveguide sidewalls, the researchers used e-beam lithography and reactive ion etching. For TM-polarized light, this OADM achieved a Q as high as 2500 and FSR of 26.5nm. For the resonant wavelengths, attenuation at the through port was 10 dB and ER at the drop port was 25 dB. Due to loss inside the ring, however, attenuation or insertion loss at the drop port was 5.5 dB. The biggest contributor to this loss was waveguide side wall roughness. To minimize side wall scattering loss, other groups devoted considerable efforts to optimize their lithography and etch processes. They also used wider waveguides to reduce the interaction between the optical mode and side walls. Duman *et al.* [94] used deep UV lithography to define waveguides  $0.5 \mu\text{m}$  wide and  $0.22 \mu\text{m}$  tall. TE transmission loss was as low as 2.4 dB/cm, of which nearly half was believed to be loss due to leakage to the substrate through the insufficiently thick  $1 \mu\text{m}$  bottom oxide cladding. For their OADM demonstration, the researchers used a racetrack design, which enabled more precise control of the coupling efficiency between the ring and bus waveguides by controlling the coupling length rather than just the separation. This racetrack resonator, with  $3 \mu\text{m}$  coupling length and  $5 \mu\text{m}$  radius, achieved a Q-factor more than 3000 and FSR of  $\sim 14 \text{ nm}$ . Compared to the previous RR, its add-drop cross-talk is significantly improved. At resonance, attenuation at the through port is nearly 25 dB, ER at the drop port is nearly 20 dB, and attenuation at the drop port is  $\sim 3 \text{ dB}$ . By focusing on optimization of their e-beam lithography and plasma etching processes, Tsuchizawa *et al.* [95] demonstrated even better waveguide transmission in their study of compact Si RRs; they achieved transmission losses as low as 2.8 dB/cm for  $0.2 \times 0.4 \mu\text{m}$  waveguides. Using a  $5 \mu\text{m}$  ring, positioned  $0.3 \mu\text{m}$  away from the bus waveguides, they made channel-dropping filters with Q-factors as large as 13 000, FSR of 17.9 nm, ER greater than 20 dB, and insertion loss less than 3 dB.

While these researchers were investigating small strip waveguides for RR applications, others were experimenting with larger rib waveguides. One motivation was that it is much easier to minimize waveguide losses of these larger devices to achieve even higher Q values for improved wavelength selectivity. Kiyat *et al.* [96] used a racetrack resonator design with  $1 \times 1 \mu\text{m}$  rib waveguides and  $0.8 \mu\text{m}$  ring-to-bus separation. They increased the ring radius from 150 to 500  $\mu\text{m}$  and found the Q-factor increase because of the associated reduction in bend loss of the ring. For 350  $\mu\text{m}$  radius, the Q-factor was 119 000 and ER was greater than 10 dB. Nonetheless, because of the large ring radius, FSR was only  $\sim 290 \text{ pm}$ . All the RRs discussed so far are polarization dependent because it is challenging to simultaneously achieve polarization insensitive waveguides and ring-to-bus coupling. Headley *et al.* [97] were able to do exactly this using  $1 \times 1.35 \mu\text{m}$  rib waveguides and a racetrack resonator whose coupling length was carefully selected to yield similar coupling efficiency for both TE and TM polarizations. While the coupling length required to transfer optical power from one waveguide to the other is different for TE polarization compared with TM, the researchers were able to



**Figure 7.27** (a) Spectral response of the fabricated resonator; (b) close-up of the peak at 1576.624 nm. Reproduced from [97] by permission of American Institute of Physics

design a polarization-independent coupler by using a coupling length that allowed the TE optical power to make multiple complete transitions and the TM to make a single transition. Their resonator consisted of a racetrack with 400  $\mu\text{m}$  radius and a single bus waveguide. As Figure 7.27 illustrates, the resonator indeed demonstrated polarization-independent response with the TE and TM spectra aligned to each other within 1 pm. The Q-factor was 90 000 and ER was  $\sim$ 12 dB. Again, because the resonator has a large radius, FSR is only 190 pm.

Like the F-P resonator, Si RRs have also been used to make active devices. Xu *et al.* [98] created an electro-optic modulator based on the plasma dispersion effect, where a change in free-carrier density changes the refractive index of the material. The modulator consists of a  $p-i-n$  diode with a 6  $\mu\text{m}$  radius Si ring as the intrinsic region. The ring and single bus waveguides have a cross-section of  $0.25 \times 0.45 \mu\text{m}$  and exhibits filter response with  $Q > 39\,000$ . The area inside the ring is doped  $p$ -type, and area outside the ring is doped  $n$ -type. The effective refractive index of the ring can be changed using carrier injection when a forward-bias is applied to the device. With this change in effective index, the resonant wavelength shifts, which changes or modulates the optical power at the through port. The dynamic response of the modulator was measured using both non-return-to-zero and return-to-zero driving signals, data transmission was demonstrated at both 0.4 and 1.5 Gb/s, respectively.

## 7.9 Conclusion

In this chapter, we have described some selected passive photonic devices such as waveguide bends, directional couplers, MMI couplers, Y-splitter, MZIs, waveguide Bragg gratings, Fabry–Perot resonators, and ring resonators. An attempt was made to provide fundamental principles of operation and their manifestation in CMOS-compatible technologies. We have described the state-of-the-art results in the silicon based fundamental building blocks. We believe that these most basic and critical elements covered in this chapter as well as other important active elements such as modulators, detectors, and lasers will enable low-cost integrated silicon photonic chips for various applications.

One of the unique features in silicon photonics is the possibility to construct high-confinement and sub-micrometer-size waveguides by using silicon and oxide. Because of the high refractive index contrast in the silicon waveguide material system, one can design truly

miniaturized integrated optical circuits with bend radii of micrometers and MMI couplers, tens of micrometers long allowing very small footprints of devices, modules, and systems. We note that all the components described in this chapter have been demonstrated in other material systems and have shown excellent performance. Silicon-based counterparts demonstrated similar or comparable performance in many cases. However, in few cases, additional progress is required to provide fully mature solutions for better performance; nevertheless, it is clear in all given examples that silicon adds unique unprecedented features to integrated photonics.

The huge potential in silicon photonics technology (miniaturization, integration with microelectronics, power reduction, optical interconnects, and significant cost reduction) has started to be explored and it is very likely that during the next decade silicon will become the main photonic material for the benefit of diverse areas of applications.

## References

- [1] E. A. J. Marcatili, ‘Bends in optical dielectric guides’, *Bell Sys. Tech. J.* **48**, 1969, 2103–2132.
- [2] M. Heiblum and J. H. Harris, ‘Analysis of curved waveguides by conformal transformation’, *IEEE J. Quantum Electron.* **QE-11**, 1975, 75–83.
- [3] J. Gu, P. Besse and H. Mechior, ‘Method of lines for the analysis of the propagation characteristics of curved optical waveguides’, *IEEE J. Quantum Electron.* **QE-27**, 1991, 531–537.
- [4] T. Yamamoto and M. Koshiba, ‘Numerical analysis of curvature loss in optical waveguides by finite-element method’, *J. Lightwave Technol.* **11**, 1993, 1579–1583.
- [5] T. Kitoh, N. Takato, M. Yasu and M. Kawachi, ‘Bending loss reduction in silica-based waveguides by using lateral offsets’, *J. Lightwave Technol.* **13**, 1995, 555–562.
- [6] S. Kim and A. Gopinath, ‘Vector analysis of optical dielectric waveguide bends using finite-difference method’, *J. Lightwave Technol.* **14**, 1996, 2085–2092.
- [7] S. S. A. Obayya, B. M. A. Rahman, K. T. Grattan and H. A. El-Mikati, ‘Beam propagation modeling of polarization rotation in deeply etched semiconductor bent waveguides’, *IEEE Photon. Technol. Lett.* **13**, 2001, 681–683.
- [8] L. Prkna, M. Hubalek and J. Ctyroky, ‘Vectorial eigenmode solver for bent waveguides based on mode matching’, *IEEE Photon. Technol. Lett.* **13**, 2004, 2057–2059.
- [9] R. Jedidi and R. Pierre, ‘Efficient analytical and numerical methods for the computation of bent loss in planar waveguides’, *J. Lightwave Technol.* **23**, 2005, 2278–2284.
- [10] Details available from [www.rsoftinc.com](http://www.rsoftinc.com).
- [11] Details available from [www.photond.com](http://www.photond.com).
- [12] S. L. Chuang, *Physics of Optoelectronic Devices*, John Wiley, New York, 1995, pp. 270–273.
- [13] Y. A. Vlasov and S. J. McNab, ‘Losses in single-mode silicon-on-insulator strip waveguides and bends’, *Opt. Express*, **12**, 2004, 1622–1631.
- [14] G. B. Cao, L. J. Dai, Y. J. Wang, J. Jiang, H. Yang and F. Zhang, ‘Compact integrated star coupler on silicon-on-insulator’, *IEEE Photon. Technol. Lett.* **17**, 2005, 2616–2618.
- [15] H. Kogelnik, ‘Coupled wave theory for thick hologram gratings’, *Bell Sys. Tech. J.* **48**, 1969, 2909–2947.
- [16] D. Marcuse, ‘The coupling of degenerate modes in two parallel dielectric waveguides’, *Bell Sys. Tech. J.* **50**, 1971, 1791–1816.
- [17] A. W. Snyder, ‘Coupled mode theory for optical fibers’, *J. Opt. Soc. Am.* **62**, 1972, 1267–1277.
- [18] A. Yariv, ‘Coupled mode theory for guided-wave optics’, *IEEE J. Quant. Elect.* **9**, 1973, 919–933.
- [19] A. Hardy and W. Streifer, ‘Coupled mode theory of parallel waveguides’, *J. Lightwave Tech.*, **3**, 1985, 1135–1146.
- [20] W. Streifer, M. Osinski and A. Hardy, ‘Reformulation of coupled-mode theory of multiwaveguide systems’, *J. Lightwave Tech.*, **5**, 1987, 1–4.
- [21] D. Marcuse, *Theory of Dielectric Optical Waveguides*, Academic Press, New York, 1991.
- [22] H. A. Haus, W. P. Huang, S. Kawakami and N. A. Whitaker, ‘Coupled mode theory of optical waveguides’, *J. Lightwave Tech.*, **5**, 1987, 16–23.

- [23] S. L. Chuang, 'A coupled mode formulation by reciprocity and a variational principle', *J. Lightwave Tech.*, **5**, 1987, 5–15.
- [24] H. Kogelnik, 'Theory of dielectric waveguide', In: *Integrated Optics*, T. Tamir (ed.), Berlin Springer, 15–79, 1975.
- [25] N. Izhaky and A. Hardy, 'Analysis of grating-assisted backward coupling employing the unified coupled-mode formalism', *J. Opt. Soc. Am. A*, **16**, 1999, 1303–1311.
- [26] S. E. Miler, 'Integrated optics: An introduction', *Bell Sys. Tech. J.*, **48**, 1969, 2059–2069.
- [27] P. D. Trinh *et al.*, 'Integrated optical directional couplers in silicon-on-insulator', *Elect. Lett.*, **31**, 1995, 2097–2098.
- [28] G. B. Cao *et al.*, 'Directional couplers realized on silicon-on-insulator', *IEEE Photon. Tech. Lett.*, **17**, 2005, 1671–1673.
- [29] I. Kiyat *et al.*, 'A compact silicon-on-insulator polarization splitter', *IEEE Photon. Tech. Lett.*, **17**, 2005, 100–102.
- [30] H. Yamada *et al.*, 'Optical directional couplers based on Si-wire waveguides', *IEEE Photon. Tech. Lett.*, **17**, 2005, 585–587.
- [31] L. B. Soldano and E. C. M. Pennings, 'Optical multimode interference devices based on self-imaging: principles and applications', *J. Lightwave Tech.*, **13**, 1995, 615–627.
- [32] O. Bryngdahl, 'Image formation using self-imaging techniques', *J. Opt. Soc. Amer.* **63**, 1973, 416–419.
- [33] R. Ulrich, 'Image formation by phase coincidences in optical waveguides', *Opt. Commun.* **12**, 1975, 259–264.
- [34] 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 Tech.*, **10**, 1992, 1843–1850.
- [35] D. S. Levy, R. Scarmozzino and R. M. Osgood, Jr., 'Length reduction of tapered  $N \times N$  MMI devices', *IEEE Photon. Tech. Lett.*, **10**, 1998, 830–832.
- [36] D. S. Levy, K. H. Park, R. Scarmozzino, R. M. Osgood, Jr., C. Dries, P. Studenkov and S. Forrest, 'Fabrication of ultracompact 3 dB  $2 \times 2$  MMI power splitters', *IEEE Photon. Tech. Lett.*, **11**, 1999, 1009–1011.
- [37] H. Wei, J. Yu, Z. Liu, X. Zhang, W. Shi and C. Fang, 'Fabrication of  $4 \times 4$  tapered MMI coupler with large cross section', *IEEE Photon. Tech. Lett.*, **13**, 2001, 466–468.
- [38] U. Fischer, T. Zinke and K. Petermann, 'Integrated optical waveguide switches in SOI', *Proc. 1995 IEEE International SOI Conference*, pp. 141–142.
- [39] P. D. Trinh, S. Yegnanarayanan, F. Coppinger and B. Jalali, 'Compact multimode interference couplers in silicon-on-insulator technology', *Proc. CLEO'97*, p. 441.
- [40] D. Dai, S. Liu, S. He and Q. Zhou, 'Optimal design of an MMI coupler for broadening the spectral response of an AWG demultiplexer', *J. Lightwave Tech.*, **20**, 2002, 1957–1961.
- [41] X. Jiang, X. Li, H. Zhou, J. Yang, M. Wang, Y. Wu and S. Ishikawa, 'Compact variable optical attenuator based on multimode interference coupler', *IEEE Photon. Tech. Lett.*, **17**, 2005, 2361–2363.
- [42] T. Kamalakis and T. Sphicopoulos, 'Numerical study of the fabrication tolerance of conventional and MMI-flattened AWG's', *IEEE Photon. Tech. Lett.*, **16**, 2004, 1876–1878.
- [43] D. Xu, S. Janz and P. Cheben, 'Design of polarization-insensitive ring resonators in silicon-on-insulator using MMI couplers and cladding stress engineering', *IEEE Photon. Tech. Lett.*, **18**, 2006, 343–345.
- [44] E. C. M. Pennings, R. van Roijen, M. J. N. van Stralen, P. J. de Waard, R. G. M. P. Koumans and B. H. Verbeek, 'Reflection properties of multimode interference devices', *IEEE Photon. Tech. Lett.*, **6**, 1994, 715–718.
- [45] Y. Shibata, S. Oku, M. Yamamoto and M. Naganuma, 'Reflection characteristics and cascadability of a multimode interference 3 dB coupler', *IEE Proc. Optoelect.*, **149**, 2002, 217–221.
- [46] Y. P. Li and C. H. Henry, 'Silica-based optical integrated circuits', *IEE Proc. Optoelect.*, **143**, 1996, 263–280.
- [47] W. K. Burns *et al.*, *Appl. Phys. Lett.*, **29**, 1976, 790.
- [48] H. Sasaki and R. M. De La Rue, *Elect. Lett.*, **12**, 1976, 459.
- [49] Y. Silberberg *et al.*, 'Digital optical switch', *Appl. Phys. Lett.*, **51**, 1987, 1230–1232.
- [50] A. G. Rickman and G. T. Reed, 'Silicon-on-insulator optical rib waveguides: loss, mode characteristics, bends and y-junction's', *IEE Proc. Optoelect.*, **141**, 1994, 391–393.
- [51] C. K. Tang, A. K. Kewell, G. T. Reed, A. G. Rickman and F. Namavar, 'Development of library of low-loss silicon-on-insulator optoelectronic devices', *IEE Proc.-Optoelect.*, **143**, 1996, 312–315.
- [52] P. Dumon *et al.*, 'Basic photonic wire components in silicon-on-insulator', *IEEE International Conference on Group IV Photonics*, 2004, pp 1–3.
- [53] Y. L. Liu *et al.*, 'Silicon  $1 \times 2$  digital optical switch using plasma dispersion', *Elect. Lett.*, **30**, 1994, 130–131.
- [54] P. I. Borel *et al.*, 'Topology optimized broadband photonic crystal Y-splitter', *Elect. Lett.*, **41**, 2005, xxx–yyy.

- [55] N. Izhaky *et al.*, ‘Intelligent switches of integrated lightwave circuits with core telecommunication functions’, *Proc. SPIE*, **4284**, 2001, 54–63.
- [56] G. T. Reed and A. P. Knights, *Silicon Photonics an introduction*, John Wiley, 2004.
- [57] T. Goh *et al.*, ‘High extinction ratio and low loss silica-based 8 × 8 thermooptic matrix switch’, *IEEE Photon. Tech. Lett.*, **10**, 1998, 810–812.
- [58] C. K. Tang, G. T. Reed, A. J. Walton and A. G. Rickman, ‘Low loss single mode optical phase modulator in SIMOX material’, *J. Lightwave Tech.*, **12**, 1994, 1394–1400.
- [59] R. M. H. Atta *et al.*, ‘Fabrication of a highly efficient optical modulator based on silicon-on-insulator’, *Proc. SPIE*, **5116**, 2003, 495–500.
- [60] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu and M. Paniccia, ‘A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor’, *Nature* **427**, 2004, 615–618.
- [61] L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, O. Cohen, U. D. Keil and T. Franck, ‘High speed silicon Mach-Zehnder modulator’, *Opt. Express* **13**, 2005, 3129–3135.
- [62] A. Liu, L. Liao, D. Rubin, H. Nguyen, B. Ciftcioglu, Y. Chetrit, N. Izhaky and M. Paniccia, ‘High-speed optical modulation based on carrier depletion in a silicon waveguide’, *Opt. Express* **15**, 2007, 660–668.
- [63] R. L. Espinola *et al.*, ‘Fast and low-power thermooptic switch on thin silicon-on-insulator’, *IEEE Photon. Tech. Lett.* **15**, 2003, 1366–1368.
- [64] M. Harjanne *et al.*, ‘Sub-μs switching time in silicon-on-insulator Mach-Zehnder thermooptic switch’, *IEEE Photon. Tech. Lett.* **15**, 2004, 2039–2041.
- [65] B. Jalali *et al.*, ‘Guided-wave optics in silicon-on-insulator technology’, *IEE Proc.-Optoelectron.* **143**, 1996, 307–311.
- [66] Y. Liu *et al.*, ‘Silicon-on-insulator waveguide Mach-Zehnder wavelength combiner’, *IEEE International Conference on Group IV Photonics*, 2004, pp 536–537.
- [67] A. House *et al.*, ‘Silicon waveguide integrated optical switching with microsecond switching speed’, *OFC*, 2003, 449–450.
- [68] T. K. Liang and H. K. Tsang., ‘Silicon-on-insulator waveguide TE/TM polarization beam splitter’, *IEEE International Conference on Group IV Photonics*, 2004, pp 855–856.
- [69] Z. Wang *et al.*, ‘Rearrangeable nonblocking thermo-optic 4 × 4 switching matrix in silicon-on-insulator’, *IEE Proc.-Optoelectron.*, **152**, 2005, 160–162.
- [70] Y. Li, J. Yu and S. Chen, ‘Rearrangeable nonblocking SOI waveguide thermooptic 4 × 4 switch matrix with low insertion loss and fast response’, *IEEE Photon. Tech. Lett.*, **17**, 2005, 1641–1643.
- [71] L. Gu *et al.*, ‘Silicon-on-insulator-based photonic-crystal Mach-Zehnder interferometers’, *Proc. SPIE*, **6128**, 2006, 1–7.
- [72] A. Othonos and K. Kalli, *Fiber Bragg Gratings: Fundamentals and Applications in Telecommunications and Sensing*, Artech House, London, 1999.
- [73] V. R. Subramanian, R. G. DeCorby, J. N. McMullin, C. J. Haugen and M. Belov, ‘Fabrication of aperiodic gratings on silicon-on-insulator (SOI) rib waveguides using e-beam lithography’, *Proc. SPIE*, **4654**, 2002, 45–53.
- [74] T. E. Murphy, J. T. Hastings and H. I. Smith, ‘Fabrication and Characterization of Narrow-Band Bragg-Reflection Filters in Silicon-on-Insulator Ridge Waveguides’, *J. Lightwave Technol.*, **19**, 2001, 1938–1942.
- [75] T. Chu, H. Yamada, S. Ishida and Y. Arakawa, ‘Reconfigurable Optical Add-Drop Multiplexer Based on Silicon Nano-Wire Waveguides’, *ECOC Proceedings*, **2**, 2005, 245–246.
- [76] D. J. Moss, V. G. Ta’eed, B. J. Eggleton, D. Freeman, S. Madden, M. Samoc, B. Luther-Davies, S. Janz and D.-X. Xu, ‘Bragg gratings in silicon-on-insulator waveguides by focused ion beam milling’, *Appl. Phys. Lett.*, **85**, 2004, 4860–4862.
- [77] J. T. Hastings and H. I. Smith, ‘Apodized optical filters based on silicon-on-insulator waveguides with sidewall gratings’, *CLEO Proceedings*, **2**, 2004, 16–21.
- [78] G. J. Lee, J. Park, E. K. Kim, Y. P. Lee, K. M. Kim, H. Cheong, C. S. Yoon, Y.-D. Son and J. Jang, ‘Microstructure of femtosecond laser-induced grating in amorphous silicon’, *Opt. Express*, **13**, 2005, 6445–6453.
- [79] L. Liao, A. Liu, S. Pang and M. J. Paniccia, ‘Tunable Bragg grating filters in SOI waveguides’, *OSA Integrated Photonics Research Technical Digest*, 2004, Paper IThE2.
- [80] R. Jones, M. J. Paniccia and S. Merritt, ‘Tunable waveguide based external cavity laser using a silicon/poly-silicon Bragg grating’, *OSA Integrated Photonics Research Technical Digest*, 2004, Paper IThE4.
- [81] G. Cocorullo, F. G. Della Corte, M. Iodice and I. Rendina, ‘Simple and low-cost silicon Fabry-Perot filter for WDM channel monitoring’, *Digest of the LEOS Summer Topical Meetings*, 2000, 45–46.

- [82] T. Niemi, M. Uusimaa, S. Tammela, P. Heimala, T. Kajava, M. Kaivola and H. Ludvigsen, ‘Tunable silicon etalon for simultaneous spectral filtering and wavelength monitoring of a DWDM transmitter’, *IEEE Photon. Technol. Lett.*, **13**, 2001, 58–60.
- [83] Y. H. Zuo, R. W. Mao, Y. Y. Zheng, X. Shi, L. Zhao, W. H. Shi, B. W. Cheng, J. Z. Yu and Q. M. Wang, ‘A Si-based tunable narrow-band flat-top filter with multiple-step-type Fabry–Perot cavity structure’, *IEEE Photon. Technol. Lett.*, **17**, 2005, 2134–2136.
- [84] V. Heikkinen, J. Aikio, T. Alajoki, J. Hiltunen, A.-J. Mattila, J. Ollila and P. Karioja, ‘Single-mode tuning of a 1540-nm diode laser using a Fabry–Perot interferometer’, *IEEE Photon. Technol. Lett.*, **16**, 2004, 1164–1166.
- [85] M. S. Unlu, M. K. Emsley, O. I. Dosunmu, P. Muller and Y. Leblebici, ‘High-speed Si resonant cavity enhanced photodetectors and arrays’, *J. Vac. Sci. Technol.*, **22**, 2004, 781–787.
- [86] C. A. Barrios, V. R. Almeida, R. R. Panepucci, B. S. Schmidt and M. Lipson, ‘Compact silicon tunable Fabry–Perot resonator with low power consumption’, *IEEE Photon. Technol. Lett.*, **16**, 2004, 506–508.
- [87] G. Lammel, S. Schweizer and P. Renaud, ‘MEMS infrared gas spectrometer based on a porous silicon tunable filter’, *The 14th IEEE International Conference on Micro Electro Mechanical Systems*, 2001, 578–581.
- [88] L. De Stefano, L. Moretti, A. Lamberti, O. Longo, M. Rocchia, A. M. Rossi, P. Arcari and I. Rendina, ‘Optical sensors for vapors, liquids and biological molecules based on porous silicon technology’, *IEEE Transactions on Nanotechnology*, **3**, 2004, 49–54.
- [89] D. Cristea, M. Kusko, C. Tibeica, R. Muller, E. Manea and D. Syvridis, ‘Integrated tunable optical interferometer obtained using <111> oriented silicon micromachining’, *International Semiconductor Conference*, **1**, 2003, 159–162.
- [90] R. Wolffenbuttel, ‘State-of-the-art in integrated optical microspectrometers’, *IEEE Transactions on Instrumentation and Measurement*, **53**, 2004, 197–202.
- [91] K. K. Lee, D. R. Lim and L. C. Kimerling, ‘Fabrication of ultralow-loss Si/SiO<sub>2</sub> waveguides by roughness reduction’, *Opt. Lett.*, **26**, 2001, 1888–1890.
- [92] D. K. Sparacin, S. J. Spector and L. C. Kimerling, ‘Silicon waveguide sidewall smoothing by wet chemical oxidation’, *J. Lightwave Technol.*, **23**, 2005, 2455–2461.
- [93] A. Vorckel, M. Monster, P. Haring Bolivar, H. Kurz and W. Henschel, ‘Fabrication and characterization of ultra compact silicon-on-insulator micro-ring resonator add-drop-multiplexer for dense wavelength division multiplexing’, *Conference on Lasers and Electro-Optics*, 2003, 866–867.
- [94] P. Dumon, W. Bogaerts, V. Wiaux, J. Wouters, S. Beckx, J. Van Campenhout, D. Taillaert, B. Luyssaert, P. Bienstman, D. Van Thourhout and R. Baets, ‘Low-loss SOI photonic wires and ring resonators fabricated with deep UV lithography’, *IEEE Photon. Technol. Lett.*, **16**, 2004, 1328–1330.
- [95] T. Tsuchizawa, K. Yamada, H. Fukuda, T. Watanabe, J. Takahashi, M. Takahashi, T. Shoji, E. Tamechika, S. Itabashi and H. Morita, ‘Microphotronics devices based on silicon microfabrication technology’, *IEEE J. Selected Topics in Quantum Electron.*, **11**, 2005, 232–240.
- [96] I. Kiyat, A. Aydinli and N. Dagli, ‘High-Q silicon-on-insulator optical rib waveguide racetrack resonators’, *Opt. Express*, **13**, 2005, 1900–1905.
- [97] W. R. Headley, G. T. Reed, S. Howe, A. Liu and M. Paniccia, ‘Polarization-independent optical racetrack resonators using rib waveguides on silicon-on-insulator’, *Appl. Phys. Lett.*, **85**, 2004, 5523–5525.
- [98] Q. Xu, B. Schmidt, S. Pradhan and M. Lipson, ‘Micrometre-scale silicon electro-optic modulator’, *Nature*, **435**, 2005, 325–327.

# 8

# Integration

Cary Gunn

## 8.1 Motivation for Integration

The great promise of silicon photonics is manifested by integration. A CMOS transistor is not known to be a particularly high-performance transistor, but the ability to manufacture large numbers of them integrated on a single substrate has undoubtedly changed the course of technology. Similarly, most silicon photonics optical devices don't perform as well as they might in another material system, but the ability to integrate them in large numbers alongside CMOS circuitry holds great promise. Examples of integration are available today that dramatically shrink conventional optical technology and place tens of optical devices on single die with hundreds of thousands of transistors, and in the future, many believe that silicon photonics links will replace copper connections between microprocessors and memory, and even connect circuit blocks within multicore processors, as shown in Figure 8.1 [1, 2].

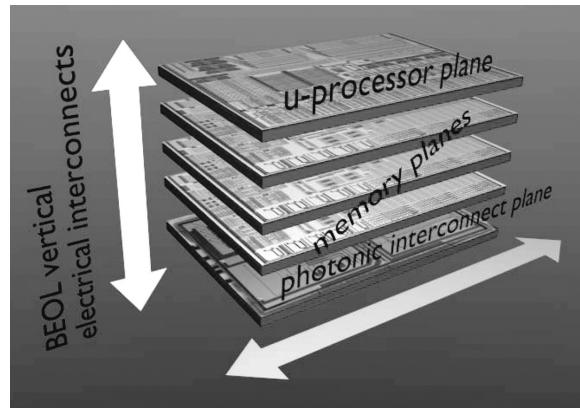
But perhaps the true advantages for integration are still unknown. Just as the first designers of integrated circuits could not have foreseen the modern microprocessor, who can say what clever designers will come up with when presented this technology as a basic 'toolkit' from which to start designing.

## 8.2 CMOS Integration Approaches

There are a variety of ways to approach integration of silicon photonics with CMOS circuitry, depending on which point in the CMOS process to integrate the optical functionality, and how complete the library of optical functions is to be.

### 8.2.1 Hybrid Integration

A simple approach is to fabricate the photonic devices in one substrate, and the electronic devices in the other, and simply connect the chips via flip-chip or wirebonding. This hybrid



**Figure 8.1** A concept for integration of CMOS photonics with traditional CMOS and memory technologies through wafer bonding and 3D integration (courtesy J.A. Kash, IBM Research)

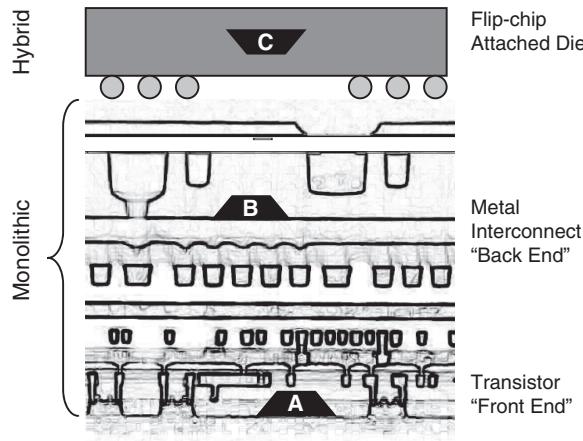
approach has been used by Intel to demonstrate silicon modulation [3] where the MZI modulator was fabricated in one die, and the driver was fabricated in another die, which was then wirebonded to the first one. The primary drawback of this technique is the lack of scalability due to the limited number of wirebonds that can be used to stitch the two chips together. A similar approach is to flip-chip one die on top of the other die. While this technique allows large numbers of circuits on one die to interface with a large number of components on the other die, it does not allow subsequent flip-chip mounting of the assembly into a standard packaging approach, severely limiting this approach for commercial applications. The main benefit of both of these approaches, however, is that circuit design can be accomplished in a standard process, such as that available from a foundry, and the silicon photonics components can be optimized independently of the constraints of CMOS compatibility.

### 8.2.2 Monolithic Integration into CMOS

Monolithic integration is being pursued due to its broad commercial appeal. Today, complete monolithic integration has yet to be demonstrated, particularly related to the light source. However, a wide variety of silicon photonic devices have been monolithically integrated with CMOS circuitry, allowing a large number of applications with the option of an external or hybrid-attached light source.

Both economic and performance justifications favor monolithic integration over the hybrid approach. Often, 50–75% of the costs of an optical system or component are related to assembly and test. Monolithic integration obviously impacts assembly, but it also makes system testing much easier and more affordable by allowing a majority of testing to be accomplished at the wafer level. Ultimately, as achieved with the microprocessors today, only a wafer probe need be performed before the part can be confidently packaged and sold.

When considering performance, an important benefit of monolithic integration is the improved sensitivity of an integrated receiver compared with that of an equivalent discrete or hybrid implementation. As discussed later in Section 8.3, a monolithically integrated silicon



**Figure 8.2** A cross-section of a CMOS wafer illustrating three options for integration of optics: (a) integration at the active device level; (b) integration in the back end metallization dielectrics; (c) hybrid or bonded devices

photronics receiver will offer substantially improved sensitivity, which dramatically impacts system design and the economics of the final solution.

A number of monolithic integration approaches are being pursued by companies and research institutions. Among these, a key differentiator is whether to integrate into the front-end or the back-end of the CMOS process. This difference is illustrated in Figure 8.2.

### 8.2.2.1 Front-end Integration

Integration of silicon photonics in the front-end of the CMOS process involves starting with a SOI CMOS wafer and simultaneously processing both transistors and optical devices in the silicon film. Subsequently, a standard back-end is constructed and used to connect the transistors and photonic components into the appropriate circuitry. To date, this is the only approach that has successfully realized monolithic integration of photonic components with CMOS circuitry, and will be covered in great detail in subsequent sections of this chapter.

### 8.2.2.2 Back-end Integration

An alternative approach integrates the silicon photonics components in the back-end of the CMOS process [4]. This can be done following two paths: (a) growth and patterning of low-temperature waveguide material such as amorphous silicon while processing in the back end; or (b) use of 3D integration approach to perform layer transfer of a silicon photonics layer on or within the back-end of a CMOS wafer. The major advantages of both back-end integration approaches are that the sensitive front end processing need not be modified, and a SOI transistor process is not required. The major drawback, on the other hand, is that the modification of the highly optimized thickness and properties of the back-end materials, requiring these steps to be redeveloped.

The amorphous silicon integration approach has the benefits of being readily compatible with back-end processing temperatures, but has severe limitations regarding active devices. Implants, for example, require an activation step at temperatures approaching 1000°C, well beyond the melting point of the metals used in the back-end processing. However, passive waveguide devices can be constructed in this fashion. Recent work on reducing the optical losses of amorphous silicon points to the attractiveness of this material as a waveguide [5].

The layer transfer process has the benefits of allowing the photonic devices to be fabricated on a separate wafer, using a custom photonics process that is unconstrained by the electronics process [6]. For example, growth of germanium and activation of implants can occur on the photonics wafer, which is then bonded at low temperature to the electronics wafer. It then becomes necessary to process the vias required to connect to, and to electrically transverse this layer.

While being pursued by a number of institutions, neither back-end integration approach has been realized as of the date of this writing.

## 8.3 Detailed Description of the Front-end Integration Approach

### 8.3.1 Transistor and CMOS Process Selection

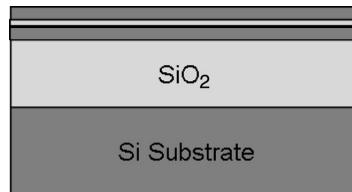
#### 8.3.1.1 Significance of SOI and the 130 nm CMOS Process Node

The two aspects of silicon manufacturing that make it useful for optical structures are the recent emergence of silicon-on-insulator (SOI) wafers, and the progression of DUV lithography below the quarter-wavelength of light in these waveguides. Single mode waveguides of the type developed in this work tend to have an effective index of refraction  $n_e$ , between 2.5 and 3.0. Assuming operation at 1550 nm and taking the least favorable end of that range, the quarter wavelength is

$$\frac{\lambda_0}{4n_e} = 0.129 \mu\text{m}$$

Thus, it's no accident that the technology became viable with the advent of 0.13 μm CMOS processing. Once the technology node was identified for integration into CMOS, it was important to carefully consider the transistor technology with which integration was to occur.

SOI transistors come in two basic flavors, depending on how far the gate depletion region extends into the body of the transistor. Transistors where the depletion region does not fully deplete the body of charge are called 'partially depleted', and their counterparts are called 'fully depleted transistors'[7]. The selection of which transistor to use is of critical importance to the optical properties of the technology because of the strong dependence on the thickness of the film employed. All SOI transistors have a distinct advantage when it comes to parasitic effects, such as parasitic capacitances associated with the body of the transistors, although fully depleted devices are slightly superior. Since the substrate has been replaced with an oxide, parasitic capacitance associated with the junctions in the substrate is greatly reduced.



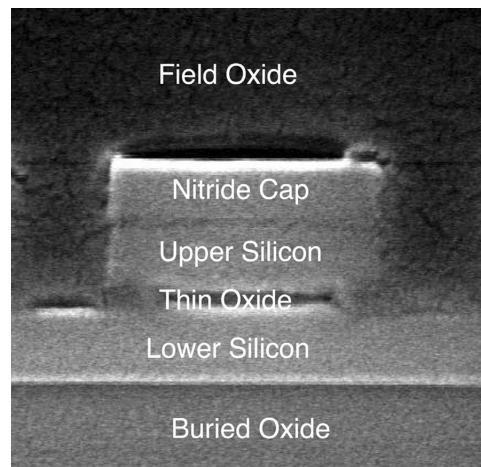
**Figure 8.3** Cross-section of a double-stack wafer. A thin oxide breaks the top silicon electrically, but maintains single mode operation optically

### 8.3.1.2 Integration with Fully Depleted Transistor Technology

Fully depleted transistors are generally considered superior for digital applications, but at the voltage levels in 0.13  $\mu\text{m}$  CMOS, the films required are approximately 50 nm thick. For light of 1550 nm, this means that waveguides constructed in this technology would be well below the single-mode cut-off thickness of  $\sim$ 250 nm.

As a result, the optical mode would extend well outside of the silicon film, requiring a very thick cladding on top and bottom of the film, and detrimentally impacting the bend radius of the waveguide. Additionally, since the most relevant electro-optic effect available in Si photonics is the interaction with free carriers in the semiconductor, a large overlap of the optical mode with Si is required to achieve compact, high-efficiency modulators. To optimize independently the electrical and optical requirements, a double-stack SOI wafer such that depicted in Figure 8.3 can be used as the substrate.

In a double-stack SOI wafer, the active Si film has been divided into two layers, the top layer being the appropriate thickness for full or partially depleted transistor body, and the combined thickness of the two layers being appropriate for waveguiding. This approach has been successfully implemented for a partially depleted SOI process requiring an active film thickness of  $\sim$ 100 nm. A waveguide constructed by etching the top film is shown in Figure 8.4.



**Figure 8.4** SEM of a double-stack wafer etched to form a waveguide

Though allowing the use of standard SOI transistors, the main drawback of this approach is the cost and development complexity associated with the starting wafer.

### 8.3.1.3 Partially Depleted Transistor Technology

Partially depleted (PD) SOI transistors are typically constructed in SOI films between 70 and 250 nm thick. These transistors will have charge underneath the channel when the transistor is operating, creating nonideal transistor behavior, such as the ‘kink’ effect. For digital circuitry, this effect is tolerable, and perhaps even advantageous in certain situations, but for analog and RF circuitry, the kink poses a design problem. A common approach used to address this concern is to integrate a body tie to the layout of each individual transistor. This makes the layout slightly larger, and poses limitations on the allowed gate width of high-frequency transistors. Since many optical components require high-performance analog and RF circuitry, consideration of this effect must occur during the selection of the CMOS process, and the wafer chosen for implementation.

However, these film thicknesses are more in line with the requirements of a silicon waveguide, and as such, front-end integration of silicon photonics is achievable with a conventional single-layer starting material. In this approach, the proper silicon film thickness is chosen for waveguiding operation, and the SOI transistors are constructed in this film. There is some impact on the transistor performance, as a result of this optimization.

### 8.3.2 CMOS Process Selection

Silicon manufacturing processes have evolved over the last two decades due to the growing application fields, driving transistor performance needs, in particular in the mixed-signal and analog areas. Silicon-only transistors in CMOS processes have long dominated the industry when it came to digital applications, while more exotic materials, such as III–V semiconductors (GaAs) were used for the rest of the application field.

With the constant improvement in manufacturing processes and ever-shrinking device sizes, however, silicon transistors have quickly reached levels of performance (in particular high-frequency operation) that make them suitable for use in fields such as communication (cell-phone chipsets) where III–Vs used to reign supreme. Additionally, recent development in introducing SiGe alloys into standard CMOS manufacturing processes have helped address some of the traditional issues with simple CMOS transistors, such as  $1/f$  and broadband noise, while retaining the capability of large integration of digital and analog circuitry on the same chip.

Discussing CMOS photonics integration then first brings to question which silicon-based technology platform to use. This depends entirely on the application targeted with the integrated optoelectronic product. In the traditional application field of optical transceivers for high-data rates, for instance, the performance is primarily driven by the needs of the high-speed analog circuits (TIAs, LAs, CDRs, laser/modulator drivers, etc.). For speeds of 10 Gb/s and beyond, submicrometer CMOS processes may not be the appropriate technology due to the noise of the transistors and the low voltages available.

SOI CMOS platforms are available commercially (Freescale, IBM) currently and are in fact used in many of today’s high-end CPUs (AMD, PowerPCs). In that context, they benefit from aggressive scaling of minimum dimensions, leading to ever-increasing performance of transistors. Because of their commercial availability and well-documented design flows, many

designers have demonstrated most of the building block circuits for 10 Gb/s optical links with satisfactory performance, typically in the 180 nm and shorter nodes [8–11].

### 8.3.3 *Passive CMOS Waveguide*

A silicon waveguide, upon integration into the front-end of a CMOS process, looks quite different than a simple strip of etched silicon. There are a variety of additional dielectrics added to the vicinity of the waveguide, which impact performance in one way or another. The makeup of these dielectric films is highly dependent on the process of choice, but will typically entail silicon nitrides, silicon dioxides (both thermal and CVD) in addition to low- $k$  spacer dielectric layers, and in the future, high- $k$  gate materials. It is important that all materials within vicinity of the optical waveguide be carefully characterized, and a complete waveguide model be constructed. Also, this device will be the baseline design for the far more interesting active waveguide, discussed in the following section.

### 8.3.4 *Active CMOS Waveguide*

Previous chapters in this book have presented classical integrated optoelectronic devices that can be built using a silicon waveguiding core. These are well known overall and have been discussed at length in the literature. However, the reported designs of these devices are typically based on restricted assumptions of ‘CMOS compatibility’, mainly the use of an SOI substrate, etching and implanting of silicon, and simplified metallization.

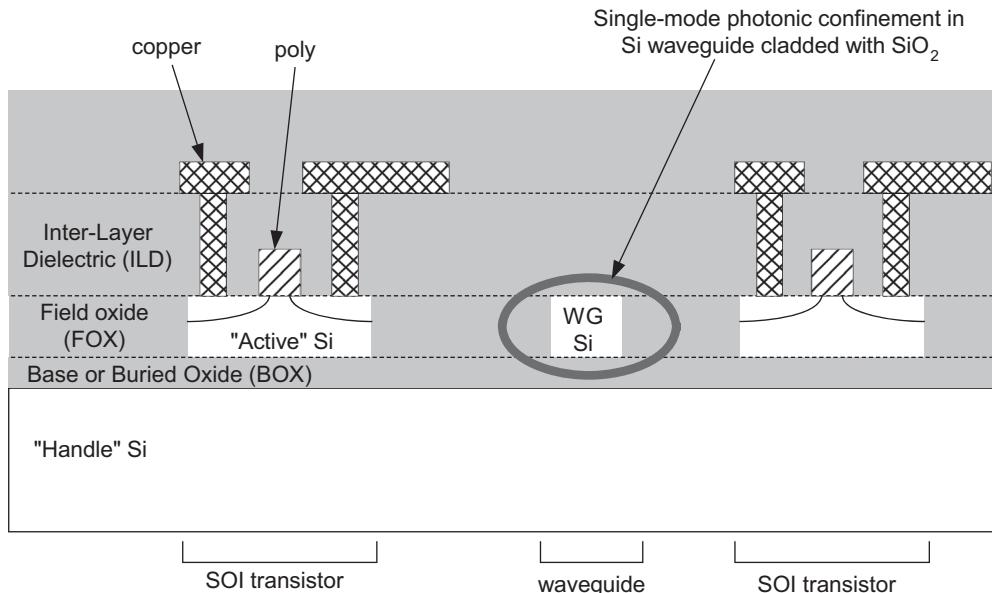
A true CMOS process, however, brings into play many other parameters into the performance of the optical devices. The optical mode in the waveguide devices overlaps not only with the side walls that have been etched and passivated (for instance with a thermal oxidation/strip/clean cycle), but also with many other dielectrics beyond simple thermal or CVD-grown oxides, for instance low- $k$  back-end dielectrics or front-end dielectrics (gate oxides, spacers, etc.). Additionally, the optical devices often have to be protected from unwanted implants, and have to rely ideally on the same contacting scheme as for transistors.

Figure 8.5 shows a typical cross-section of an opto-electronic device in a typical CMOS process environment at the end of line. We have chosen here a  $p-i-n$  phase/amplitude modulator as a representative example to discuss integration issues.

In this instance, the waveguide has been etched using a standard silicon etch, and was subjected to the same sidewall treatment as the body of the SOI transistors, which is useful for controlling and reducing optical loss and random propagation phase errors.

The waveguiding section is surrounded on the sides by field oxide, usually CVD-grown. The entire wafer at this stage will usually go through a polishing step to provide planarity for subsequent process steps. The introduction of silicon optical devices must avoid changing the density of silicon features for which the polishing process was tuned and thus permit the same process latitude. Additionally, the polishing step itself must not degrade the performance of the optical devices.

The gate process module usually involves deposition, patterning, and etching of polysilicon, followed by deposition of various dielectrics on and around the gate that are used to create spacers for subsequent implant steps. Of course, one must make sure that no poly-silicon is left close to the waveguide in any amount, due to the large propagation loss it would create.



**Figure 8.5** Cross-section of an opto-electronic integrated circuit including a *p-i-n* phase/amplitude modulator (left) a waveguide (center) and a transistor (right)

These dielectrics, depending on the process, may be left over the top of the waveguiding region, thus definitely affecting at the very least the mode shape and its effective and group indices, and potentially also altering its propagation loss and creating additional sources of random variations in the optical propagation properties. These dielectrics have to be characterized for their optical properties and taken into account in any modeling work.

The entire contact module (contact to S/D silicon and G poly) needs to be well understood and analyzed. Typically one would prefer not to modify anything, from using the S/D implants as contact implants to the optoelectronic devices, to the salicidation process encountered in most CMOS flows, to the actual metallization of the contact plugs.

Implant design of the custom implants becomes critical to ensure that they overlap well enough with the S/D implants, and provide low-resistance paths to the optoelectronic regions.

Salicidation is the process most used to create ohmic contact between the contact metal plug and the silicon/polysilicon. It relies on the interaction of the exposed silicon/polysilicon surface with, for instance, cobalt, titanium, or nickel. Needless to say, the presence of any of these silicides near the optical mode would create large propagation losses. Thus, the waveguiding regions need to be protected from the salicidation process. Most CMOS processes offer an unsalicidated poly resistor, which is achieved by depositing and patterning a dielectric over a poly region, preventing the reaction of the metals with exposed poly. This same mask can be used to pattern a protective dielectric cap over the waveguide. This dielectric layer may or may not be part of the spacer layers of the gate module, depending on the integration approach. They may or may not also block the S/D implants, so particular care must be taken to design this layer properly.

Once the very close environment of the waveguide has been analyzed and designed around properly, still the back-end of line of the CMOS process, with its various dielectrics and metals, needs to be taken into account.

### 8.3.5 Implants and Activation

If the implants needed for optimal performance of the optoelectronic devices are not identical to already existing implant steps in the CMOS process (i.e., wells, source/drain, extensions), then custom implant steps must be inserted. Depending on the energy/dose of these implants, only certain types and thicknesses of photoresists can be used to ensure nonpenetration in undesirable regions. This can restrict in turn the minimum feature sizes and alignment accuracies of the lithography. If the device design does not rely on self-aligned implants (transistors typically use self-alignments and thus can use relaxed tolerances), this can turn into a problem, for instance transforming a desired abrupt lateral junction into a graded junction, or bringing high-density implant regions too close to the optical mode, thus creating high propagation loss, as could happen in our  $p-i-n$  diode in Figure 8.5.

Another critical aspect of ion implantation for optoelectronic device design, that can easily be overlooked, is the angle of implantation. To avoid a channeling effect, especially for low-dose implants, most CMOS processes now rely on off-normal implantation. Again, in the case of an abrupt junction design, shadowing effects due to the photoresist thickness could lead to variations in performance between nominally identical diodes laid out in different orientations.

The insertion of the custom implants in the CMOS flow also need to be considered with great care, because all implant usually require a thermal activation step, and these need to be minimized in order not to modify the behavior of the transistors. A perhaps clever way to introduce implants is to add them in the same area of the process flow as some of the existing implants, especially if they are close in energy. Then the same activation step can be used for multiple implants at the same time.

### 8.3.6 Edge-coupling vs Surface-coupling and Reliability

The typical dimensions of the waveguides discussed here is in the hundreds of nanometers in cross-section. While there are many advantages derived from such small cross-sections and high confinement, at the same time this design choice creates a major problem: coupling light into and out of these devices efficiently. Desirable mode sizes for external optical devices range in the many micrometers in diameter (laser modes, single- and multi-mode fibers). Direct butt-coupling from a small silicon photonic waveguide to such an external optical device will create huge insertion losses, making the silicon photonic system completely useless.

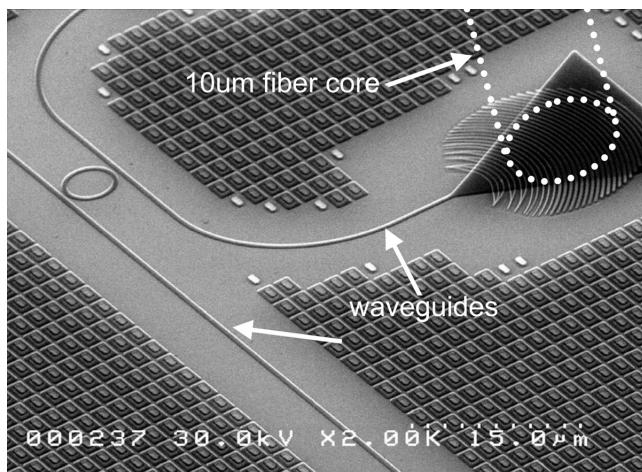
Additionally, butt-coupling typically requires cleaving/dicing and polishing of facets, and this presents two major issues. The first one is the inability to test photonic devices or systems at the wafer level. Wafer-scale testing inherently allows for lower cost manufacturing by allowing early testing and screening of defective parts rather than having to go through an entire assembly process to figure out which parts are functional and to specifications. A prime example of this advantage is in the comparison of a VCSEL versus a communications grade edge-emitter. Despite being more complex and costly to grow, the VCSEL is substantially lower cost, which can be primarily attributed to the ability to perform wafer-scale testing.

The second major issue is that SOI CMOS chips typically require an edge-seal/anti-crack ring around the edge of the die. The first purpose of this seal is to prevent contamination of the circuits from mobile ions (e.g. potassium) that can drastically change transistor properties over time as they travel through the buried oxide and lodge themselves into the gate oxide. The second purpose is to protect the circuits from cracks that could happen during wafer dicing and prevent these cracks from propagating into the circuits. Any butt-coupling scheme for a silicon photonics system would require an opening in such an edge-seal which is usually built of a set of metal ‘walls’, going all the way from the active to the surface of the chip.

The only way to address these issues is through surface-normal coupling into and out of the chip. For that purpose, as is well-known in the literature, a diffractive element can be used, typically a holographic lens. As mentioned previously, the 130 nm node is where feature sizes achievable through high-volume optical lithography and etching have become small enough to allow the design of these types of devices using standard lithography tools and processes.

The availability of optical lithography allows the designers of holographic lenses to tailor the shape of the grating to match in principle any optical mode (size, shape, and phase front) by creating a nonuniform grating. Such designs have been developed for matching to single-mode fibers (in the 10  $\mu\text{m}$  mode-size range) and smaller laser modes (in the 2–3  $\mu\text{m}$  mode-size range) with efficiencies as high as  $-1 \text{ dB}$  and with a bandwidth wide enough to cover most of a relevant wavelength band, as for instance the communication C-band (Figure 8.6).

There is one critical additional consideration when designing a holographic lens into a CMOS process. The back-end of line also has a direct impact on the performance of holographic lenses if they are designed to emit and receive light through the back-end rather than through the substrate. Figure 8.6 shows an example of a cross-section of a holographic lens with the light having to propagate through the back-end of the chip. As mentioned above, there are many layers present in the back-end, and in fact more importantly many interfaces between these layers of different refractive indices. This can create a random interference pattern, since the thickness of these layers will vary across wafers, and from wafer to wafer. This interference can dramatically reduce the overall efficiency of coupling. An obvious important



**Figure 8.6** Oblique view of a holographic lens device

side-note here is also that no metals of any kind should be present in the area of light travel above the holographic lens.

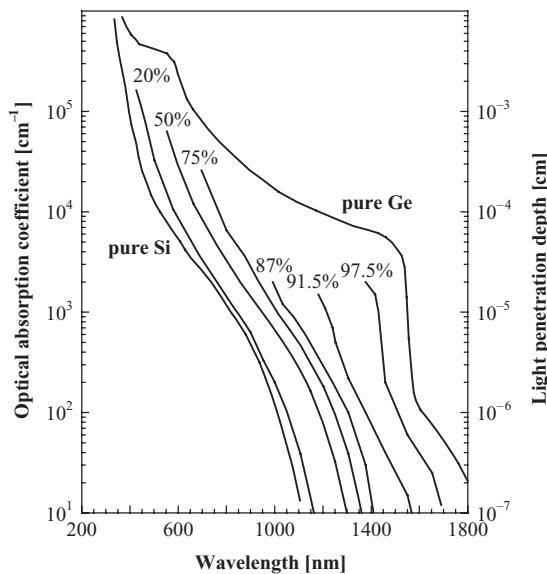
A simple way to overcome this effect would be to locally remove part of or all of the back-end dielectrics, and potentially refill the hole thus created with a single dielectric, for instance a CVD-grown oxide. However, performing this operation has the potential to create additional issues due to the fact that it punches through the passivation layer that is usually present as the last or next-to-last layer in the back-end. Leaving a fully open, uncapped, unpassivated area can lead to issues such as mobile ion contamination.

In practice, it is possible to achieve good coupling efficiency through the back-end film stack by considering the back-end to be a thin-film filter. Selection and design of metal and via thicknesses will modify the filter characteristics, and it is possible to make a selection that does not strongly impact transmission through the back-end.

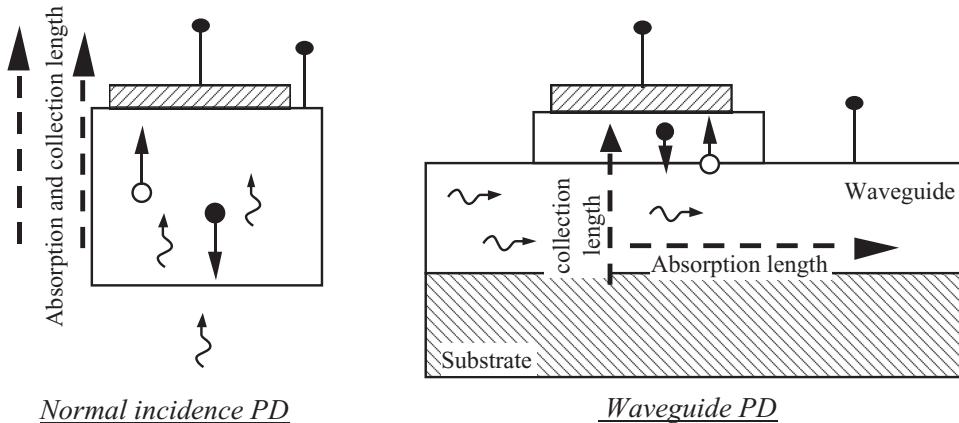
### 8.3.7 Germanium Integration

SiGe and Ge are being considered for integration in the CMOS process thanks to the higher mobility of carriers compared with Si [12]. In other approaches, a SiGe or Ge substrate is used for the growth of strained Si [13]. SiGe recessed source/drain regions have been proposed as well to induce compressive strain in the channel, thus improving transport properties [14]. Beyond CMOS, SiGe alloys having Ge concentration in the 10–20% range are used in the base of heterojunction bipolar transistors (HBTs) where the graded alloy creates a quasi-electric field that shortens the transit time of minority carriers [15].

SiGe and Ge have been considered for their optical properties, too [16]. Thanks to the smaller bandgap, compared with Si, they enable photodetection at the wavelengths used in optical communications. Ge, in particular, shows high absorption coefficient at both 1.3 and 1.55  $\mu\text{m}$  (Figure 8.7).



**Figure 8.7** Absorption coefficient and corresponding light penetration depth for SiGe alloys



**Figure 8.8** Light absorption and photocarrier collection in normal incidence (left) and waveguide (right) photodetectors

Normal-incidence Ge-on-Si photodetectors with high efficiency and speed have been demonstrated, in recent years [17–19], however, the integration of these devices in a typical CMOS process is difficult due to thermal budget and topography constraints. In particular, the Ge film thickness required to efficiently absorb light at  $1.55\text{ }\mu\text{m}$  (more than  $2\text{ }\mu\text{m}$ ), hardly fit in the CMOS landscape where the typical thickness of poly gate lines is more than one order of magnitude smaller.

Waveguide photodetectors (WPDs) solve this issue by having the absorption path parallel to Si surface (Figure 8.8). An additional advantage of waveguide configuration is that the usual trade-off between speed (transit time) and efficiency (absorption length) of normal incidence detector is eliminated by having photocarriers traveling perpendicular to photons. This allows one to independently optimize detector length (absorption) and thickness/width (transit time).

Waveguide detectors, of course, rely on the availability of efficient coupling of light from the external sources to the Si waveguide, such as the holographic lens discussed in the first part of this chapter.

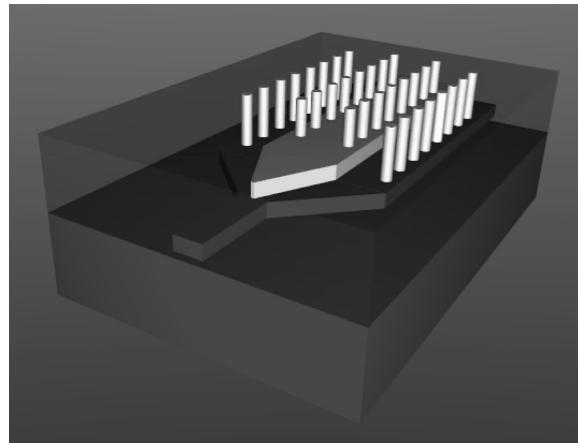
Depending on the electrode configuration, Ge-on-Si WPDs can be homojunction or heterojunction devices. The former has both anode and cathode on the Ge film, the latter has one electrode on Ge and the other in Si. Due to valence and conduction band alignment at the heterojunction, best results are obtained when Ge is the anode and Si the cathode.

Figure 8.9 shows a 3D view of a Ge on Si heterojunction waveguide photodetector.

### 8.3.7.1 Epitaxy

Ge WPD body is created by the epitaxial growth of a patterned Ge film on the Si surface. The latter can be accomplished by depositing a blanket film followed by a lithographic patterning step. A more effective approach uses the selectivity of Ge CVD epitaxy, under appropriate growth conditions. A silicon oxide patterned layer can be used as a hard mask to this extent.

Ge epitaxy on Si is affected by the large lattice mismatch (about 4%) existing between the film and the substrate. The latter causes 3D film growth in the Stransky-Krastanov mode,



**Figure 8.9** A Ge-on-SOI waveguide photodetector

unless special techniques are used to relax the film by introduction of misfit dislocations at the heterointerface. The most successful technique, to date, is that introduced in [20] which basically consists of growing a thin Ge buffer layer at low temperature (about 350°C) where insertion of misfit is favored compared with islanding, thanks, also, to the surfactant effect of hydrogen. Buffer thickness must be enough to fully relax the film (about 50 nm). Once the film is fully relaxed, the growth temperature can be increased and growth continued as in the homoepitaxial case.

Figure 8.10 shows a TEM cross-section of a Ge-on-Si epitaxial film grown using the aforementioned technique on a 8" Si wafer: misfits at the interface are evidenced. The technique can be applied to selective growth, as well.

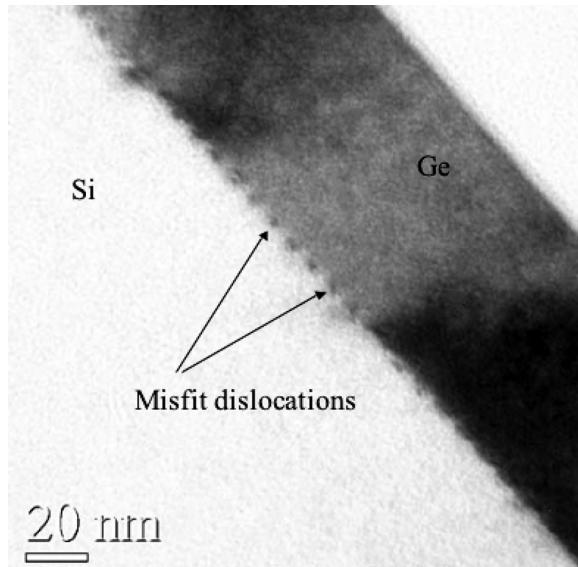
### 8.3.7.2 Integration Issues

When introducing the Ge module into an existing CMOS process, a number of constraints and interactions must be considered. These can be broadly grouped into three categories: thermal budget constraints, topography limitations and contamination concerns. Moreover, all the process steps used in the Ge module must be ‘manufacturable’, meaning that they have to stay within specific limits in terms of duration, use of resources, and reproducibility.

The thermal budget of Ge epitaxy is mostly driven by the pre-epi cleaning step required to create an atomically clean Si surface. This is usually accomplished with a two-step process: wet clean in buffered HF followed by an *in situ* bake in hydrogen atmosphere. The latter implies the use of temperature in the 800°C range for an extended time (compared with typical rapid thermal annealing cases). This can result in salicide agglomeration (especially in fine line modern CMOS technologies) [21] and doping diffusion.

On the other side, the high-temperature steps used for doping activation in Si can drive excessive dopants diffusion in Ge thus impairing device performance.

Planarization techniques are commonly used in modern CMOS technologies to smooth topography features, such as those created by the poly gates, and enable high-resolution



**Figure 8.10** Cross-section of an epitaxial Ge on Si film. Misfit dislocations at the interface are evidenced

lithography required for contact placement. The introduction of additional features on the Si surface must withstand the limits imposed by the capability of the planarization process which basically translates in limiting the maximum thickness to a few hundred nanometers. As discussed earlier in this chapter, this can only be accomplished by adopting the waveguide configuration for the detector.

Contamination is one of the most important concerns in semiconductor foundries: fast diffuser and electrically active species, if not properly blocked, can find their way to transistor and change their electrical characteristics, thus impairing yield and reliability of products. Fortunately, Ge is not a big concern under this aspect: it shares with Si the same group in the periodic table, thus it does not impact electrical conductivity, in addition, it is quite a slow diffuser. Moreover, the electronic industry has developed a good familiarity with Ge thanks to former works on HBTs and, more recently on strained CMOS. Nonetheless, since Ge is much more susceptible to be etched by chemical solutions that are perfectly safe for Si (such as SC1, one of the standard RCA baths commonly used to clean Si surface), care has to be taken to avoid exposure of Ge surface to these baths during processing.

### 8.3.7.3 Photodetector Devices

The design of a high-speed/high-efficiency Ge-on-Si WPD has to deal with a number of trade-offs and constraints.

In order to obtain high responsivity, the input light must be efficiently coupled from the Si waveguide to the detector body. This can be done in different ways, ranging from pure end fire (sometimes using a short taper), to evanescent coupling. Long, adiabatic, tapers are not suitable due to the high absorption coefficient of Ge. In both cases, the higher refractive index of Ge, compared with that of Si, facilitates the localization of the optical mode in the Ge waveguide, thus maximizing the overlap with the absorption medium.

Once in the detector body, light is absorbed on a length which shows a large dependency on wavelength and temperature. At room temperature, indeed, the onset of direct transitions in the band structure of Ge happens at 0.8 eV (1550 nm). This edge moves to longer wavelengths as the temperature increases. Ge on Si is normally slightly compressively strained due to the difference in thermal expansion coefficients of film (Ge) and substrate (Si) [22]. The strain shifts the edge to a longer wavelength, thus improving absorption in the infrared. Depending on the limited overlap of the optical mode with Ge, about 10–20  $\mu\text{m}$  of propagation length is required to efficiently absorb light at 1550 nm.

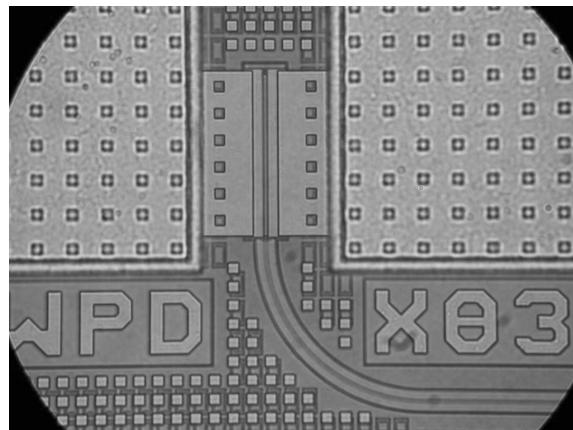
Along propagation, light can be scattered out from the waveguide by the presence of discontinuities such as film nonuniformity and/or metal contacts. Contacts are unavoidable; therefore it is very important to place them in order to minimize their overlap with the optical mode, while maintaining a low-resistance connection to the junction.

Once the photons are absorbed, photocarriers are generated and drift under the effect of the built-in/applied field towards the contacts. In a homojunction WPD, carriers drift is parallel to the heterojunction and speed is limited by transit time, i.e. by cathode-to-anode implant spacing. Since the electrode area is very small, capacitance is, normally, not a limiting factor for the speed of these devices. In heterojunction WPDs, where the current flow is perpendicular to the heterojunction, transit time is very short due to the very thin Ge layer, and capacitance effects usually dominate the frequency response. In any case, Ge WPDs show much smaller capacitance than discrete photodiodes: this characteristic can be exploited to improve the sensitivity of high-speed receivers.

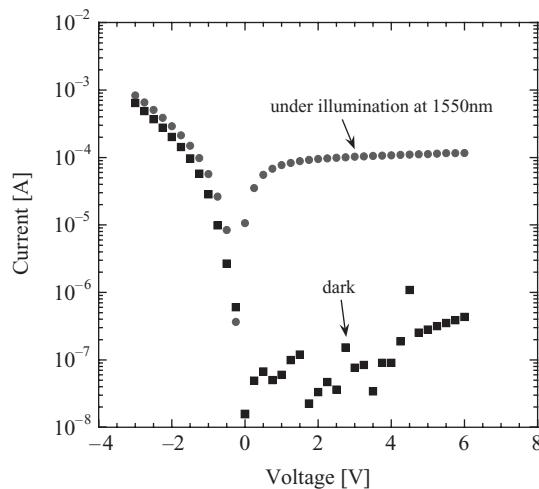
Figure 8.11 shows a Ge WPD integrated in Freescale's HIP7\_SOI CMOS process. The Ge waveguide, in the center of the device, is partially hidden by the metal stripes and is 30  $\mu\text{m}$  long. Light is coupled through the Si waveguide shown in the bottom part of the figure.

The current–voltage characteristics of a Ge on Si heterojunction WPD in the dark and under illumination at 1550 nm are shown in Figure 8.12. A responsivity of 0.7 A/W at 1550 nm has been measured.

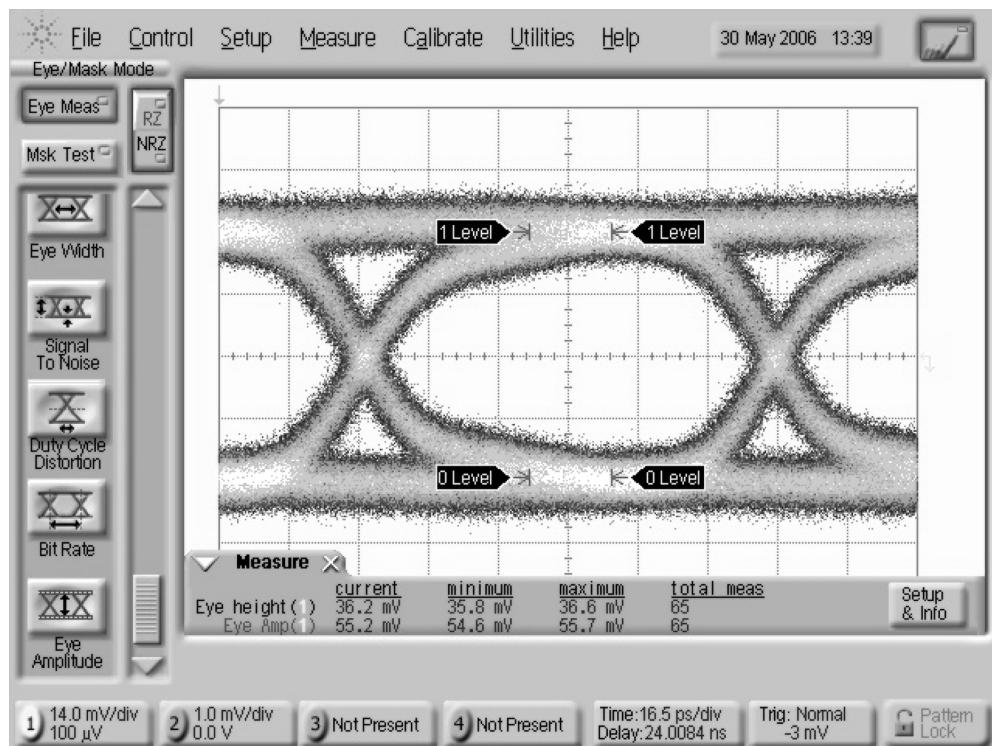
Thanks to the very small capacitance and short transit time, heterojunction WPD can be very fast. Figure 8.13 shows the eye diagram at 10 Gbs measured by connecting the photodiode to a high-speed sampling scope (no TIA) and exciting it at 1550 nm.



**Figure 8.11** Optical microscope picture of a Ge on Si waveguide photodetector



**Figure 8.12** Current–voltage characteristic of a heterojunction WPD in the dark and under illumination  
Light at 1550 nm is coupled through the Si waveguide



**Figure 8.13** Eye diagram of an avalanche waveguide photodetector when excited with a 10 Gbs NRZ  
PRBS optical signal

## 8.4 State of the Art

### 8.4.1 An Integrated 40 Gbps Wavelength Division Multiplexed Transceiver

One of the compelling aspects of CMOS integration is the potential to scale by placing many transceivers on a single die. Since this can be done with a relatively low impact on the cost of the silicon die, it is expected that such scaling will dramatically impact the cost for a given amount of bandwidth. Also, since the cost of an optical fiber can be high for transmission over a considerable distance, use of WDM and silicon photonics allows a single CMOS die and a single strand of fiber to carry a large amount of bandwidth inexpensively. This is the vision behind the DARPA EPIC program, which has sponsored the development of a 40 Gb transceiver at Luxtera, and eventually a 100 Gb transceiver.

In this section, we will describe a four wavelength WDM transceiver on a single chip operating at 10 Gbps on each wavelength, for an aggregate bandwidth of 40 Gb in each direction down a single fiber pair.

A block diagram of the transceiver is shown in Figure 8.14. Each transmitter receives continuous wave (CW) light from an off-chip laser, which is coupled into the silicon using a surface relief holographic lens (HL) [23]. The HL is used to couple light both in and out of the die. The four wavelengths are separated by 200 GHz and are located in the C-band, but they are not locked to the telecommunications grid. Mach-Zehnder interferometer (MZI) modulators located in the die perform amplitude modulation of the light according to the electrical signal input to each modulator driver. The modulation is performed with a 2.5 V swing, which induces free-carrier phase modulation in the arms of the MZI as appropriate to create either constructive or destructive interference. The outputs of the modulators are then fed into a two-stage interleaver structure, which multiplexes the signal into a single output waveguide. The light is then transmitted down a fiber via a HL structure.

When the light arrives at the receiver, it is coupled into the chip via a HL and demultiplexed using another two-stage interleaver. After the interleaver, the light is coupled to a receive photodetector. There are two types of photodetector which have been demonstrated. The first type is a traditional InP Pin photodetector that is flip-chip mounted to the die in proximity to the receiver transimpedance amplifier (TIA). The light is coupled into the active area of the photodetector via a specially designed HL. Another technique is to construct a germanium photodetector directly on the receive waveguide after the demultiplexer, and adjacent to the TIA. After the TIA, the signal is further amplified by a limiting amplifier (LA) and transmitted off-chip electrically.

A similar interleaver is used at both the transmitter and receiver. Two stages of cascaded unbalanced MZI stages are employed. By tuning the phase relationship between the arms of them MZI, the filter is tunable to any wavelength plan. The tuning of the arms of the interleaver is achieved by heating the arms appropriately, and utilizing the change in Si refractive index as a function of temperature. The interleaver was selected due to its low insertion loss for a four channel system. The measured transfer function of the multiplexer interleaver is shown in Figure 8.15. One can observe a 200 GHz spacing and cross-talk suppression in excess of 20 dB. The demultiplexer interleaver is constructed and operated in a similar manner. The optical spectrum of the transmitter output is shown in Figure 8.16, illustrating all four wavelengths in the output fiber.

A schematic of the modulator driver circuit is shown in Figure 8.17. It is constructed from a three stage 1.5 V predriver followed by a cascaded differential out stage driven from a 5 V

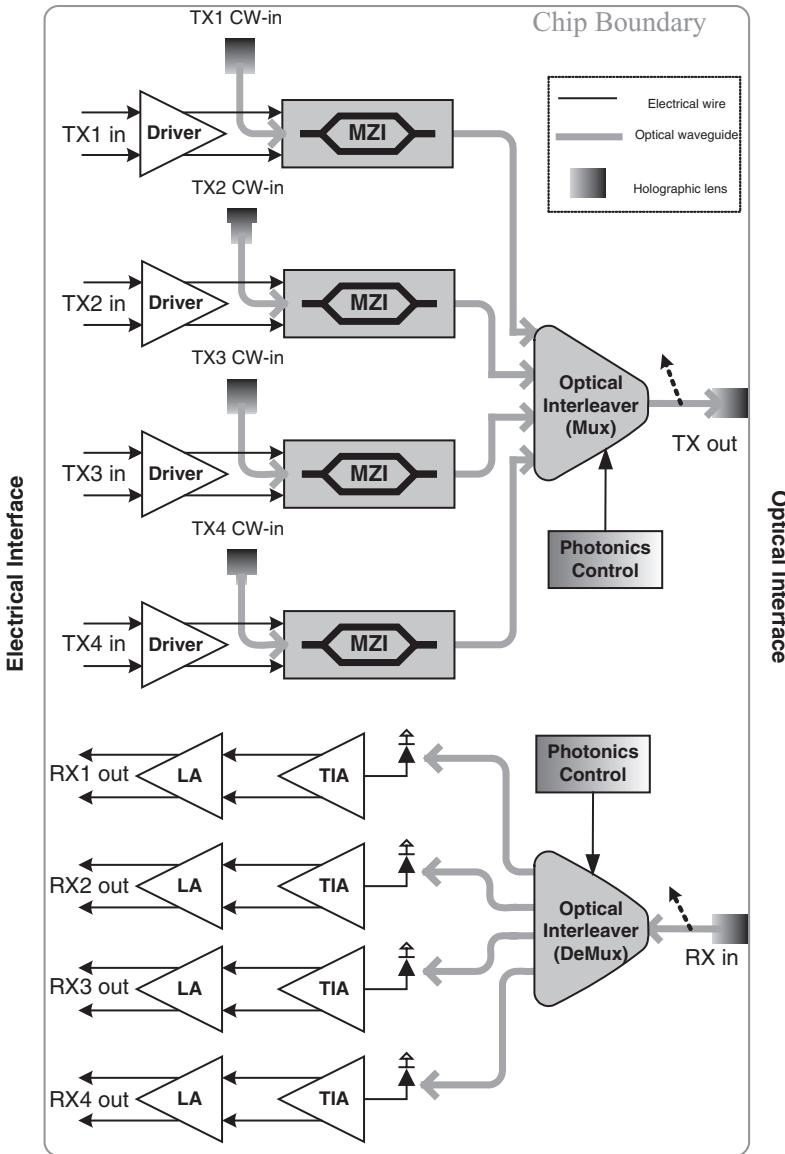
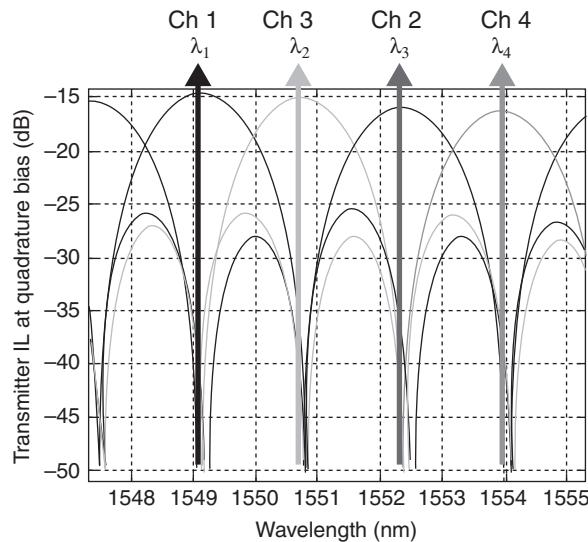
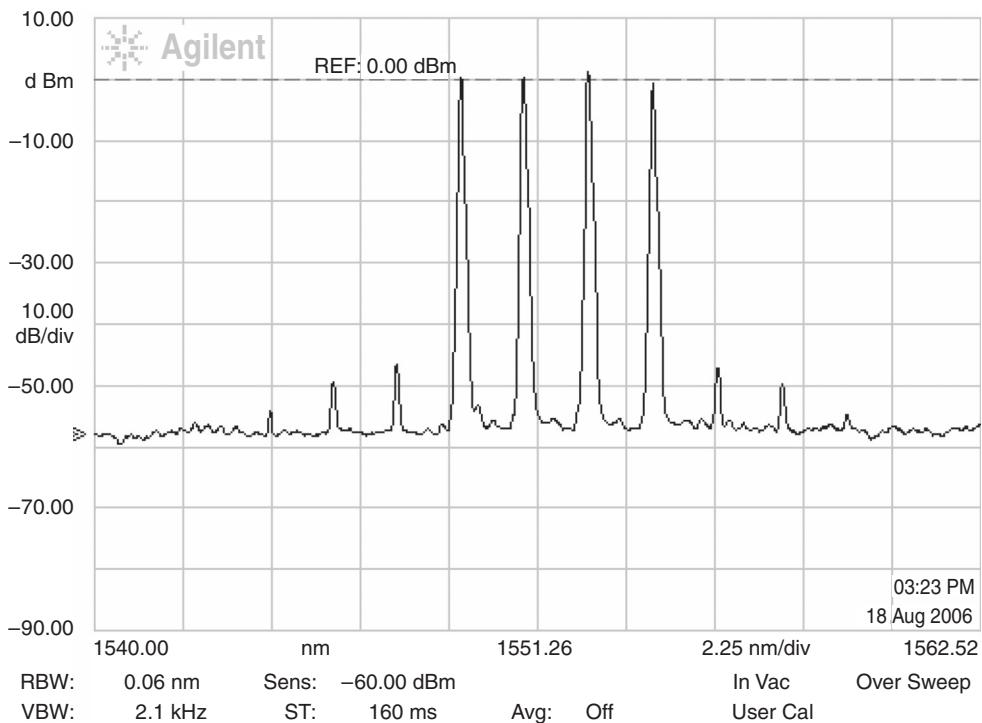


Figure 8.14 Block diagram of Transceiver

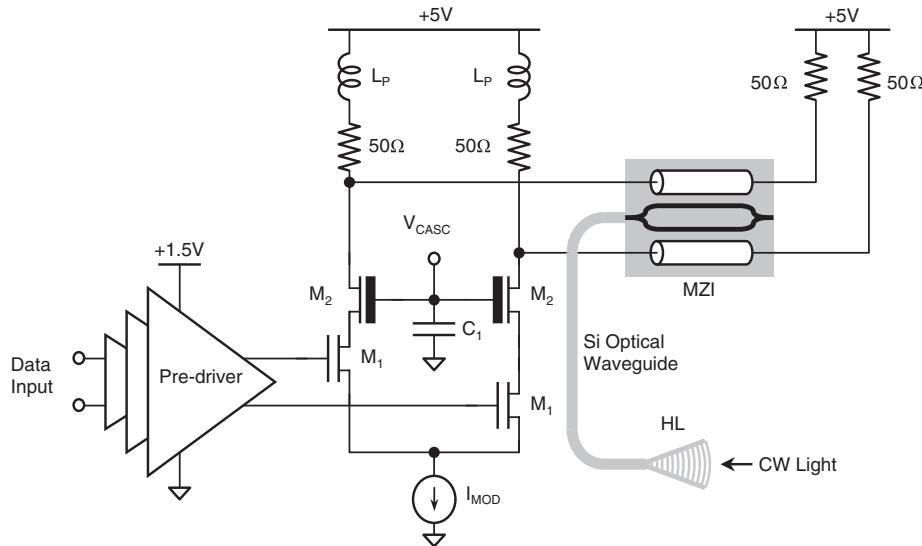
supply. A  $50\ \Omega$  waveguide connects the driver to the MZI modulator, and is terminated with  $50\ \Omega$  to prevent reflections. Achieving the 2.5 V necessary to drive the MZI is difficult with thin oxide CMOS transistors, so the final stage utilizes thick oxide transistors typically used to drive chip I/O. However, these transistors are substantially slower and require inductive peaking and a strong predriver signal in order to achieve the fast transitions required to operate



**Figure 8.15** Transfer function of a multiplexer/interleaver



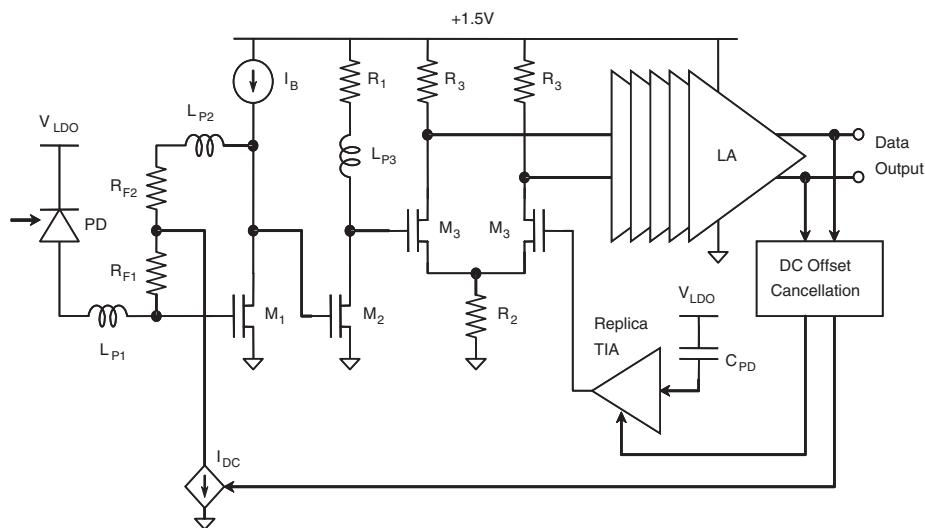
**Figure 8.16** Output spectrum of the WDM transceiver showing the four wavelengths multiplexed in the same fiber



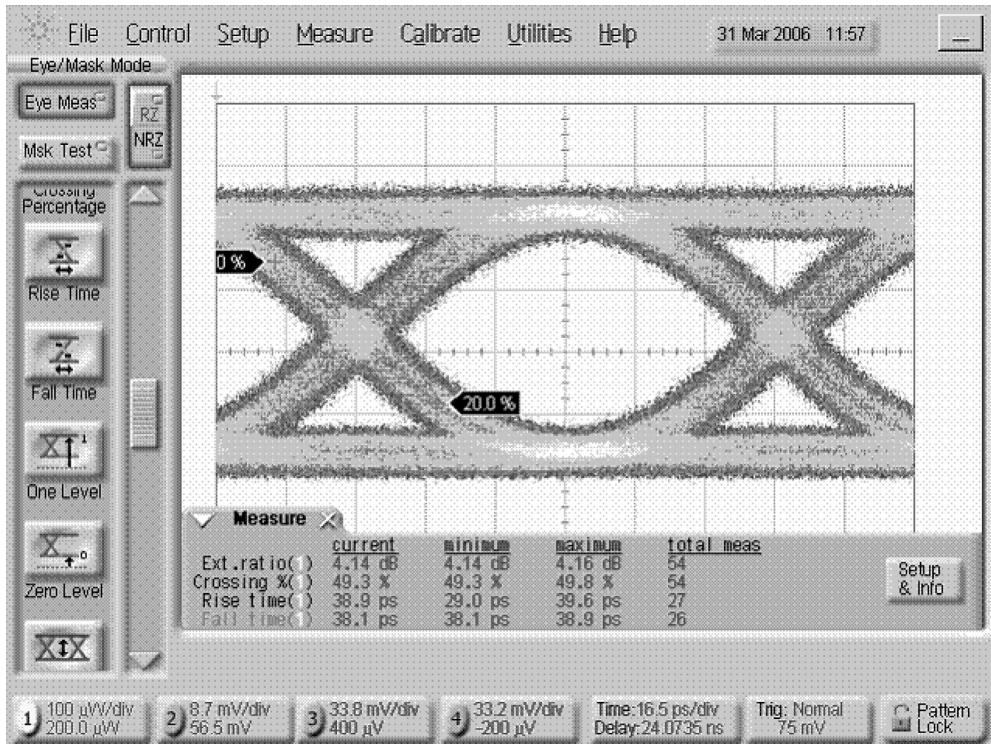
**Figure 8.17** Block diagram of the modulator driver circuitry

at 10 Gbps. Under nominal conditions, each driver consumes 575 mW. Only 50 mW of peak-to-peak input signal is required at the input to the predriver, which will result in an extinction ratio of approximately 4 dB.

Figure 8.18 illustrates the receiver front-end, as described earlier. The TIA input stage is standard, consisting of a resistive feedback structure and allowing high-speed operation



**Figure 8.18** Block diagram of the receiver front-end



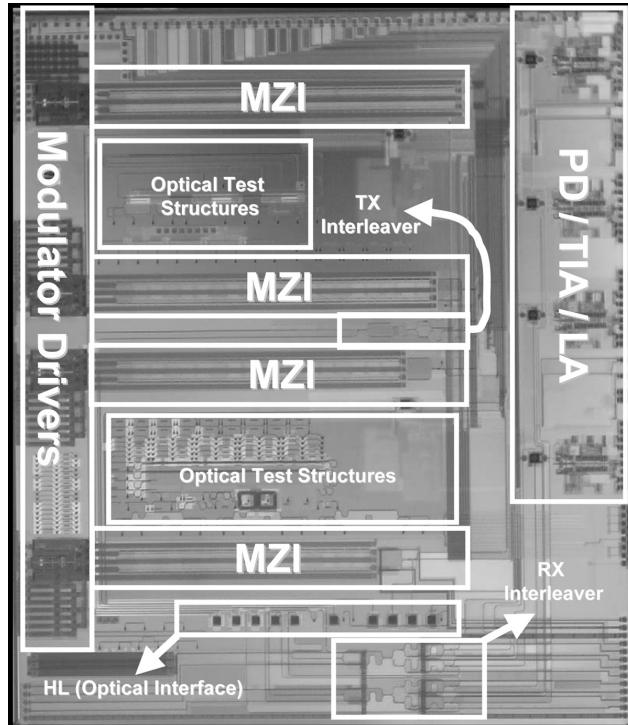
**Figure 8.19** Eye diagram at the receiver of the WDM system

[24, 25]. Subsequent stages impedance match the signal and convert it to a differential output. The limiting amplifier is differential and employs inductive peaking. A filter on the output of the LA is used to extract the DC component of the signal, which is used in an offset cancellation loop, which compensates for the offset induced by a nonperfect extinction ratio. The LA drives an output buffer with a  $50\ \Omega$  impedance.

The receiver circuitry all operates from a 1.5 V supply and consumes 120 mW. A sensitivity of better than  $-15\text{ dBm}$  average power is achieved for a BER of  $10^{-12}$  at 10Gbps, measured over 500 s. The eye diagram of the receiver is shown in Figure 8.19. This eye was taken with un-correlated data running on each of the four channels. Figure 8.20 shows a photograph of the die.

#### 8.4.2 An Integrated 10 GHz Opto-electronic RF Oscillator

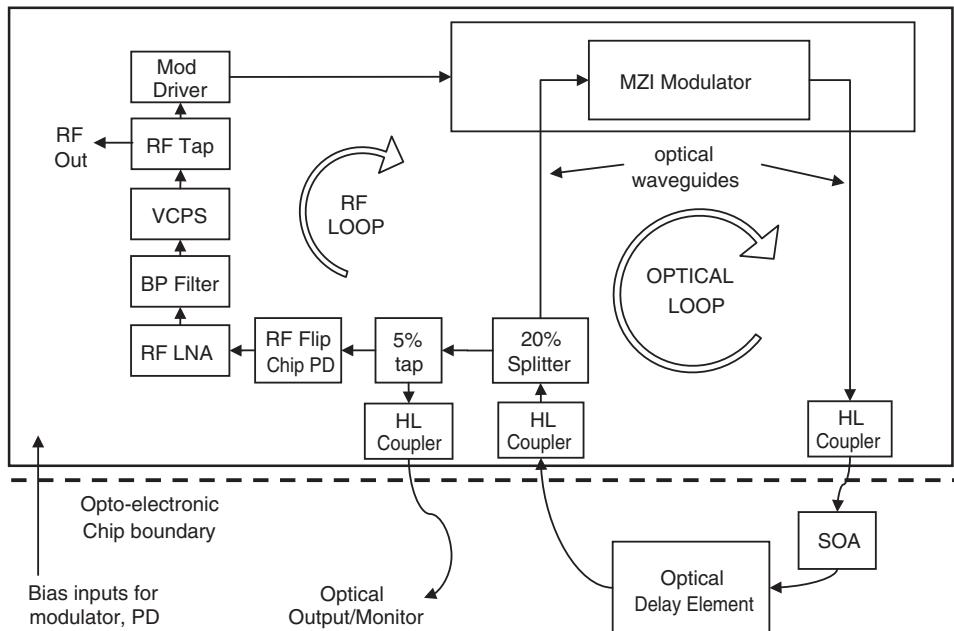
The emerging wireless communications standards at 60 and 80 GHz are challenged by the availability of high-performance, yet cost-effective oscillators. The bandwidth of a wireless communications system is heavily dependent on the phase noise of the local oscillator, meaning that spectral efficiency at these frequencies is limited to a fraction of that achieved in lower-frequency applications where local oscillators (LO) are readily available. LO performance also limits radar system performance, particularly impacting millimeter-wave phased arrays.



**Figure 8.20** Picture of the WDM system die showing the different subsystems

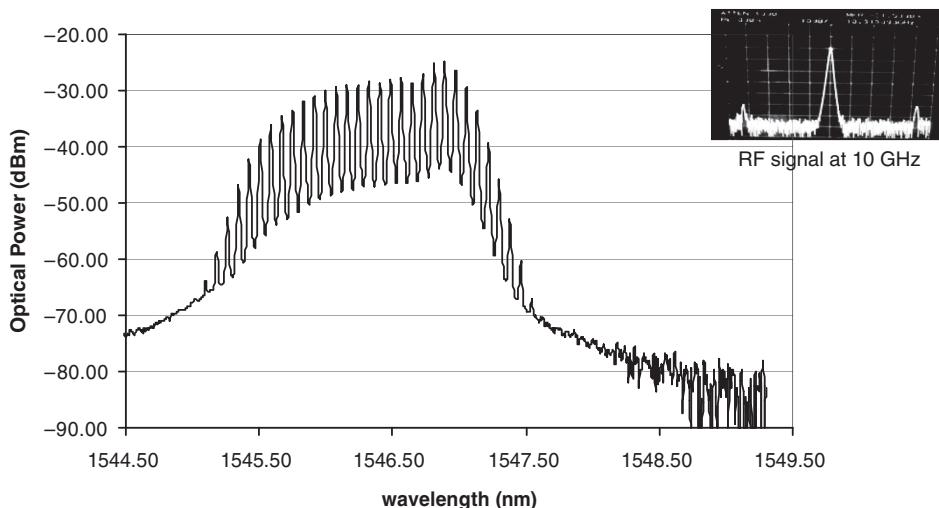
Optoelectronic oscillators (OEO) have been known to achieve better phase noise performance than traditional frequency synthesis approaches, examples of which are: phase-locked dielectric resonators, YIG, and quartz crystal oscillators [26, 27]. The improvement in phase noise achievable in an OEO is fundamentally attributed to direct frequency generation from an optical source at the frequency of interest, rather than through electronic multiplication, which also multiplies jitter. Historically, OEO systems are constructed from discrete devices, require large volume and power consumption, and are expensive. Vibration issues encountered with discrete parts also cause an issue for both military and commercial environments. Recently, this circuit has been implemented in 130 nm SOI CMOS photonics technology [28]. The result is a low-power, small-form-factor, low-phase-noise OEO.

The OEO is constructed from two coupled loops (Figure 8.21); one providing optical gain, and the other providing RF gain. The optical loop resembles a mode-locked laser with a long time delay in the optical cavity. The RF loop determines the operation frequency of the OEO and provides the RF output signal. The optical delay results in a low phase noise and high spectral purity. This delay is achieved with a long length of fiber (though a high-Q resonator could be substituted). The optical loop is constructed from a silicon Mach-Zehnder interferometer (MZI) modulator along with a semiconductor optical amplifier (SOA) and a fiber-based optical delay element (ODE). The fiber output is connected to the input of the modulator, forming a loop with more gain than loss. The silicon MZI was achieved adequate bandwidth and extinction ratio while maintaining a low insertion loss (<4 dB).

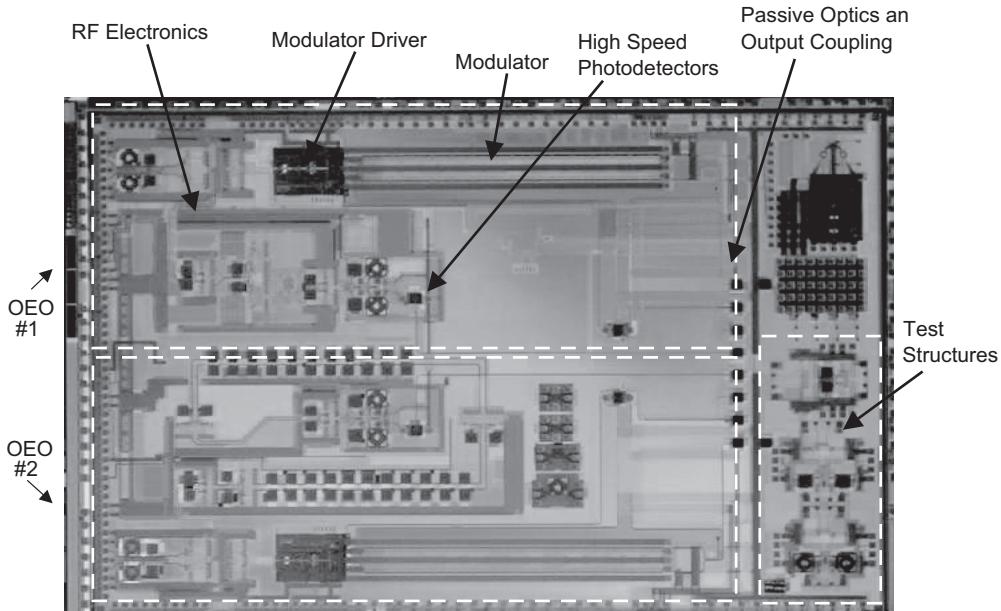


**Figure 8.21** Block diagram of the optoelectronic oscillator evidencing the two coupled loops

The optical energy is split and a portion is directed to a photodetector (PD). The electrical output of the PD feeds the RF loop. The locking of the optical modes is shown in Figure 8.22. Note that they are spaced by 10 GHz, which results in a 10 GHz RF signal at the PD. The PD output signal is amplified with a low-noise amplifier realized in CMOS, filtered with a bandpass filter (BPF) to set the operation frequency, and used as input to the optical modulator



**Figure 8.22** Optical spectrum at the output of the modulator showing mode locking operation



**Figure 8.23** Picture of the oscillator die

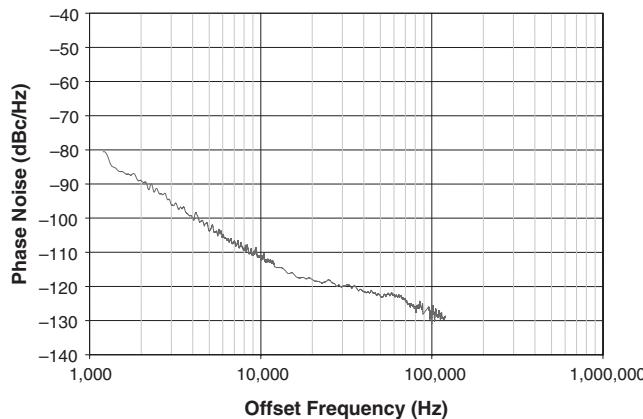
driver. Adequate phase alignment between the two closed-loop systems is achieved by use of a voltage-controlled phase shifter (VCPS). The complete circuit was designed to oscillate at 10.24 GHz, with a RF output of 0 dBm.

All RF circuitry was monolithically integrated on chip, and the reader interested in the electrical design is referred to [29, 30].

The fabricated chip is shown in Figure 8.23. There are two partial OEO circuits in this single die, which are identical, with the exception of two different VCPS designs. The chip was fiber pigtailed and packaged, and the circuit was completed with an external ODE and an SOA. Optical locking and RF oscillation at 10.2 GHz was achieved. The phase noise was  $-112 \text{ dBc/Hz}$ , measured at a 10 kHz offset (Figure 8.24). The CMOS RF electronics clearly limit the performance (separate measurements with external RF components achieved a phase noise of  $-139 \text{ dBc/Hz}$  at 10 kHz). When run open-loop, the RF phase noise was  $-125 \text{ dBc/Hz}$ , measured at a 10 kHz offset. Improved RF component design is expected to significantly improve performance.

While the savings in size and weight of the integrated portions of the OEO was dramatic, there was also a significant reduction in the power consumption of the integrated version. The discrete version contains several amplifiers which may use as much as 10–20 W, while the entire integrated RF chain was realized in less than 200 mW. In addition, the modulator driver typically uses 575 mW.

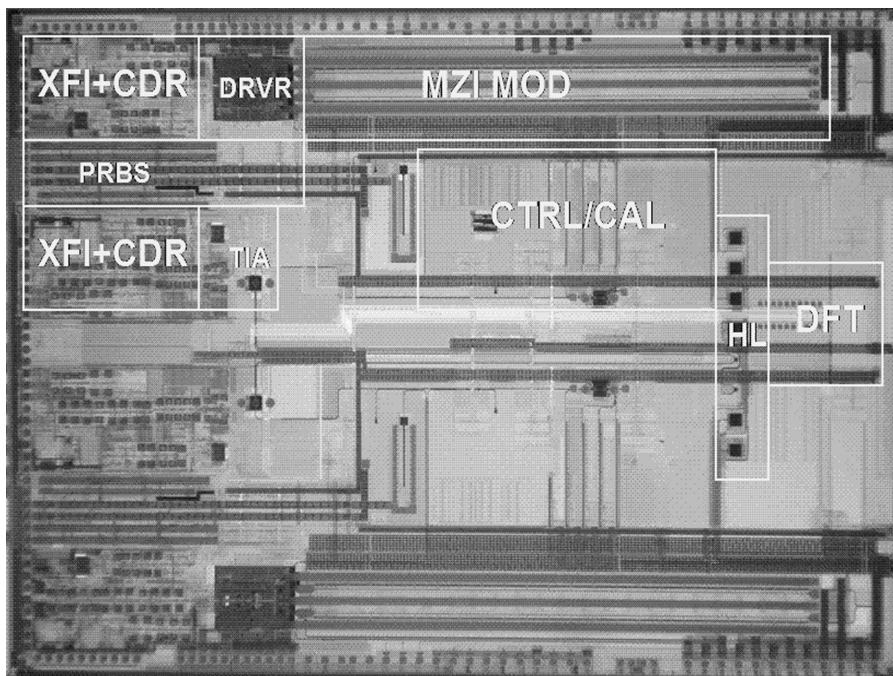
As the reader may ascertain, silicon photonics is capable of implementing high-performance systems in a fraction of the size and weight of traditional systems, and using less power. This is the first successful integration attempt of silicon photonics for a RF application, although other efforts are currently underway at MIT Lincoln Labs as part of the DARPA EPIC program.



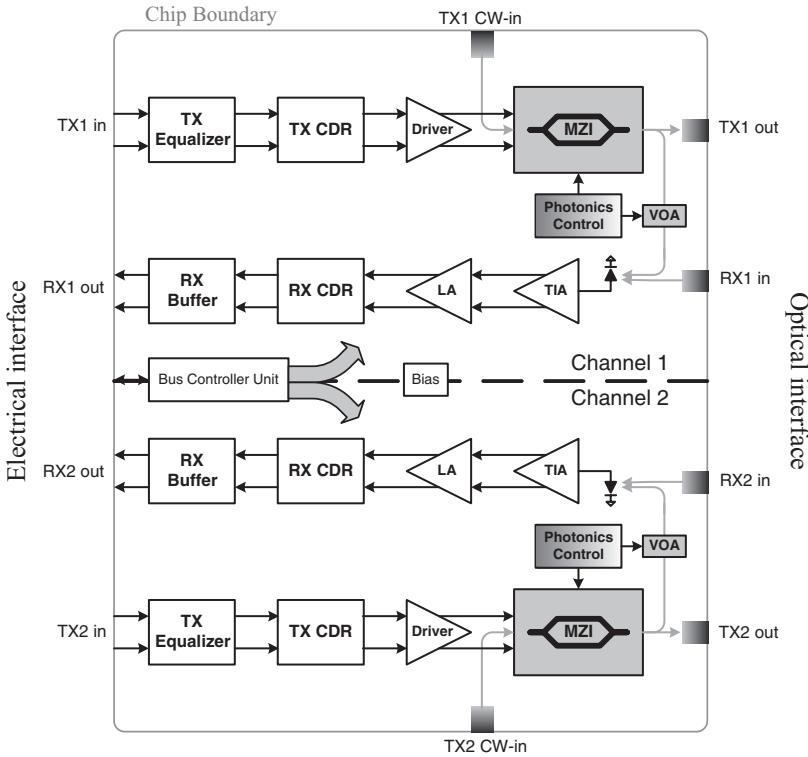
**Figure 8.24** Phase spectral noise of the opto-electronic oscillator

#### 8.4.3 Aurora

Perhaps the highest level of integration achieved to date is illustrated in the 20 Gb transceiver pictured in Figure 8.25. This die contains two near-complete XFP transceivers, each operating at 10 Gbps. These transceivers are independent and thus transmit down two separate fiber pairs. An external microprocessor and passive components accompany this die in operation.



**Figure 8.25** Picture of the Aurora die, evidencing the two 10 Gbps each transceivers integrated on the same CMOS IC



**Figure 8.26** Block diagram of the two-channel, 20 Gbps optical transceiver shown in Figure 25

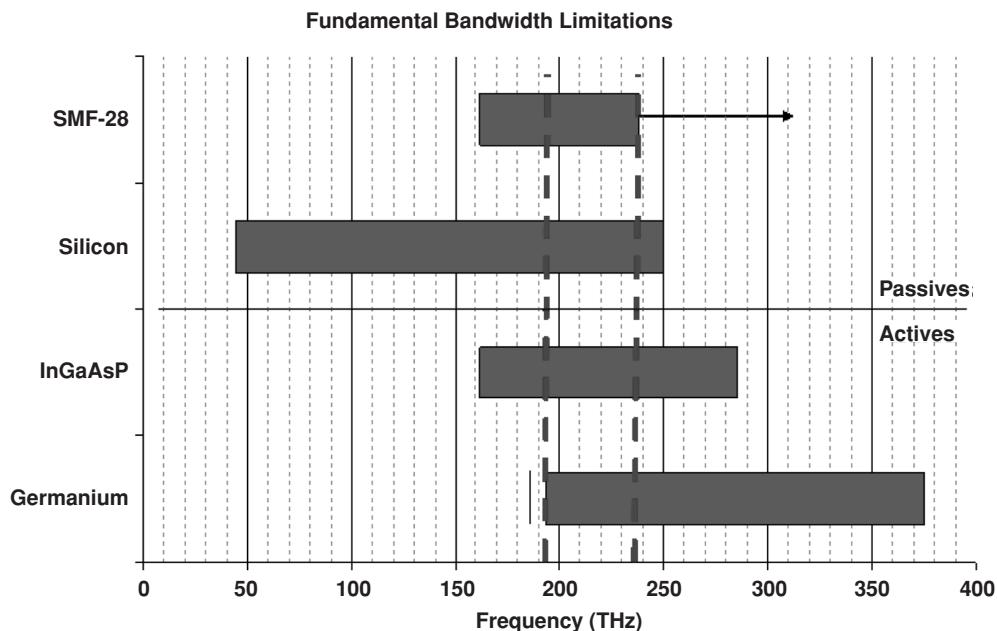
Figure 8.26 shows a block diagram of the transceiver. Note that the transceivers are nearly identical and placed with mirror symmetry on the die.

The major blocks include both optical components such as an integrated MZI modulator with the associated driver, and challenging RF circuitry such as the XFI interface, which is compliant with the 10 Gbps XFP signaling standard and includes CDR. This device has been published in detail in [31].

## 8.5 Conclusions

One interesting question to consider is what impact silicon-based photonic technology will have on the fundamental limitations of communications bandwidth. Figure 8.27 represents the major bounding constraints on optical bandwidth from a materials point of view. The bandwidth of SMF-28 is limited both by the material absorption, and by the single-mode cut-off frequency. The other limitations are materials systems, and it can be seen that there's a substantial overlap where the available III-V light source technology, silicon transparency window, and the germanium detector capability all overlap.

Today's WDM systems attain 0.10 bits/Hz spectral efficiency by spacing 10Gb payloads every 100 GHz. Ultradense systems attain 0.20 bits/Hz efficiency by utilizing a 50 GHz spacing



**Figure 8.27** Fundamental limitations of silicon photonics technology from a materials and media perspective

for the same payload. Due to the availability of plentiful bandwidth, optical systems have not been driven to acquire the same spectral efficiency that RF communications has, but it's quite reasonable to make the assumption that 0.25 bits/Hz can be achieved. This indicates that the technology ultimately has the potential to place 10 Tb of data in a single fiber.

## References

- [1] J. W. Goodman, F. J. Leonberger, S. -Y. Kung and R. A. Athale, 'Optical interconnections for VLSI systems', *Proc. IEEE*, **72**, 1984, 850–866.
- [2] D. A. B. Miller, 'Rationale and challenges for optical interconnects to electronics chips', *Proc. IEEE*, **88**, 2000, 728–749.
- [3] A. Liu, L. Liao, D. Rubin, H. Nguyen, B. Ciftcioglu, Y. Chetrit, N. Izhaky and M. Paniccia, 'High-speed optical modulation based on carrier depletion in a silicon waveguide', *Opt. Express*, **15**, 2007, 660–668.
- [4] J. M. Fedeli, R. Orobouchouk, C. Seassal and L. Vivien, 'Integration issues of a photonic layer on top of a CMOS circuit', *Proc. SPIE*, **6125**, 2006, 97.
- [5] A. Harke, M. Krause and J. Mueller, 'Low-loss singlemode amorphous silicon waveguides', *Electron. Lett.*, **41**, 2005, 1377.
- [6] P. Rojo Romeo, J. Van Campenhout, P. Regreny, A. Kazmierczak, C. Seassal, X. Letartre, G. Hollinger, D. Van Thourhout, R. Baets, J. M. Fedeli and L. Di Cioccio, 'Heterogeneous integration of electrically driven microdisk based laser sources for optical interconnects and photonic ICs', *Optics Express*, **14**, 2006, 3864.
- [7] G. G. Shahidi, A. Ajmera, F. Assaderaghi, R. J. Bolam, H. Hovel, E. Leobandung, W. Rausch, D. Sadana, D. Schepis, L. F. Wagner, L. Wissel, K. Wu and B. Davari, 'Device and circuit design issues in SOI technology', *Proceedings of the IEEE on Custom Integrated Circuits*, 1999, 339.

- [8] Wei-Zen Chen, Ying-Lien Cheng and Da-Shin Lin, 'A 1.8-V 10-Gb/s Fully Integrated CMOS Optical Receiver Analog Front-End', *IEEE J. Solid-State Circuits*, **40**, 2005, 1388.
- [9] C. L. Schow, L. Schares, R. A. John, L. S. Fischer and D. Guckenberger, '25 Gbit/s transimpedance amplifier in 0.13  $\mu\text{m}$  CMOS', *Electron. Lett.* **42**, 2006.
- [10] B. Analui, A. Hajimiri, 'Bandwidth enhancement for trabnsimpedance amplifiers', *IEEE J. Solid State Circuits*, **39**, 2004, 1263.
- [11] C. Kromer, G. Sialm, T. Morf, M. L. Schmatz, F. Ellinger, D. Erni and H. Jäckel, 'A low-power 20-GHz 52-dBOHM transimpedance amplifier in 80-nm CMOS', *IEEE J. Solid State Circuits*, **39**, 2004, 885.
- [12] M. Lee and E. A. Fitzgerald, *IEDM*, 2003, p. 429.
- [13] J. S. Goo *et al.*, *EDL-24*, 2003, 351.
- [14] T. Ghani, *et al.*, *IEDM*, 2003, 978.
- [15] J. D. Cressler and G. Niu, *Silicon-Germanium Heterojunction Bipolar Transistor*, Artech House, Boston, London, 2003.
- [16] G. Masini, L. Colace and G. Assanto 'Germanium thin films on silicon for detection of near-infrared light', In: *Handbook of Thin Films Materials*, H. S. Nalwa (ed.), Academic Press, 2001.
- [17] L. Colace *et al.*, 'Ge on Si  $p$ - $I$ - $n$  photodiodes operating at 10 Gbit/s', *Appl. Phys. Lett.*, **88**, 2006, 101111.
- [18] G. Dehlinger *et al.*, 'High Speed Ge-on-SOI Lateral PIN photodiodes', *IEEE PTL*, **16**, 2004, 2547.
- [19] O. I. Dosunmu *et al.*, 'Resonant Cavity Enhanced Ge Photodetectors for 1550 nm Operation on Reflecting Si Substrates', *IEEE J. Sel. Topics in Quantum Electr.*, **10**, 2004, 694.
- [20] L. Colace, G. Masini, G. Assanto, F. Galluzzi, G. Capellini, L. DiGaspare, E. Palange and F. Evangelisti, 'Metal–Semiconductor–Metal near-infrared light detector based on epitaxial Ge/Si', *Appl. Phys. Lett.*, **72**, 1998, 317.
- [21] J. Y. Dai *et al.*, 'Formation of cobalt silicide spikes in 0.18  $\mu\text{m}$  complementary semiconductor process', *Appl. Phys. Lett.*, **78**, 2001, 3091.
- [22] J. M. Hartmann, *et al.*, 'Reduced pressure-chemical vapor deposition of Ge thick layers on Si (001) for 1.3–1.55  $\mu\text{m}$  photodetection', *J. Appl. Phys.*, **95**, 2004, 5905.
- [23] A. Huang, C. Gunn, G. Li, Y. Liang, S. Mirsaidi, A. Narasimha and T. Pinguet, 'A 10 Gb/s photonic modulator and WDM MUX/DEMUX integrated with electronics in 0.13  $\mu\text{m}$  SOI CMOS', *ISSCC 2006 Dig. Tech. Papers*, 2006, 244–245.
- [24] S. B. Alexander, *Optical Communication Receiver Design*, SPIE Optical Engineering Press and the IEEE, 1997.
- [25] E. Sackinger, *Broadband Circuits for Optical Fiber Communication*, John Wiley & Sons, 2005.
- [26] X. S. Yao and L. Maleki, 'Optoelectronic oscillator for photonic systems', *IEEE J. Quant. Elec.*, **32**, 1996, 1141–1149.
- [27] X. S. Yao, L. Davis and L. Maleki, 'Coupled optoelectronic oscillators for generating both RF signal and optical pulses', *IEEE J. Lightwave Tech.*, **18**, 2000, 73–78.
- [28] C. Gunn, D. Guckenberger, T. Pinguet, D. Gunn, D. Eliyahu, B. Mansoorian, D. A. Van Blekkom and O. Salminen, 'A low phase noise 10 GHz opto-electronic RF oscillator implemented using CMOS photonics', *ISSCC 2007 Dig. Tech. Papers*, 2007.
- [29] D. L. C. Leung and H. C. Luong, 'A Fourth-order CMOS bandpass amplifier with high linearity and high image rejection for GSM receivers', *Proc. 1999 ISCAS*, vol. 2, 1999, 589–592.
- [30] W. M. C. Sansen and R. G. Meyer, 'Distortion in Bipolar Transistor Variable-Gain Amplifiers', *IEEE J. Solid State Circuits*, **8**, 1973, 275–282.
- [31] B. Analui, D. Guckenberger, D. Kucharski and A. Narasimha, 'A Fully Integrated 20-Gb/s Optoelectronic Transceiver Implemented in a Standard 0.13  $\mu\text{m}$  CMOS SOI Technology', to appear in *IEEE J. Solid State Circuits*, Dec. 2006.

# 9

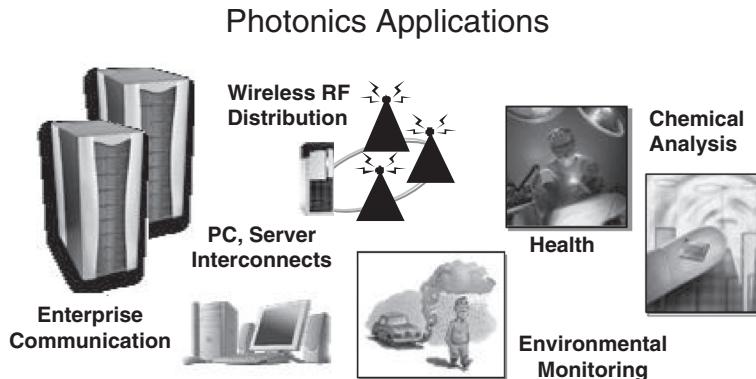
# Silicon Photonic Applications

Richard Jones, Haisheng Rong, Hai-Feng Liu and Mario Paniccia

## 9.1 Introduction

The previous chapters have discussed the significant advances in silicon photonic device performance over the past few years. These advances demonstrate that silicon can be considered as a material to build future optical devices. Silicon offers the opportunity for low-cost, highly scalable, opto-electronic components, and progress is moving at a rapid rate to both further improve device performance and to ‘commercialize’ this technology. If successful, silicon may come to dominate the optical communications arena as it already has the electronics industry. Figure 9.1 shows some of the many applications for silicon photonics such as the traditional communication areas ranging from long-haul telecommunications down to low-cost chip-to-chip interconnects. In addition radio-over-fiber (RoF) is emerging as a new area where silicon photonics could be used for distributing and transmitting wireless connectivity. Furthermore, integrated silicon lasers, detectors, and waveguide-based sensors may be applied to new and emerging applications in medical, healthcare and diagnostic arenas. All of these applications have the common requirement of low cost, small form factor, and high-volume manufacturability, which will only be possible if one can deliver these technologies on an integrated silicon platform.

In this chapter we will provide an overview of the various applications and benefits of silicon photonics. The chapter is arranged as follows; first optical communications and interconnect applications will be discussed. The various ‘distance’-based applications for interconnects will be discussed as well the transition from electrical to optical interconnects. In addition some key benefits and requirements for silicon photonics in each of these regions will be highlighted. Secondly radio-over-fiber applications are discussed and some key requirements for future RF networks are highlighted. Nonlinear effects in silicon are then discussed and their potential applications such as optical amplifiers, lasers and wavelength converters. Finally, a new and emerging area based on optical sensing is discussed and its related applications, including

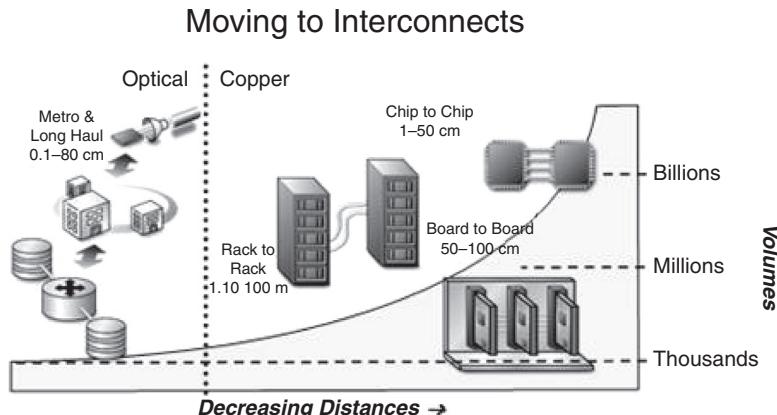


**Figure 9.1** Applications areas where silicon photonics could be applied. Today the cost of optical technologies is prohibiting its application into many of these areas

physical and chemical sensors and the technologies needed to produce an often talked about ‘lab on a chip’.

## 9.2 Communications and Interconnects

Figure 9.2 depicts the yearly estimated volume of interconnects as a function of link distance for both optical and electrical interconnects. Today for long-haul and metro applications, that is for data transmission at 10 Gb/s and distances above 100 m, optical is the only solution that is technically viable. This is in comparison to distances shorter than 100 cm where today’s interconnects are all copper based. In the mid-reach distance, in the (i.e. 10–100 m domain), optical

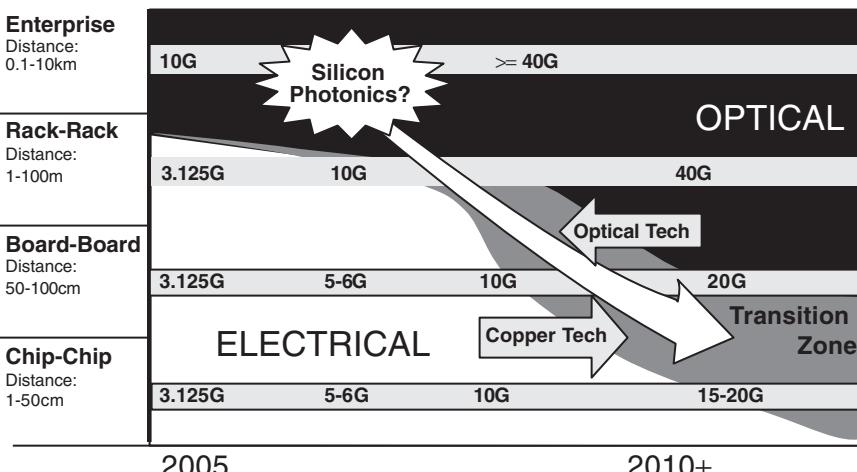


**Figure 9.2** Simple schematic showing various types of interconnects and the yearly estimated volumes as a function of distance. When taken as a whole a majority of the worlds interconnects are still copper based.

is starting to be discussed as a needed technology due to the fundamental speed limitations of copper based interconnects over these distances. It is interesting to note that if one looks at the *total* world wide volume of interconnects today, the number of optical interconnects remains a relatively small percentage when compared with those based on copper. Annual shipments of optical modules are measured in the hundreds of thousands to millions per year. For comparison, the number of computer systems shipped annually is in the hundreds of millions per year and each of these systems have hundreds if not thousands of copper interconnects per system. This takes the number of annual shipment of copper-based interconnects into the billions per year.

There are many benefits of moving from electrical to optical for future server and computer systems. These include removal of the length dependency when scaling to increased data rates which limits the link distance in electric interconnects at speeds >10 Gb/s. Cross-talk and electromagnetic interference (EMI) can also be minimized, along with the interconnect weight reduction when moving from bulky copper cables to thin optical fibers. New architectures become available from the ability to re-partition the server and PC architecture away from the microprocessor as a result of the increased optical interconnect length. Even with all these benefits however the transition from copper to optical interconnects will be determined predominately by cost. Here then is a tremendous opportunity for silicon photonics, but only if one can deliver photonic technology in high volume with low cost.

Figure 9.3 is a data-driven plot depicting the transition from copper to optical interconnects over different distances for various data rates. You can see from Figure 9.3 that there is a transition zone between electrical and optical interconnects that has some uncertainty. This is due to the competition between the reduction of optical costs over time that comes with maturing technology, with the constant and never ending pursuit and progression of Moore's

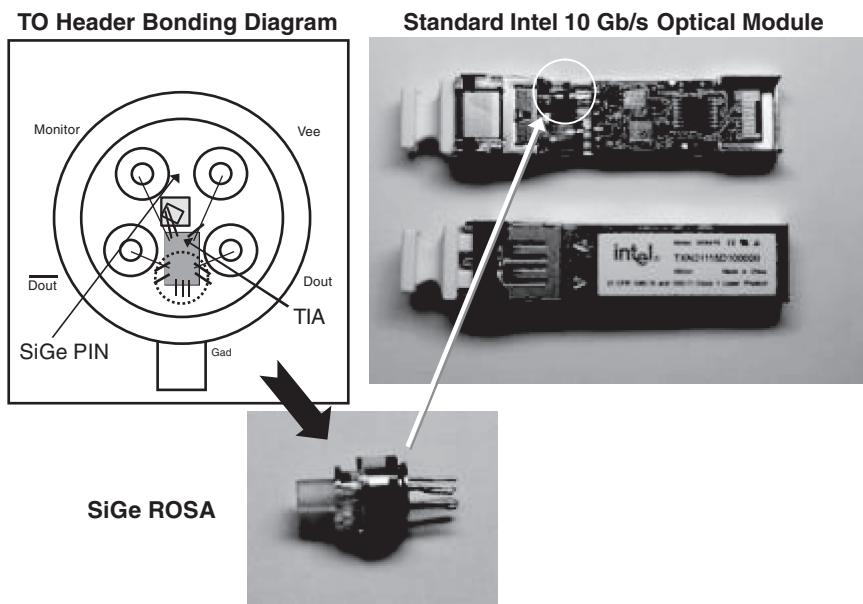


**Figure 9.3** Timeline depicting transition from copper to optical as function of distance and data rates. The data rates listed for each region are in Gb/s. Reproduced from Chapter 2 in *Silicon Photonics*, edited by Pavesi and Lockwood, Springer (ISSN 0303-4216), with kind permission of Springer Science and Business Media.

law. The combination of continued improvement in transistor performance and innovative architectures and equalization schemes result in the continuing bandwidth improvements of copper signaling over time. Recent work has demonstrated electrical signaling over copper FR4 traces at data rates of 20 Gb/s over 6 inch (15 cm) distances [1,2]. In addition, moving to improved channel materials such as flex cable has recently been used to demonstrate a 20 Gb/s interconnect over 30 inches (75 cm). [3]. It is expected that, for very short distances (i.e. <12 inches, 30 cm) and for data rates less than 20 Gb/s, copper will continue to be applicable for the foreseeable future. At higher data rates (i.e.  $\gg$ 20 Gb/s), optical interconnects will begin to have an advantage, and this is the place where silicon photonics could play a role. This opportunity will be revisited later when discussing Tera-scale computing.

For communications the opportunity for silicon photonics is to first find insertion points into longer-distance applications where discrete component replacement could allow for direct bill of material (BOM) savings over existing III-V components. This discrete component replacement will allow for product learning, reliability testing and qualification of silicon photonics. Then, over time, derivatives and more importantly higher-integrated photonic devices may be driven into the shorter-distance regions.

An example of a direct component replacement, or BOM savings opportunity, is a simple SiGe photodetector. This could be a direct silicon photonic component replacement for GaAs photodetectors used in applications such as fiber channel and Ethernet. Recently there has been significant amount of work in the area of SiGe-based photodetectors [4–6] to achieve the key requirements at 850 nm. These roughly are: low dark current (ideally of order  $\sim$ 100 nA), responsivity of approximately  $\sim$ 0.60 A/W; and bandwidths  $>$ 10GHz. Figure 9.4 shows



**Figure 9.4** Experiment where a SiGe pin PD was assembled and bonded into a receiver optical sub-assembly (ROSA) package, then integrated into a commercial 10 Gb/s optical module.



**Figure 9.5** Measured eye diagrams at 10 Gb/s, comparing GaAs (left) to SiGe (right) photodetectors.

a photograph of a SiGe photodetector that was assembled into a commercial 10 Gb/s optical transceiver. Figure 9.5 are measured eye diagrams showing the performance comparison of SiGe-based a photodetector compared with that of an identical commercial transceiver based on a GaAs photodetector. One can see from Figure 9.5 that the performance of the SiGe photodetector is comparable to that of the existing III-V photodetector. More detailed results of the SiGe photodetector reliability and testing can be found elsewhere [6].

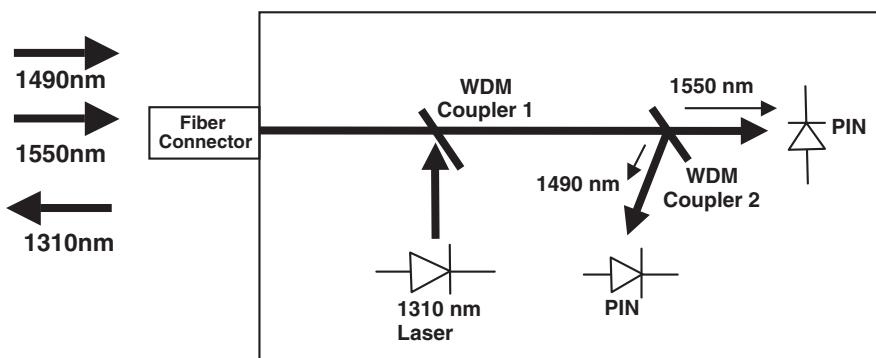
Another potential direct component replacement is a silicon-based modulator. Silicon modulators have been experimentally demonstrated with upto 40 Gb/s data transmission capability [7, 8], and simulations show that scaling to even higher speeds is possible [9]. The costs of lithium niobate and III-V-based electro-absorption modulators (EAMs) are still very high which may allow silicon-based modulators to penetrate into long-haul, metro or enterprise arenas. Key performance parameters to trade off are the extinction ratio (ER), drive voltage, size and insertion loss. More importantly integration of multiple modulators into a silicon platform could allow one to use more complex modulation formats such as dual-phase shift keying (DPSK) to achieve higher (i.e. >40 Gb/s) more bandwidth efficeint data rates [10]. This integration capability clearly favors silicon-photonic-based devices over discretes.

Internet traffic is increasing due to data, voice, video, along with wireless and cellular traffic, telecommuting and new multimedia functions, and this is pushing the industry to consider higher baseline data rates than the 10 Gb/s found today in the majority of backbone networks. The next transition is expected to be at 40 Gb/s for SONET and 100 Gb/s for Ethernet, with these technology transitions expected to ramp in volume by the end or early part of the next decade (i.e.  $\sim$ 2010). At these higher data rates the focus will be on cost, form factor and power dissipation, with integration providing real and needed value to enable these next-generation technology transitions. As an example, for 100 Gbit Ethernet, one could imagine delivering an integrated transmitter technology using ten wavelength division multiplexed channels each modulated at 10 Gb/s using modulator arrays and a demux, all integrated together onto one piece of silicon. Or one could make this using five channels running at 20 Gb/s or three channels at 40 Gb/s. These architecture trade-offs are only possible with integration capability that silicon photonics provides. A similar analogy can be made for the integrated receiver portion of the link.

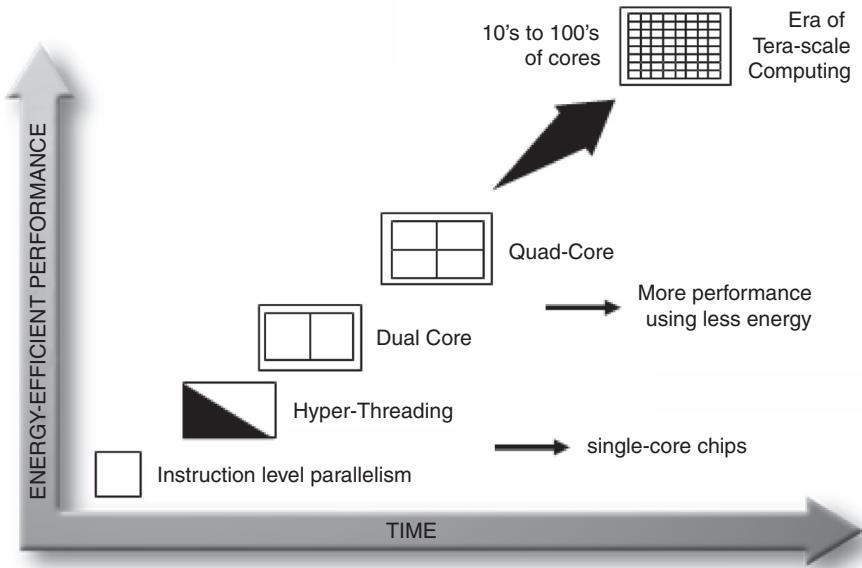
There is discussion that future high-speed ( $>40\text{Gb/s}$ ) networks may also be based on ‘all-optical’ networks rather than the more conventional optical-electrical-optical conversion (O-E-O) points used today. In addition, these future networks will be more limited and/or affected by chromatic dispersion (CD) and polarization mode dispersion [11] and other impairments due to nonlinear effects. To address these issues technologies based on all optical amplification, wavelength conversion and all-optical switching will be needed. The ability to provide these types of technologies on a silicon based platform will have profound impact to the communications arena. Some of these applications will be discussed in more detail in Section 9.3.

This high-speed core optical communication network is also beginning to penetrate into the access domain with a growing demand for high-quality multi-media services such as video on demand services. This is putting increased pressure for high bandwidth capabilities into the home, resulting in the new and emerging market of fiber to the home (FTTH), or fiber to the curb or fiber to the X (FTTX). The requirements for these applications are very low cost, low loss, bandwidth, and possibly interoperability with wireless-based technologies moderate (100 Mbps-1 Gbps) that are beginning to become more pervasive. Relevant photonic devices for FTTH include  $1 \times N$  splitters, multiplexer and demultiplexer devices and most common triplexers. The triplexer consists of a transmitter, typically at 1310 nm (often referred to as upstream) and two photodetectors for the longer-wavelength downstream signals. One photodetector receives digital data, typically 1490 nm, and the other analog receives data for broadcasting services such as cable television signal, at 1550 nm [12,13]. Figure 9.6 is a schematic of a simple triplexer typically found in FTTH applications. The key to developing triplexers is highly selective filtering and transmission of the up and downstream optical signals as well as the design and integration of low electrical cross-talk receivers. Today most triplexers are manufactured using discrete thin-film filters, power monitors, lasers and photo detectors. The uniqueness and value of silicon photonics here would be the ability to integrate the entire triplexer onto one piece of silicon, eliminating the numerous components, allowing for simple alignment of the devices, reducing the overall size and cost of the triplexer module.

Figure 9.7 is a plot showing the future transition of microprocessor architectures over time, as the industry moves from a single microprocessor core to multiple processing cores within a single chip. The ability to fabricate multiple cores on one die is possible due to the continuation



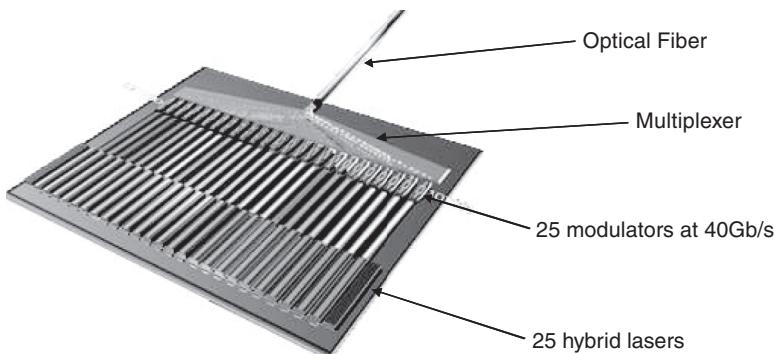
**Figure 9.6** Schematic of triplexer used for FTTH application.



**Figure 9.7** Plot showing future transition of microprocessor architectures over time evolving from single cores to multiple cores to in the future possibly hundreds of cores integrated onto a single chip

of Moore's law and the continued scaling of the size of future transistors. As can be seen from Figure 9.7, processor cores will scale from 4 to 8 to 16 cores and continue doubling into the future. The result is that in the near future, we will have possibly hundreds of processing cores on a single chip. This level of parallelism will usher in an era of what is being called Tera-scale computing when computers will deliver Teraflop performance on a single chip [14]. Each of these future computing systems will not only be processing Terabytes of data, but also sending and receiving Terabits of information. It is clear that optical interconnects will play a significant role in this Tera-scale computing era as scaling copper signaling to  $>1\text{ Tb/s}$  over distances of a few meters will become increasingly difficult. The key enabler here will be the ability to integrate all the optical functions needed for a high-bandwidth optical link onto silicon with low-cost, high-volume manufacturability. In addition interoperability with the electrical die in terms of low latency, low jitter and low power consumption need to be addressed. Significant amount of innovation is still needed to address the packaging and electrical integration issues for enabling high-bandwidth optical devices in and around the PC and server environment.

Figure 9.8 shows a schematic of a proposed terabit silicon photonic chip, which incorporates 25 WDM channels modulated at 40 Gb/s by an array of silicon modulators before being combined using a silicon multiplexer. One of the key remaining challenges for fabricating such a device is the realization of electrically pumped lasers on silicon. While an electrically pumped laser fabricated from silicon has yet to be developed, a new approach, based on hybrid silicon/AlGaInAs integration has shown promising results. This approach uses AlGaInAs offset quantum wells bonded to silicon waveguides to fabricate a hybrid, evanescently coupled, waveguide where the mode is guided by the silicon layer, and can achieve electrically driven gain from the III-V layer. The advantage of this approach is that no critical alignment is



**Figure 9.8** Schematic drawing of a terabit silicon photonic chip showing an array of evanescent lasers fabricated on a silicon chip, all self-aligned to silicon modulators and multiplexed to a single output fiber. Reproduced from 'A Hybrid Silicon Laser', Richard Jones, Oded Cohen, Mario Paniccia, Alexander W. Fang, Hyundai Park and John Bowers, *Photonics Spectra* (January 2007) by permission of Laurin Publishing Co., Inc.

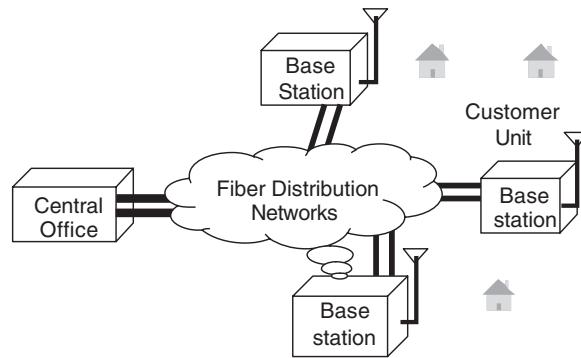
needed between the silicon waveguides and the III–V wafer prior to bonding, and thus many hybrid lasers can be fabricated from a single bond step that are self-aligned to other silicon photonic devices on the same chip. This highly scalable approach has enormous advantages over the conventional pick-and-place approach of bonding individual laser dies to individual passive waveguides in terms of both cost and assembly time, especially when many sources are required [15].

### 9.2.1 Radio-over-fiber (RoF) RF Applications

The use of fiber optical links for the transmission of modulated analog RF signals has found a variety of commercial applications, including distribution of cable TV signals and antenna remoting [16–17]. Recently, radio-over-fibers (RoF), where high-capacity optical links are integrated with the flexibility of wireless radio networks, has emerged as an attractive technology to meet the increasing demands for high-capacity broadband access networks and personal communication system applications [18–20]. An example application of this is for distributing short-reach wireless antennas throughout a building; these short-reach antennas would be connected to the primary antennas on the roof via optical links to achieve better wireless or cellular coverage.

The main challenges for RoF optical links are in the optical-to-electrical (O/E) and electrical-to-optical (E/O) conversions at the antenna sites. Silicon photonics offers value here as the ability to integrate optical and electrical devices on one silicon platform could enable small form factor, lower-power highly integrated modules. In this section we will briefly discuss the technical details of these systems and the key optical requirements and components needed for such systems.

In RoF systems instead of using many complex and expensive base stations (BSs), a central office (CO) is used to perform most of the signal processing (coding, multiplexing, RF generation and modulation, etc.) and the signals are transmitted to and from lower-cost, simplified,



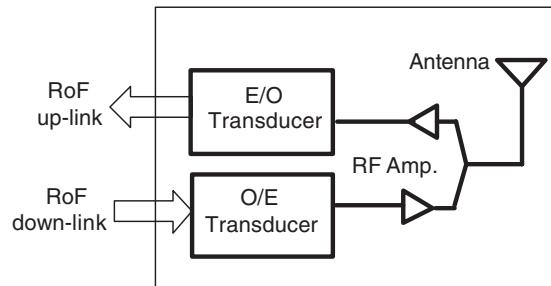
**Figure 9.9 Schematic diagram of radio-over-fiber systems**

BSs via fiber optical networks, as shown schematically in Figure 9.9. This architecture with shared COs significantly simplifies the design of BSs, which essentially only consist of optical-to-electrical and electrical-to-optical transducers, an antenna and some microwave circuitry, as shown in Figure 9.10.

In addition to taking full advantage of existing high-capacity optical networks, and the short-range broadband point-to-multipoint radio link, RoF offers flexible remote interface with multiple BSs and can reduce system complexity. Because analog optical links are typically multi-channel and have much higher power compared with digital links, there are two important measures for the RF performance in addition to gain and bandwidth, these are noise figure and spur-free dynamic range (SFDR). In particular, it is often the SFDR that constrains the system performance and is defined as two-thirds the difference between the third order intercept point and the minimum discernable signal. For typical RoF systems, the required SFDR can range from 95 to 130 dB-Hz<sup>2/3</sup>, while for some military applications, there is never enough SFDR.

There are basically three schemes to transport RF signal over fibers.

- RF-over-fiber
- IF-over-fiber
- Baseband-over-fiber.



**Figure 9.10 Basic E/O and O/E building blocks in a typical base station**

Among them, RF-over-fiber has the simplest BS design while the frequency translation in the other schemes reduces the effects of fiber dispersion considerably. Each scheme has its own unique advantages and technical challenges, therefore leading to different BS designs [21], but highly linearized O/E and E/O transducers as well as linear power amplifiers and low-noise amplifiers are the key building blocks for achieving good RF link performance. Over the past 15 years, there have been tremendous research efforts in designing various RF photonic links using different photonic devices to improve the link performance [22–26].

Three widely used E/O transducers are directly modulated semiconductor lasers, electro-optic modulators (EOM) and electro-absorption modulators (EAM). While direct modulation links have large SFDR at low frequencies ( $125 \text{ dB-Hz}^{2/3}$  at  $< 1 \text{ GHz}$ ), it degrades significantly at higher frequencies, primarily due to the relaxation oscillation resonance [27]. On the other hand, external modulation links can have relatively flat SFDR over a wider frequency range, and since Mach-Zehnder interferometer (MZI)-based EOMs can be biased to create zero second-order distortion, the challenge is to minimize the third-order nonlinearities. Links with SFDR as high as  $112 \text{ dB-Hz}^{2/3}$  have been demonstrated using MZI EOMs at frequencies up to 17 GHz [28]. Recently, it was also shown that controlling the modulator bias voltage is an effective way to optimize carrier-to-noise ratio in RoF optical links [29]. Although EAMs have also been demonstrated as an attractive E/O transducer for many applications, including RoF [30], the achieved SFDR is usually not as high as that using EOM due to its poor power handling ability.

It should be noted that although all E/O transducers are intrinsically nonlinear devices, their SFDR can be improved by using a variety of linearization techniques including feed-forward, and pre-distortion to compensate for the nonlinearities of the electrical-to-optical transfer function of the E/O transducers. There have been extensive research activities to linearize all three types of E/O transducers. Both feed-forward and pre-distortion have been shown to be effective in suppressing second- and third-order distortion in directly modulated and externally modulated analog links. The challenge, however, is to linearize the transfer function over broadband. A high performance 6–12 GHz, 20 km analog optical link with  $108 \text{ dB-Hz}^{2/3}$  SFDR is shown with EOMs linearized by pre-distortion [31]. One interesting development is the use of CMOS-based pre-distortion circuit integrated with EOM, which showed a 14 dB improvement in SFDR over 1 GHz frequency range [32].

High-speed photodetectors have been used as the O/E transducers in RF photonic links. The desired link performance calls for photodetectors that can respond efficiently and linearly to modulated high-frequency optical signals [33]. Furthermore, as high SFDR can be achieved in modulator-based links at high launch power, detectors with high saturation currents are needed to fully exploit the benefits of an improved link SFDR. The recently developed uni-traveling carrier photodetectors have exhibited a 50 mA saturation current with a 55 GHz bandwidth [34]. The use of velocity-matched distributed photodetectors has been shown to be another promising approach, where 45 mA linear photocurrent was obtained in a device with 35 GHz bandwidth [35].

In addition to developing highly linear devices, there have also been tremendous efforts in linearizing photodetectors by various compensation techniques to maximize the SFDR of RoF links. Adaptive post-distortion is one such technique, which has been successfully applied to linearize photodetectors with a monolithic CMOS polynomial generator [36]. Integrating such CMOS-based electronic circuitry with silicon-based high-speed photodetectors would be an interesting development demonstrating the values of silicon photonics.

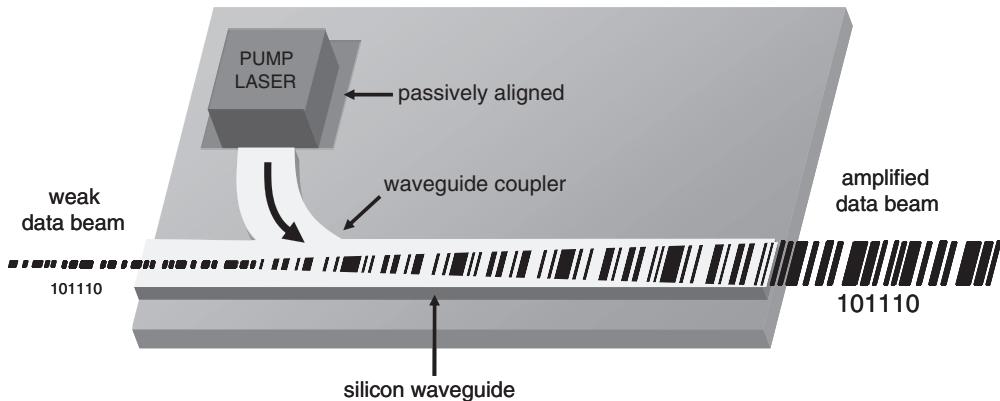
Another important building block is the RF amplifiers at millimeter-wave frequencies. As CMOS technology delivers transistors with  $f_t$  and  $f_{max}$  in excess of 100 GHz, fully integrated circuit blocks from 1 to 100 GHz have been demonstrated over the past decade [37]. This capability, along with the development of CMOS-based low-noise amplifiers, enables CMOS radios operating at millimeter-wave frequencies. Of particular interest is the 7 GHz unlicensed frequency band near 60 GHz, which can offer a high degree of frequency reuse as a result of the high oxygen absorption in air. One interesting development would be the combination of CMOS radio and RoF, which would lead to a fiber-distributed 60 GHz wireless personal area communication network providing indoor high-data-rate omnidirectional wireless links, home video streaming, and outdoor line-of-sight high-data-rate backbone link.

One of the main challenges for wide-scale deployment of RoF is the development of cost-effective and integrated high-speed photonic devices. As has been discussed previously, over the past few years there have been quite a number of significant breakthroughs in silicon-based photonic devices [38]. The feasibility of using Si material to make high-speed optical modulators and photodetectors, two key building blocks in RoF system, has been shown with 10 Gb/s Si-based modulators [7, 8, 39], and SiGe photodetectors with a bandwidth exceeding 35 GHz being demonstrated [5]. A key focus for RoF systems will be the ability to integrate multiple optical components together along with complicated RF electronics and circuitry to achieve better performance at lower costs. As the infrastructure costs for deploying RF systems scale with the number of multiple antenna elements, the ability to integrate while reducing size and cost is essential. In addition, unlike digital systems, the key focus here is on linearity of the modulators and photodetectors. This is an area that will need future research and development to understand the applicability of silicon photonics in this RF arena.

### 9.3 Nonlinear Optical Effects in Silicon and Applications

One of the more recent and important branches in silicon photonics is the study of various nonlinear optical effects in silicon to create active photonics devices. Due to the high index contrast of silicon to its oxide, light can be confined in silicon waveguides more tightly than in glass fibers or silica-based waveguides. The optical modal area in silicon waveguides can be more than  $10^3$  times smaller than in a single-mode fiber. Moreover, many optical nonlinear effects are much stronger in silicon than in glass fibers, for instance, the Raman gain coefficient in silicon is 3–4 orders of magnitude greater than in an ordinary glass fiber. Many of the concepts and applications already developed today that are based on nonlinear optical effects in silica fibers can be adapted for silicon waveguides to produce compact, chip-scale active photonics devices for various applications, including optical amplifiers, lasers, and wavelength converters, as well as ultra-fast switching, pulse shaping, and slow-light generation devices.

In recent years, rapid progress has been made in this area and several breakthroughs have been reported. In particular, stimulated Raman scattering (SRS) has been successfully employed to demonstrate all optical amplifiers [40–48] and lasers [49–54] in silicon. In the following sections we will discuss some recent progress in the area of silicon photonics to produce optical amplifiers, lasers and wavelength converters.

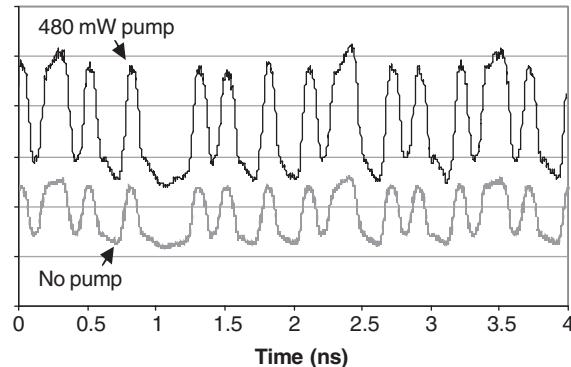


**Figure 9.11** Simplified schematic of a silicon-based on-chip optical amplifier

### 9.3.1 Silicon Amplifiers and Lasers

Although silicon has a much higher Raman gain coefficient and silicon waveguides provide stronger light confinement when compared with silica fiber, the optical loss in silicon is significantly higher than in fiber, especially the nonlinear loss due to two-photon absorption (TPA)-induced free-carrier absorption (FCA) [55–58]. Net optical gain was achieved for the first time in silicon waveguides using a pulsed pump configuration [40], and with narrower pump pulses, on-off peak Raman gain as high as 20 dB has been reported [43]. CW amplification and lasing is significantly more challenging due to the competing nonlinear optical loss mechanism, as TPA-induced FCA causes the optical loss to increase with pump power [56]. By introducing a reverse biased  $p-i-n$  diode structure embedded in a silicon waveguide, the nonlinear absorption can be efficiently reduced and CW net Raman gain has been demonstrated [46]. Figure 9.11 is a simple schematic of an all-optical amplifier; a pump source is directed into the silicon waveguide and is combined with the incoming data stream passing through the silicon waveguide. By using this configuration on-chip amplification of nearly  $\sim 3$  dB of a 10 Gb/s data stream and on-off gain of 2.3 dB of a 40 Gb/s data stream has been recently demonstrated (Figure 9.12) [59, 60].

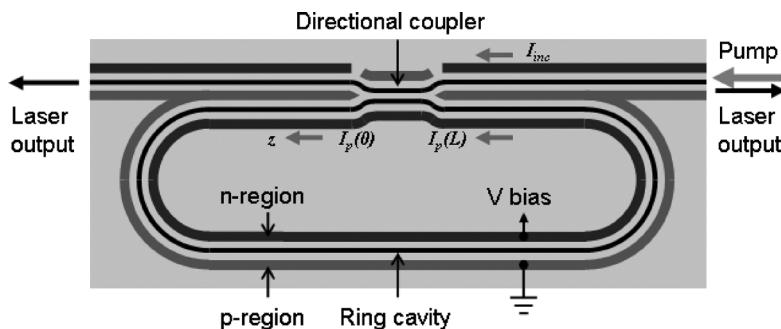
Pulse Raman lasing in silicon was first reported in a hybrid laser cavity formed by using an 8-m-long external fiber loop [49]. An all-silicon, chip-scale continuous wave (CW) Raman silicon laser was realized and published soon after in a laser cavity formed by applying multi-layer optical coatings onto the chip facets of a linear cavity [51]. After these first proof-of-principle lasing demonstrations, recently a fully monolithic integrated ring cavity Raman silicon laser has been developed [52]. This ring-based architecture removes the need for optical coatings on the silicon waveguide facets and allows for smaller and more integrated laser devices. This ring laser design allows for on-chip integration with other silicon photonics components to provide a highly integrated and scalable monolithic device. The ring laser cavity is constructed from a silicon-on-insulator (SOI) rib waveguide forming a racetrack-shaped ring resonator (Figure 9.13). A bus waveguide is connected with the ring cavity via a directional coupler which couples both pump and laser light into and out of the cavity.



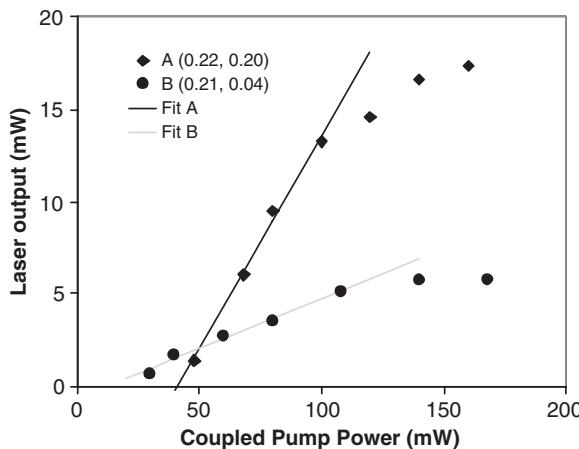
**Figure 9.12** 10 Gb/s optical PRBS data stream passing through a silicon waveguide with (top) and without (bottom) optical pump. On/off gain is  $\sim 3$  dB [59]. Reproduced from ‘Recent development on silicon-based Raman lasers and amplifiers’, Haisheng Rong, Ying-Hao Kuo, Shengbo Xu, Oded Cohen, Omri Raday, and Mario Paniccia in *Proc. SPIE* **6389**, 638904 (2006) by permission of SPIE.

The coupling ratio depends on the input wavelength and polarization and can be varied by changing the gap and/or length of the coupler. A  $p-i-n$  diode structure, as depicted in Figure 9.13, is embedded along the waveguides to reduce the nonlinear optical loss due to two-photon absorption (TPA)-induced free-carrier absorption (FCA) [46].

From this design a single-mode CW laser with less than 20 mW threshold and a 23% slope efficiency has been achieved. (Figure 9.14). The side-mode suppression ratio (SMSR) exceeds 70 dB and the laser line-width is  $<100$  kHz (Figure 9.15). Such superior spectral purity from this type of laser can provide unique advantages for applications in coherent optical communications, spectroscopy, metrology, and optical sensing where high spectral purity is of significant value. In addition, Raman lasers have the advantage of being flexible in wavelength



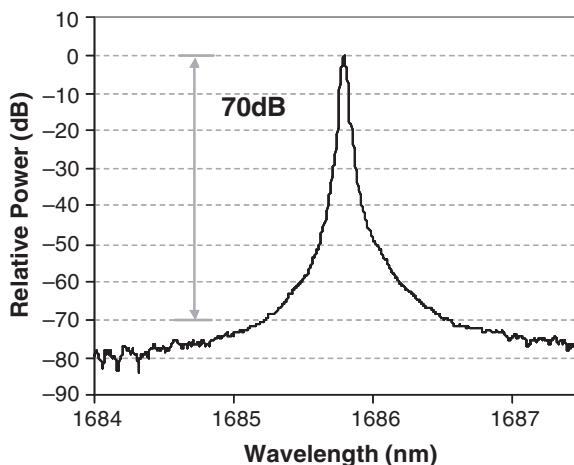
**Figure 9.13** Layout of the silicon ring laser cavity with a  $p-i-n$  structure along the waveguides. Reproduced with permission from Figure 1 in H. Rong, Y. -H. Kuo, S. Xu, A. Liu, R. Jones, M. Paniccia, O. Cohen, and O. Raday, ‘Monolithic integrated Raman silicon laser,’ *Opt. Express*, **14**, 6705–6712 (2006). <http://www.opticsinfobase.org/abstract.cfm?URI=oe-14-15-6705>



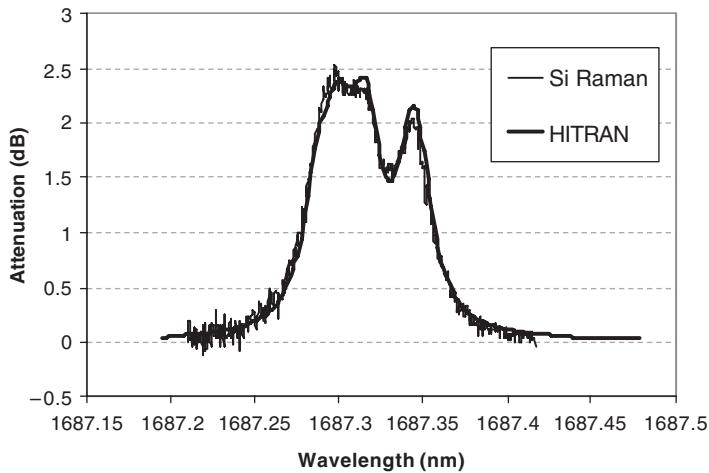
**Figure 9.14** Laser output power as a function of the coupled input pump power for two ring laser cavities, A and B, with different coupling ratios. The numbers in the parentheses indicate the coupling ratios for the pump at 1550 nm, and the laser wavelength at 1686 nm, respectively

as, by changing the pump wavelength, the Raman lasing wavelength can be tuned over a broad wavelength range that is not easily achievable with other types of lasers.

An example of how one could use the wavelength tunability and narrow line-width of the Raman ring silicon laser is demonstrated in a spectroscopy experiment for gas sensing applications. Figure 9.16 shows a methane absorption spectrum recorded using a CW Raman



**Figure 9.15** Silicon ring laser spectrum measured with a grating-based optical spectrum analyzer showing a side mode suppression of >70 dB. Reproduced from Figure 6 in H. Rong, Y. -H. Kuo, S. Xu, A. Liu, R. Jones, M. Paniccia, O. Cohen, and O. Raday, ‘Monolithic integrated Raman silicon laser,’ *Opt. Express*, **14**, 6705–6712 (2006). <http://www.opticsinfobase.org/abstract.cfm?URI=oe-14-15-6705> by permission of the Optical Society of America

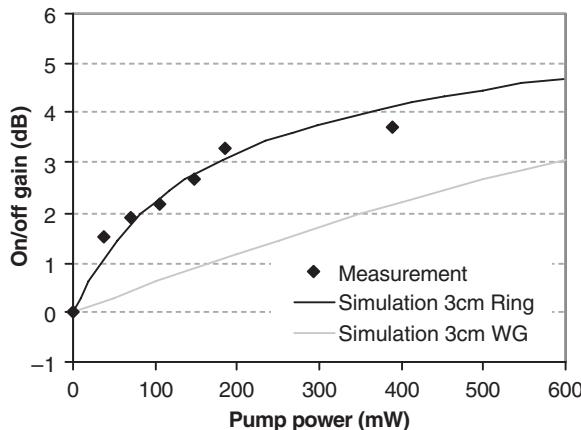


**Figure 9.16** Methane absorption lines at 1687 nm taken with a CW Raman silicon ring laser with >10 GHz mode-hop-free tuning

silicon laser. The laser frequency tuning was achieved by scanning the pump laser's frequency, achieving a mode-hop-free tuning range of greater than 10 GHz. Figure 9.16 shows the profile for a group of methane absorption lines around 1687 nm. The measurement data agree very well with the calculated line profile based on the HITRAN data base [61].

Using the ring resonator architecture, one can also construct a chip-scale Raman amplifier that takes advantage of the resonance effect to increase the effective pump power and reduce the overall size of the device. If the coupler is designed in such a way that the pump is critically coupled into the resonator, the pump power inside the ring resonator is most effectively enhanced. If the coupling for the Stokes wavelength is high enough, lasing will not occur and instead when a signal beam at the Stokes wavelength is passed through the resonator it is amplified by the strong SRS inside the ring resonator. For the signal wavelength, this resonator has low  $Q$ , therefore the amplification bandwidth will not be affected significantly by the cavity resonance. Figure 9.17 plots the measured and calculated Raman amplification vs. waveguide-coupled input pump power for a ring resonator of length 3 cm, compared with a simple straight waveguide of the same length. In this simulation the parameters used were: 20% coupling for the pump, 100% coupling for the signal wavelength, and a waveguide loss of 0.3 dB/cm. As shown in Figure 9.17, more than 3 dB on-off gain can be achieved with 200 mW pump using a ring configuration as compared with 600 mW pump when using a straight waveguide. The pump power is reduced by factor of 3 to achieve the same gain level.

Future work is still needed to improve these devices, including reduction of the TPA-induced FCA, either by reducing the free-carrier lifetime and/or improving the  $p-i-n$  efficiency. Also important for commercial applications is reducing the reverse bias voltage, as well as reducing the optical pump power needed for these nonlinear silicon-based devices. Recently, it has been shown that net Raman gain can be achieved with zero bias, i.e. no electrical power dissipation [48]. Alternatively, the free-carrier lifetime can be shortened significantly by implanting defects

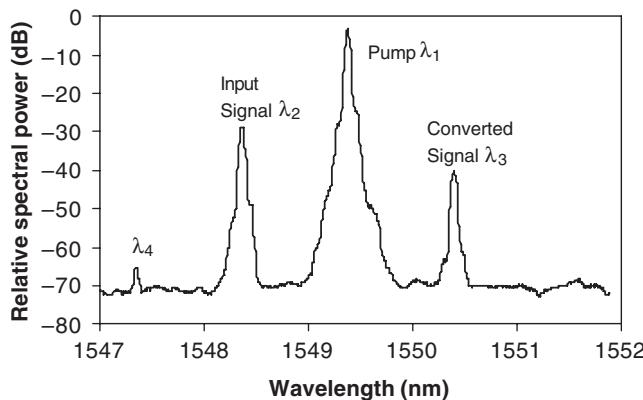


**Figure 9.17** Measured and simulated on/off gain of a Raman ring silicon amplifier compared with a linear amplifier with the same 3-cm-long waveguide

into silicon waveguides, e.g. by implanting helium ions into silicon, net CW Raman gain has been observed [47]. Reducing the waveguide cross-sectional dimensions and/or introducing impurities into the waveguides in combination with a *p-i-n* diode structure could reduce the carrier lifetime to below 1ns, thus reducing lasing threshold and increasing laser output power. In addition, other cavity configurations can be considered to minimize the device size and increase efficiency. A monolithic integrated cavity structure based on photonic bandgap has been proposed where a sub-milliwatt lasing threshold is predicted [62]. Similar to fiber Raman ring or micro-cavity lasers [63,64], cascaded lasing is also possible using silicon ring cavity lasers. This would allow the lasing wavelength to be extended into the mid-infrared region where a suitable room-temperature semiconductor laser is in high demand. For many applications ranging from high-resolution and ultra-sensitive detection of molecules for trace gas analysis, pollution and toxic gas monitoring, and biomedical sensing, to free space communications.

### 9.3.2 Wavelength Conversion

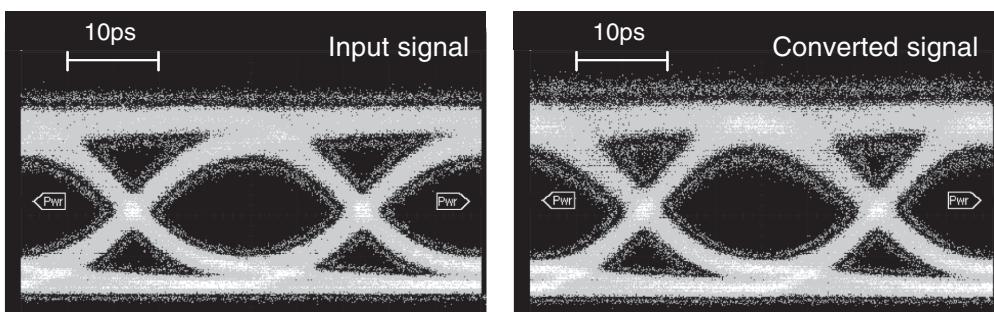
All-optical wavelength conversion based on nonlinear interactions of light beams of different wavelengths inside silicon waveguides is another attractive application area. Coherent anti-Stokes Raman scattering or four-wave mixing (FWM) in silicon waveguides have been studied [65–69]. Using a CW pump laser, a FWM conversion efficiency as high as –8.5 dB has been demonstrated with a *p-i-n* waveguide under reverse-biasing (Figure 9.18) [69]. This level of conversion efficiency is comparable to LiNbO<sub>3</sub>-based devices [70,71]. However, compared with wavelength converters based on difference-frequency generation in periodically poled lithium niobate waveguides, silicon waveguide devices are not subject to photorefractive damage causing degradation of the quasi-phase matching conditions and therefore should have superior reliability. It has also been demonstrated that high-speed optical data stream at 40 Gb/s on one DWDM channel in the C-band can be converted to another channel with practically no waveform distortion (Figure 9.19)[72].



**Figure 9.18** FWM spectrum of a silicon-based wavelength converter [69]. Reproduced Figure 3 in H. Rong, Y. -H. Kuo, A. Liu, M. Paniccia, and O. Cohen, ‘High efficiency wavelength conversion of 10 Gb/s data in silicon waveguides,’ *Opt. Express*, **14**, 1182–1188 (2006). <http://www.opticsinfobase.org/abstract.cfm?URI=oe-14-3-1182> by permission of the Optical Society of America

By properly choosing the cross-sectional dimensions, broadband optical parametric gain in SOI channel waveguides has also been observed [73], and phase-matched FWM with 5.2 dB conversion efficiency can be obtained using pulsed pumping. It may be possible to achieve CW operation if the free-carrier lifetime can be shortened further by introducing, for instance, a reverse-biased *p-i-n* structure. However, a careful balance must be maintained in reducing free-carrier lifetime without increasing waveguide loss.

Future research work and investigations in other areas related to utilizing silicon’s nonlinear properties include optical bistability for switching and memory applications [74–77], ultra-fast, all-optical switching by cross-absorption modulation [78], pulse shaping [79], and Raman-induced slow light for tunable optical delay applications [80].



**Figure 9.19** Eye diagrams of the wavelength converted (right) signal with input (left) signal at 40 Gb/s [72]. Reproduced from Figure 4 in Y. -H. Kuo, H. Rong, V. Sih, S. Xu, M. Paniccia, and O. Cohen, ‘Demonstration of wavelength conversion at 40 Gb/s data rate in silicon waveguides,’ *Opt. Express*, **14**, 11721–11726 (2006). <http://www.opticsinfobase.org/abstract.cfm?URI=oe-14-24-11721> by permission of the Optical Society of America

## 9.4 Sensing

The integration advantages of silicon processing may also be applied to sensing applications. In this section an overview of silicon sensors will be given, split into four sections: physical sensors; chemical sensors; biochemical sensors; and integrated devices. The principles behind the commonly used architectures will be discussed along with some of the potential applications.

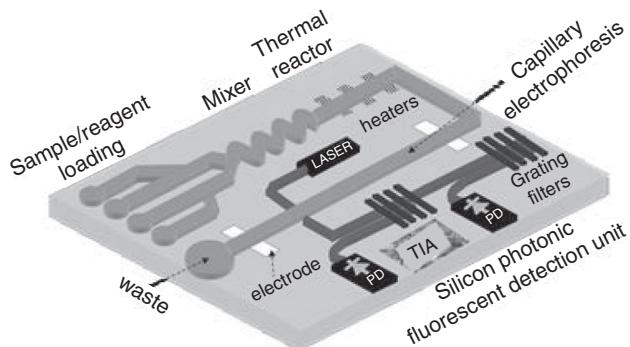
Advantages of silicon photonics and the use of silicon in this field include:

- the ability of photonic sensors to take advantage of the prior art biochemical assays based on fluorescence technology;
- the fact that silicon dioxide, glass, is one of the more commonly used substances in biochemical labs with convenient surface modification techniques to allow stable deposition of sensing layers;
- novel sensing methodologies open to waveguide based sensors, e.g. evanescent sensing;
- the ability to integrate different sensing technologies onto a single chip, e.g. silicon photonic sensors for cell identification and electrical sensors for ion concentration measurement;
- fabrication of miniature micro-fluidic channels to control the flow of liquid through the chip integrated with silicon sensors;
- MEMS technology to switch and control fluid flow through the chip;
- integration of E-field control on chip, which is useful for applications such as electrophoresis for DNA analysis;
- the high thermal conductivity of silicon, allowing for faster temperature changes, significantly speeding up temperature cycling assays, e.g. polymerase chain reaction (PCR) [81];
- preliminary indications that bulk silicon is biocompatible, [82] paving the way for implantable silicon sensors.

Integration is key here: multiple sensing elements can be multiplexed on a single die with micro-fluidic channels to control the flow of different reagents with the sample under test, allowing fabrication of miniature chemistry labs on a single silicon die in which reagents can be mixed together in tiny reactors before being fed to micro-sensors that detect the different chemical and biological species. These kinds of highly integrated systems are known as labs-on-a-chip or micro-total analysis systems ( $\mu$ -TAS) [83, 84], see Figure 9.20, and are advantageous over conventional testing methodologies in requiring smaller sample volumes, less chemical reagents and less time, thus reducing testing and assay costs. Further integration can permit sample preparation before the detection, for example DNA may be extracted from cells on a chip, chemically amplified to measurable quantities, and separated by electrophoresis prior to detection, which opens up a whole new paradigm for the way lab testing is done today.

### 9.4.1 Physical Sensors

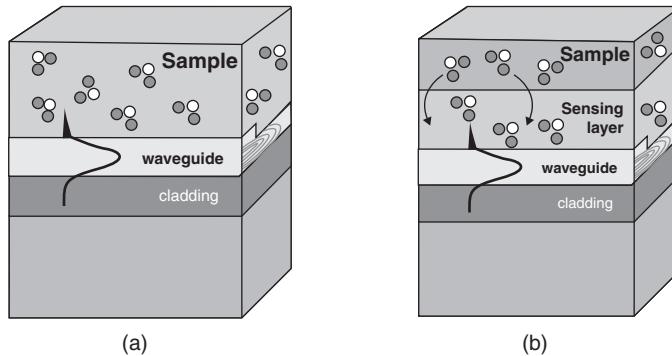
Silicon micro- or nano-devices are commonly used to detect physical quantities, such as temperature, acceleration and pressure, an obvious commercial example being the MEM's accelerometer used for airbag deployment in automobiles [85]. For both aviation and automotive applications electromagnetic interference (EMI) tolerance, low weight and low cost are critical requirements. Photonic sensors have the potential to achieve higher sensitivity, and enable use in harsher environments with greater tolerance to both temperature variations and



**Figure 9.20** Schematic of a silicon photonic lab-on-a-chip showing the integration of dark-grey optical waveguides (filters, detectors, passive alignment) with light-grey micro-fluidic channels (sample loading wells, heaters, mixing and electrophoresis channels)

electromagnetic interference (EMI) than their electrical counterparts [86]. The high bandwidth infotainment needs in cars, such as GPS and video systems; and EMI-tolerant needs of aviation have already led to the utilization of polymer optical communication systems for transferring data from electrical sensor arrays distributed throughout the vehicle (monitoring acceleration, exhaust gas analysis, tyre pressure, temperature and fluid levels) to a core processor that can automatically deploy airbags, brakes and drive- or fly-by-light systems [87]. Here, silicon photonics has the potential to converge the in-vehicle communication systems and sensors arrays on a single substrate.

Silicon photonic temperature sensors have been fabricated based on the thermo-optic effect, by monitoring the resonant wavelength shift caused by a change in the effective index of a waveguide with temperature. When implemented in a Bragg grating and Michelson architecture, this has the potential for sensitivity of  $<5 \times 10^{-4} \text{ }^{\circ}\text{C}$  [88,89]. Monitoring the luminescent lifetime of Er-doped silicon waveguides [90] has allowed low-temperature sensing in the range of 40–150 K, and simply monitoring the propagation loss of a 2D silicon slab waveguide is an effective way to monitor the dew point temperature as an alternative to the typically used bulkier chilled mirror hygrometers [91]. Strain sensors can be fabricated based on the elasto-optic effect [92], and a Mach–Zehnder interferometer (MZI) has been fabricated with arms that lie perpendicular to each other on a silicon die. When strain is applied to the die in a preferred direction the elasto-optic effect induces a change in the index of one arm, causing a change in the transmission of the MZI. This has been used to achieve a sensitivity of 0.05 micro strain in an architecture designed for negligible variation with temperature [92]. MZIs have also been used as pressure sensors, where one arm of the MZI is placed on a deformable membrane while the other arm acts as a reference path, again making use of the elasto-optic effect to transform the induced stress into a change of transmission of the MZI. These types of sensor have shown a sensitivity of  $<10 \text{ kPa}$  [93,94]. Accelerometers can be fabricated by placing waveguides on a movable mass sensor and monitoring the change in coupling efficiency to reference waveguides placed on a frame. Acceleration moves the mass sensor, causing the sensor waveguides to become misaligned, resulting in a drop in the coupling efficiency, with a sensitivity of up to 0.1 g [95]. Integrated optical gyroscopes have also been demonstrated with



**Figure 9.21** Cross-section of a waveguide based sensor showing: (a) an evanescent spectroscopy architecture where the shaded evanescent field of the mode interacts with the sample being tested; (b) chemical sensing architecture using a thick-film sensing layer deposited on a silicon waveguide

1°/s accuracy based on a Sagnac interferometer fabricated out of silica, taking advantage of an 80 cm interaction length available in a small footprint for integrated optical components [96]. Finally, the plasma-dispersion effect has been suggested to sense electric fields in a micro-disk and Fabry–Perot-coupled cavity arrangement down to a minimum detectable field of 15  $\mu\text{V/m}$  [97].

#### 9.4.2 Chemical Sensors

Chemical sensing was one of the first applications proposed for integrated optical sensors and now has a multitude of applications, from hazardous gas sensing for homeland security to water contamination in environmental monitoring, and process control in manufacturing plants to breath analysis for health monitoring. Many of these applications need *in situ* or remote monitoring and so drive low-cost, small form factor reliable sensor architectures. Silicon photonic sensors can address this need due to their cost-effective processing, small size, and the integration of multiple sensors onto a single chip.

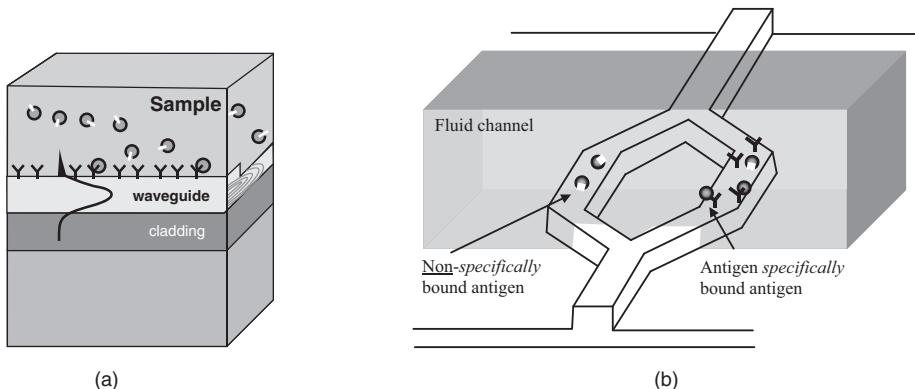
The simplest measurement technique for chemical analysis is conventional transmission spectroscopy, where the absorption spectrum of the sample under test is used to determine its constituents. This is a useful technique for gas sensing because many gases exhibit unique absorption peaks in the mid-infrared [98] and the Beer–Lambert law can be used to map the optical attenuation to the concentration of the gas under test [99]. Spectroscopy can be implemented in a waveguide geometry by using free-standing waveguides and covering them with the sample, be it liquid or gas. Here the evanescent field of the waveguide mode interacts with the sample under test, hence its name evanescent sensing, (Figure 9.21a). Siebert *et al.* [98] have used this method to detect CO<sub>2</sub> by measuring its absorption spectra in the mid-infrared; wavelengths where silicon's low optical absorption renders it superior to other waveguide materials. A final device could integrate the sensing waveguide with the optical spectrometer needed to perform the measurement, reducing both the cost and size of the overall measurement system.

For applications that require higher specificity the silicon waveguide may be coated with a film of chemically sensitive material, such as polysiloxane [100], which specifically interacts with a given chemical species, such as organic solvents (Figure 9.21b). Since the entire evanescent field interacts with the sensing film, this is known as thick-film, or homogeneous sensing. When the chemical under test is preferentially adsorbed by the sensing layer it can either cause a color change which can be detected via a change in the propagation loss of the waveguide [101], or it may cause swelling of the active material, resulting in a change of its refractive index [100]. In the latter case the change in cladding refractive index may be detected using phase sensitive detection. This can be done by inserting the sensitized waveguide into an interferometer, such as Zeeman [102] or Mach-Zehnder [103]; using resonant structures, such as ring resonators [104] or Bragg gratings [105], or input or output grating couplers [106]. Here it must be emphasized that the integrated optical device is acting only as a signal transducer and the specificity is determined solely by the chemistry of the sensing layer. Another important issue is that, in general, volume concentration is the quantity that needs to be measured, whereas these sensors and the biosensors (discussed later) usually detect the surface concentration of the analyte adsorbed into the film. Therefore the surface chemistry and adsorption dynamics of the analyte with the sensing layer are critical parameters as they affect the sensitivity, the specificity and the response time of the complete sensor. Integrated optic chemical sensors have been fabricated that are sensitive to (among others): humidity; organic gases/solvents; acid vapors; pH; metal ions; cadmium; and glucose. Multiplexing arrays of sensor elements enables sensing of multiple different chemicals on a single chip [107].

#### 9.4.3 Biochemical Sensors

Applications for biochemical sensing range from DNA sequencing, to drug development to disease and health monitoring. In general these differ from the chemical sensors described above in that significant sample preparation may be required prior to testing, leading to the development of integrated systems, discussed in the next section. The biochemical sensor layout is similar to the chemical sensing architecture described above (of coating the silicon waveguide with a sensitive layer). Here, the biochemical sensing layer usually consists of receptor molecules immobilized onto the waveguide surface using a suitable passivation layer and may comprise of antibodies specific to a certain antigens, proteins specific to other proteins or single-stranded DNA receptive to their complimentary strands. A key difference between these biochemical sensors and the chemical sensors above is that, in general, the sensing layer here is much thinner, of the order of 1–20 nm, hence its designation as thin film, or surface, sensing; see Figure 9.22(a).

Conventional immunoassays are run in the following manner: antibodies on the bottom of a 96-well plate capture the target antigen, the plate is then washed to remove unbound molecules, then a second, fluorescently labeled, noncompeting, antibody is introduced which binds to the target already captured on the plate. After an additional wash step to remove unbound fluorescent antibody the plate is excited to fluoresce and imaged, with the amount of bound fluorophore imaged in the fluorescent microscope being proportional to the amount of target molecule in the original solution. This type of assay is known as a sandwich assay [108] as the target molecule is sandwiched between a capture antibody on the plate and a fluorescently labeled antibody. Some issues with this assay type are that it is only suitable for



**Figure 9.22** (a) Cross-section of a waveguide-based sensor showing a thin-film biochemical sensing architecture where the antibody sensing layer interacts with only a part of the red evanescent field of the mode; (b) schematic of MZI-based biochemical sensor, antigens coat one arm of the biosensor allowing discrimination of nonspecific antigen binding which occurs on both arms

relatively large target molecules against which two noncompetitive antibodies are available; it takes a long time to perform due to the multiple wash steps needed; and it uses many reagents. The first planar waveguide-based assay improved on this by using the confined nature of the evanescent wave in a slab waveguide to excite only the surface-bound fluorophore rather than that in solution, overcoming the need for most of the above wash steps, hence increasing the speed and throughput of the assay [109].

An alternative approach to the sandwich assay is to directly detect the binding of the target analyte to the receptor by its effect on the effective index of the waveguide. In this case, as proteins have a higher refractive index ( $n \sim 1.4$ ) than water ( $n \sim 1.33$ ), binding of a protein to the surface of the waveguide tends to increase the effective index of the waveguide, causing a phase shift of the light passing through it. This phase shift may be monitored by placing the sensing waveguide in an interferometer. An MZI configuration has the advantage that a reference arm, with no receptor coating, may be used to filter out the effects of nonspecific binding, where target molecules bind directly to the waveguide, causing false-positive results as shown in Figure 9.22(b). Ring resonators and Fabry–Perot resonator structures have shown the higher sensitivity achievable with multi-pass resonant optical cavities [110]. The advantages of this direct detection approach are again a decrease in the assay time and reagent cost, as well as obviating the need for a second noncompetitive antibody to the target, leading to wider applicability. Furthermore, this sensing architecture allows the binding kinetics between the receptor and target to be studied in real time, allowing affinity measurements, a vital step in antibody selection for drug development.

Taking microfabrication of silicon one step further it is possible to confine light in three dimensions using photonic bandgap (PBG) structures. The dispersion of these structures can be engineered to enable low group velocity, increasing the interaction of the sample volume with the probing light beam. By creating a single defect photonic crystal cavity Loncar *et al.* showed that sensing could be carried out at the peak intensity of a photonic crystal laser [111]. This increases the interaction of light with a molecular species by confining the light to smaller and smaller mode volumes centered on the analyte, allowing sensing in very low volumes at

low concentrations. Silicon PBG sensors have shown sensitivity to differences in refractive index of  $\sim 0.002$  [112], while Schmidt *et al.* have demonstrated that silicon nano-cavities can be used to detect discrete numbers of gold nanoparticles [113].

#### 9.4.4 Integrated Lab-on-a-chip

Integration of the above chemical and biochemical sensors with other optical components on a common substrate leads to the conventional integration advantages of lowering costs and footprint size. In this case optical components that may be integrated with a sensor include: photodetectors, taps to monitor excitation power, spectrometers to measure absorption or fluorescence spectra, and wavelength filters to block the excitation light. Beyond the optical integration is the integration of optical waveguides with micro-fluidic components, a combination which has spawned the new field of opto-fluidics [114]. The ability to take a liquid sample, perform some pre-test sample preparation, followed by biochemical reactions with reagents stored on the chip before the assay is performed has enormous advantages in terms of the reduction in quantity of the reagents used, lowered costs, and smaller sample sizes, resulting in faster and more sensitive assays. This is really the concept of a lab-on-a-chip, as shown in Figure 9.20 where a raw sample is placed at the input, and the data are obtained with no lengthy delays or off-chip sample manipulation. The high thermal conductivity of silicon makes it an attractive material for those labs-on-a-chip that involve reactions requiring temperature cycling, e.g. DNA amplification. For example the polymerase chain reaction (PCR) method which is used to chemically amplify DNA requires temperature cycling from  $95^\circ$  to  $55^\circ$  to  $72^\circ\text{C}$  for typically 20–30 cycles before detectable levels are achieved. Conventionally, this temperature cycling is a bottleneck and can take up to 4 h, while shrinking the sample volume down using micro-fluidics on a silicon platform reduces this to tens of minutes [115].

Some examples where high level integration could prove useful for labs-on-a-chip are listed below:

DNA analysis is important for many applications, from drug discovery, to forensic analysis and health monitoring; all of which could take advantage of the cost, footprint, and sample size and assay-time reduction associated with micro-fluidic chips. Hong *et al.* [116] have demonstrated cell isolation, lysis and DNA purification on a micro-fluidic chip, and Lagally *et al.* [115] have demonstrated DNA amplification by PCR and capillary electrophoresis again on a micro-fluidic chip with the aid of a confocal microscope [115]. Silicon photonics integrated with micro-fluidics has the potential to fabricate such systems all together on a single chip integrated with the detection platform. The potential for this is that whole blood could be input onto the chip and its DNA analyzed on a single chip.

Another application area is fluorescence-activated cell sorting (FACS), whereby single cells or particles labeled with fluorescent tags are passed through a micro-fluidic channel to an interrogation region where they are sorted according to their fluorescent and/or light scattering properties. In [117] Fu *et al.* discuss a micro-fabricated FACS system, that miniaturizes the flow chamber, but still uses a complex, and expensive, fluorescence optical microscope for detection. Silicon photonics has the potential to integrate both the flow chamber and fluorescence detection optics onto a single chip using infrared-excitible fluorophores [118,119], dramatically reducing the cost of this technique.

Immunoassays are used in wide range of areas, from pathogen detection in blood to environmental contamination monitoring. Combining micro-fluidics with silicon photonics for this application has the potential to decentralize medical testing and diagnosis from large laboratories to doctor's surgeries or patient's homes, with significant impact of public health monitoring on a global scale [120]. As an example Cesaro-Tediac *et al.* have demonstrated a micro-fluidic chip to detect tumor necrosis containing integrated capillary systems and temperature controller [121]. As with the above examples an optical assay is used in conjunction with a fluorescence microscope, leading to high system costs that could be alleviated with silicon photonics.

## 9.5 Summary and Conclusions

The rate of advancements in silicon photonics over the past few years has been unprecedented, and device performance has been improving to a level where many silicon-based devices are operating at, or near, commercial grade performance at 10Gb/s with many devices demonstrated at 40 Gb/s. With these individual devices now functional, we are starting to enter the next era of silicon photonics where integration of these devices onto a single silicon chip will be the next challenge. As the transition from the vacuum tube to the integrated circuit enabled entirely new devices to be built which could not be produced with individual vacuum tube technology, so too will silicon photonics revolutionize optical communications with new, highly integrated, photonic devices.

In this chapter we have discussed the different applications where silicon photonics may be applied, ranging from long-haul telecommunications down to low-cost chip-to-chip interconnects. In addition, new areas such as radio-over-fiber (RoF) for transmitting wireless connectivity are emerging as a new area where silicon photonics could provide benefits. Furthermore, integrated silicon lasers, detectors, and waveguide-based sensors can be applied to new and emerging applications in the medical, healthcare and diagnostic arenas.

All of the above applications have the common requirement of low cost, small form factor, and high volume manufacturability. This will only be possible if one can deliver these technologies on an integrated silicon platform. We are at the beginning of this new technology adoption and while there is significant work ahead the future of silicon photonic looks very promising.

## Acknowledgements

The authors would like to thank Sean Koehl and Russell Hodgin for assistance in drawing the figures; Mike Morse, Olufemi Dosunmu, Hengju Cheng, Gadi Sarid and Eyal Ginsburg for the SiGe photodetector results; and Ying.-Hao Kuo, Shengbo Xu, Vanessa Sih, Oded Cohen and Omri Raday for the nonlinear and Raman results. In addition the authors would like to specially thank Tom Mader and all the members of the Photonics Technology Team and the fabrication team at Fab 8 at Intel for all their valuable contributions to this work and for stimulating discussions.

## References

- [1] J. Jaussi, B. Casper, M. Mansuri, F. O'Mahony, K. Canagasaby, J. Kennedy and R. Mooney, 'A 20 Gb/s embedded clock transceiver in 90 nm CMOS', In: *Digest IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, 5–9 February, 2006, pp 340–341, 657.
- [2] B. Casper, J. Jaussi, F. O'Mahony, M. Mansuri, K. Canagasaby, J. Kennedy, E. Yeung and R. Mooney, 'A 20 Gb/s forwarded clock transceiver in 90 nm CMOS', In: *Digest IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, 5–9 February, 2006, pp 90–91, 640.
- [3] H. Braunisch, J. E. Jaussi, J. A. Mix, M. B. Trobough, B. D. Horine, V. Prokofiev, D. Lu, R. Baskaran, P. C. H. Meier, D.-H. Han, K. E. Mallory and M. W. Leddige, 'Flex-circuit chip-to-chip interconnects', In: *Proc. IEEE Electronic Components and Technology Conf. (ECTC)*, San Diego, CA, 30–May 2 June 2006, pp 1853–1859.
- [4] M. Jutzi, M. Berroth, G. Wohl, M. Oehme and E. Kasper, 'Ge-on-Si Vertical Incidence Photodiodes With 39 GHz Bandwidth', *IEEE Photon. Technol. Lett.*, **17** (7), 2005, 1510–12.
- [5] J. Liu, J. Michel, W. Giziewicz, D. Pan, K. Wada, D. Cannon, S. Jonghammanurak, D. Danielson, L. C. Kimerling, J. Chen, F. O. Ilday, F. X. Kartner and J. Yasaitis, 'High Performance, tensile-strained Ge  $p-i-n$  photodetectors on a Si platform', *Appl. Phys. Lett.*, **87**(1), 2005, 103501–03.
- [6] M. Morse, O. Dosunmu, G. Sarid and Y. Chetrit, 'Performance of Ge-on-Si  $p-i-n$  Photodetectors for Standard Receiver Modules', *Phot. Tech. Lett.*, **18**(21), 2006, 2442–2442.
- [7] Ansheng Liu, Ling Liao, Doron Rubin, Hat Nguyen, Berkehan Ciftcioglu, Yoel Chetrit, Nahum Izhaky and Mario Paniccia, 'High-speed optical modulation based on carrier depletion in a silicon waveguide', *Opt. Express*, **15** (2), 2007, 660–668.
- [8] Huangm C. Gunn, G. Li, Y. Liang, S. Mirsaldi, A. Naraimha and T. Pinguet, 'A 10Gb/s photonics modulator and WDM Mux/Demux integrated with electronics in 0.13  $\mu\text{m}$  SOI CMOS', *Proc. 2006 IEEE Int. Solid-State Circuits Conference*, 2006, pp 24–25.
- [9] F. Y. Gardes, G. T. Reed, N. G. Emerson and C. E. Png, 'A sub-micron depletion-type photonic modulator in Silicon on Insulator', *Optics Express*, **13**, 2006, 8845–8853.
- [10] T. Kawanishi, T. Sakamoto, T. Miyazaki, M. Izutsu, T. Fujita, S. Mori, K. Higuma and J. Ichikawa, 'High Speed Optical DQPSK and FSK modulation using integrated Mach-Zehnder interferometers', *Opt. Express*, **14** (10), 2006, 4469–4478.
- [11] G. Charlet and S. Bigo, 'Upgrading WDM submarine systems to 40-Gbit/s channel bitrate', *Proc. IEEE 94* (5), 2006, 935–951.
- [12] C. Xu, X. Hong and W. Huang, 'Design optimization of integrated BiDi Triplexer optical filter based on planar lightwave circuit', *Opt. Express* **14**, 2006, 4675–4686.
- [13] S. Kim, S. Park, J. Moon and H. Lee, 'A low crosstalk Design of 1.25Gbps Optical Triplexer Module for FTTH systems', *ETRI Journal*, **28** (1), 2006, 9–16.
- [14] J. Held, J. Bautista and S. Koehl, 'From a Few Cores to Many: A Tera-scale computing research overview', [www.intel.com/go/terascale](http://www.intel.com/go/terascale)
- [15] A. W. Fang, R. Jones, H. Park, O. Cohen, O. Raday, M. J. Paniccia and J. E. Bowers, 'Integrated AlGaInAs-silicon evanescent race track laser and photodetector', *Optics Express*, **15** (5), 2315–2322; A. W. Fang, H. Park, O. Cohen, R. Jones, M. J. Paniccia and J. E. Bowers, 'Electrically pumped hybrid AlGaInAs-silicon evanescent laser', *Opt. Express*, **14**, 2006, 9203–9210.
- [16] T. Darcie and G. Bodeep, 'Lightwave subcarrier CATV transmission systems', *IEEE Trans. MTT*, **38**, 1990, 524–533.
- [17] W.L. Glomb Jr., 'Fiber optic links for antenna remoting', *Proc. SPIE*, **1703**, 1992, 523–527.
- [18] H. Ogawa, D. Polifko and S. Banba, 'Millimeter-wave fiber optic systems for personal radio communications', *IEEE Trans. MTT*, **40**, 1992, 2285–2293.
- [19] H. Al-Raweshidy and S. Komaki (eds), *Radio over Fiber Technologies for Mobile Communication Networks*, Artech House, London, Boston, 2002.
- [20] D. Novak *et al.*, 'Fiber radio systems', *Proc. ECOC*, Stockholm, Sweden, 2004.
- [21] D. Novak, A. Nirmalathas, C. Lim, C. Marra and R.B. Waterhouse, 'Fiber-radio: Challenges and possible solutions', *Proc. Microwave Photonics 2003*, 2003, pp. 49–54.
- [22] M. Nazarathy, J. Berger, A. J. Ley, I. M. Levi and Y. Kagan, 'Progress in externally modulated AM CATV transmission systems', *IEEE J. Lightwave Technol.*, **11**, 1993, 82–105.
- [23] A. Seeds, 'Microwave photonics', *IEEE Trans. MTT*, **50**, 2002, 977–887.
- [24] W. S. C. Chang (ed.), *RF Photonic Technology in Optical Fiber Link*, Cambridge University Press, 2002.

- [25] C. Cox, E. I. Ackerman, G. E. Betts and J. L. Prince, ‘Limits on the performance of RF-over-fiber links and their impact on device design’, *IEEE Trans. MTT*, **54**, 2006, 906–920.
- [26] S. Iezekiel (ed.), *Microwave Photonics*, John Wiley & Sons Inc., 2005.
- [27] S. Pappert, C. Sun, R. Orazi and T. Weiner, ‘Microwave fiber optic links for shipboard antenna applications’, *Proc. 2000 IEEE International Conference on Phased Array Systems and Technology*, Piscataway, NJ, USA, 2000, pp 345–348.
- [28] D. Baldwin, ‘Dynamic range measurements comparing DFB laser links to Mach-Zehnder modulator optical links for wide-band microwave transmission’, *Proc. Photon. Syst. Antennas Applicat. Conf.*, Monterey, CA, 1993, pp 51–54.
- [29] M. M. Sisto, S. LaRochelle and L. A. Rusch, ‘Carrier-to-noise ratio optimization by modulator bias control in radio-over-fibre links’, *IEEE Photon. Technol. Lett.*, **18**, 2006, 1840–1842.
- [30] T. Kuri, K. Kitayama and Y. Takahashi, ‘60 GHz band full duplex radio-on fiber system using two-RF-port electroabsorption transceiver’, *IEEE Photon. Technol. Lett.*, **12**, 2000, 419–421.
- [31] V. J. Urick, M. S. Rogge, P. F. Kanapp, L. Swingen and F. Bucholtz, ‘Wide band predistortion linearization for externally modulated long-haul analog fiber optic link’, *IEEE Trans. MTT*, **54**, 2006, 1458–1463.
- [32] Y. Chiu, B. Jalali, S. Garner and W. Steier, ‘Broad-band electronic linearizer for externally modulated analog fiber-optic links’, *IEEE Photon. Technol. Lett.*, **11**, 1999, 48–50.
- [33] K. Williams, L. Nicholes and E. Esman, ‘Photodetector nonlinearity limitations on a high dynamic range 3 GHz fiber optic link’, *J. Lightwave Technol.*, **16**, 1998, 192–199.
- [34] H. Ito, T. Furuta, T. Ito, Y. Muramoto, K. Tsuzuki, K. Yoshino and T. Ishibashi, ‘W-band uni-traveling-carrier photodiode module for high power photonic millimeter-wave generation’, *Electron. Lett.*, **38**, 2002, 1376–1377.
- [35] M. S. Islam, T. Jung, T. Itoh, M. C. Wu, A. Nespolo, D. L. Silvco and A. Y. Cho, ‘High power and highly linear monolithic integrated distributed balanced photodetectors’, *IEEE J. Lightwave Technol.*, **20**, 2002, 285–295.
- [36] J. Basak and B. Jalali, ‘Photodetector linearization using adaptive electronic post-distortion’, *Proc. 2005 OFC*, vol. 4, 2005, p. 101.
- [37] L. M. Franca-Neto, R. Eline and B. Balvinder, ‘Fully integrated CMOS radios from RF to millimeter wave frequencies’, *Intel Technology Journal*, **8**, 2004, 241–258, and references therein.
- [38] M. Paniccia, ‘Silicon Photonics’, *Proc. 2006 IEEE MTT-S International Microwave Symposium, TFC-6*, San Francisco, USA, 2006.
- [39] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, ‘Micrometre-scale silicon electro-optic modulator’, *Nature*, **435**, 2005, 325–327.
- [40] A. Liu, H. Rong, M. Paniccia, O. Cohen and D. Hak, ‘Net optical gain in a low loss silicon-on-insulator waveguide by stimulated Raman scattering’, *Opt. Express*, **12**, 2004, 4261–4267.
- [41] T. K. Liang and H. K. Tsang, ‘Efficient Raman amplification in silicon-on-insulator waveguides’, *Appl. Phys. Lett.*, **85**, 2004, 3343–3345.
- [42] O. Boyraz and B. Jalali, ‘Demonstration of 11dB fiber-to-fiber gain in a silicon Raman amplifier’, *IEICE Elect. Express*, **1**, 2004, 429–434.
- [43] V. Raghunathan, O. Boyraz and B. Jalali, ‘20 dB On-off Raman amplification in silicon waveguides’, *CLEO 2005, CMU1*, Baltimore, MD, May 2005.
- [44] Q. Xu, V. Almeida and M. Lipson, ‘Time-resolved study of Raman gain in highly confined silicon-on-insulator waveguides’, *Opt. Express*, **12**, 2004, 4437–4442.
- [45] X. Chen, N. C. Panoiu and R. M. J. Osgood, ‘Theory of Raman-Mediated Pulsed Amplification in Silicon-Wire Waveguides’, *IEEE J. Quant. Electron.*, **42**, 2006, 160–170.
- [46] Jones, R *et al.* ‘Net continuous wave optical gain in a low loss silicon-on-insulator waveguide by stimulated Raman scattering’, *Opt. Express*, **13**, 2005, 519–525.
- [47] Y. Liu and H. K. Tsang, ‘Nonlinear Absorption and Raman Gain in Helium Ion Implanted Silicon’, *Opt. Lett.*, **31**, 2006, 1714–1716.
- [48] Sasan Fathpour, Ozdal Boyraz, Dimitri Dimitropoulos and Bahram Jalali, ‘Demonstration of CW Raman Gain with Zero Electrical Power Dissipation in *p-i-n* Silicon Waveguides’, *CLEO 2006, Long Beach, CMK3*.
- [49] O. Boyraz and B. Jalali, ‘Demonstration of a silicon Raman laser’, *Opt. Express*, **12**, 2004, 5269–5273.
- [50] H. Rong *et al.* ‘An all-silicon Raman laser’, *Nature*, **433**, 2005, 292–294.
- [51] H. Rong *et al.* ‘A continuous-wave Raman silicon laser’, *Nature*, **433**, 2005, 725–728.
- [52] H. Rong, Y. -H. Kuo, S. Xu, A. Liu, R. Jones, M. Paniccia, O. Cohen and O. Raday, ‘Monolithic integrated Raman silicon laser’, *Opt. Express*, **14**, 2006, 6705–6712.

- [53] A. Liu *et al.* ‘Optical Amplification and Lasing by Stimulated Raman Scattering in Silicon Waveguides’, *J. Lightwave Tech.*, **24**, 2006, 1440–1455.
- [54] H. Rong *et al.* ‘Characterization of a silicon Raman laser’, *Proc SPIE*, **5931**, 2006, 59310R 1–9.
- [55] T. K. Liang and H. K. Tsang, ‘Role of free carriers from two-photon absorption in Raman amplification in silicon-on-insulator waveguides’, *Appl. Phys. Lett.*, **84**, 2004, 2745–2747.
- [56] H. Rong *et al.* ‘Raman gain and nonlinear optical absorption measurement in a low loss silicon waveguide’, *Appl. Phys. Lett.*, **85**, 2004, 2196–2198.
- [57] R. Claps, V. Raghunathan, D. Dimitropoulos and B. Jalali, ‘Role of nonlinear absorption on Raman amplification in Silicon waveguides’, *Opt. Express*, **12**, 2004, 2774–2780.
- [58] R. A. Soref and P. J. Lorenzo, ‘All-silicon active and passive guided-wave components for  $\lambda = 1.3$  and  $1.6\mu\text{m}$ ’, *IEEE J. Quantum Electron.* **QE-22**, 1986, 873–879.
- [59] H. Rong, Y.-H. Kuo, S. Xu, O. Cohen, O. Raday and M. Paniccia, ‘Recent development on silicon-based Raman lasers and amplifiers’, *Proc. SPIE*, **6389**, 2006, 638904.
- [60] V. Sih, S. Xu, Y. -H. Kuo, H. Rong, M. Paniccia, O. Cohen and O. Raday, ‘Raman amplification of 40 Gb/s data in low-loss silicon waveguides’, *Opt. Express*, **15**, 2007, 357–362.
- [61] <http://www.cfa.harvard.edu/HITRAN/docs.html>
- [62] X. Yang and C. W. Wong, ‘Design of photonic band gap nanocavities for stimulated Raman amplification and lasing in monolithic silicon’, *Opt. Express*, **13**, 2005, 4723–4730.
- [63] J. G. Naeini and K. Ahmad, ‘Raman fiber laser with two parallel couplers’, *Opt. Eng.*, **44**, 2005, 064203 1–4.
- [64] T. J. Kippenberg, S. M. Spillane, B. Min and K. J. Vahala, ‘Theoretical and Experimental Study of Stimulated and Cascaded Raman Scattering in Ultrahigh-Q Optical Microcavities’, *IEEE J. Sel. Top. Quant. Electron.*, **10**, 2004, 1219–1228.
- [65] V. Raghunathan, R. Claps, D. Dimitropoulos and B. Jalali, ‘Wavelength conversion in silicon using Raman induced four-wave mixing’, *Appl. Phys. Lett.*, **85**, 2004, 34–36.
- [66] R. L. Espinola, J. I. Dadap, R. M. Osgood, Jr., S. J. McNab and Y. A. Vlasov, ‘C-band wavelength conversion in silicon photonic wire waveguides’, *Opt. Express*, **13**, 2005, 4341–4349.
- [67] H. Fukuda, K. Yamada, T. Shoji, M. Takahashi, T. Tsuchizawa, T. Watanabe, J. Takahashi and S. Itabashi, ‘Four-wave mixing in silicon wire waveguides’, *Opt. Express*, **13**, 2005, 4629–4637.
- [68] K. Yamada *et al.*, ‘All-optical efficient wavelength conversion using silicon photonic wire waveguide’, *IEEE Photon. Technol. Lett.* **18**, 2006, 1046–1048.
- [69] H. Rong, Y.-H. Kuo, A. Liu, M. Paniccia and O. Cohen, ‘High efficiency wavelength conversion of 10 Gb/s data in silicon waveguides’, *Opt. Express*, **14**, 2006, 1182–1188.
- [70] M. H. Chou, J. Hauden, M. A. Arbore and M. M. Fejer, ‘1.5-mm-band wavelength conversion based on difference-frequency generation in LiNbO<sub>3</sub> waveguides with integrated coupling structures’, *Opt. Lett.*, **23**, 1998, 1004–1006.
- [71] Y. L. Lee, C. Jung, Y. Noh, M. Y. Park, C. C. Byeon, D. Ko and J. Lee, ‘Channel-selective wavelength conversion and tuning in periodically poled Ti:LiNbO<sub>3</sub> waveguides’, *Opt. Express*, **12**, 2004, 2649–2655.
- [72] Y. -H. Kuo, H. Rong, V. Sih, S. Xu, M. Paniccia and O. Cohen, ‘Demonstration of wavelength conversion at 40 Gb/s data rate in silicon waveguides’, *Opt. Express*, **14**, 2006, 11721–11726.
- [73] M. A. Foster, A. C. Turner, J. E. Sharping, B. S. Schmidt, M. Lipson and A. L. Gaeta, ‘Broad-band optical parametric gain on a silicon photonic chip’, *Nature*, **441**, 2006, 960–963.
- [74] V. R. Almeida and M. Lipson, ‘Optical bistability on a silicon chip’, *Opt. Lett.*, **29**, 2004, 2387–2389.
- [75] M. Notomi, A. Shinya, S. Mitsugi, G. Kira, E. Kuramochi and T. Tanabe, ‘Optical bistable switching action of Si high-Q photonic-crystal nanocavities’, *Opt. Express*, **13**, 2005, 2678–2687.
- [76] P. Barclay, K. Srinivasan and O. Painter, ‘Nonlinear response of silicon photonic crystal microresonators excited via an integrated waveguide and fiber taper’, *Opt. Express*, **13**, 2005, 801–820.
- [77] G. Priem, P. Dumon, W. Bogaerts, D. Van Thourhout, G. Morthier and R. Baets, ‘Optical bistability and pulsating behaviour in Silicon-On-Insulator ring resonator structures’, *Opt. Express*, **13**, 2005, 9623–9628.
- [78] T. K. Liang *et al.*, ‘Ultrafast all-optical switching by cross-absorption modulation in silicon wire waveguides’, *Optics Express*, **13**, 2005, 7298–7303.
- [79] T. K. Liang *et al.*, ‘Nonlinear Self-Distortion of Picosecond Optical Pulses in Silicon Wire Waveguides’, *CLEO 2006*, Long Beach, JThC44.
- [80] Y. Okawachi *et al.*, ‘All-optical slow-light on a photonic chip’, *Opt. Express*, **14**, 2006, 2317–2322.
- [81] Z-Q Zou, X. Chen, Q-H Jin, M-S. Yang and J-L. Zhao, ‘A novel miniaturized PCR multi-reactor array fabricated using flip-chip bonding techniques’, *J. Micromech. Microeng.*, **15**, 2005, 1476–1481.

- [82] X. Liu, R. K. Y. Fu, P. K. Chu and C. Ding, ‘Mechanism of apatite formation on Hydrogen plasma-implanted single-crystal silicon’, *Appl. Phys. Lett.*, **85** (16) 2004, 3623–3625.
- [83] A. Manz, N. Graber and H. M. Widmer, ‘Miniatuerized total chemical analysis systems: a novel concept for chemical sensing’, *Sens. Act.*, **B1**, 1990, 244–248.
- [84] B. H. Weigl, R. L. Bardell and C. R. Cabrera, ‘Lab-on-a-chip for drug discovery’, *Adv. Drug. Del. Rev.*, **55**, 2003, 349–377.
- [85] M. Esashi and T. Ono, ‘From MEMS to nanomachine’, *J. Phys. D: Appl. Phys.*, **38**, 2005, R223–R230.
- [86] K. Benissa and A. Nathan, ‘Silicon anti-resonant reflecting optical waveguide for sensor applications’, *Sens. Act.*, **A 65**, 1998, 33–44.
- [87] J. Lambert, ‘Resonant-cavity LEDs fire up in-car networks’, *Compound Semiconductor Magazine*, March 2006.
- [88] S-L Tsao and P-C Chun, ‘An SOI Michelson interferometer sensor with waveguide Bragg reflective gratings for temperature monitoring’, *Microwave and Optical Technology Letters*, **30** (5), 2001, 321–322.
- [89] S-L Tsao and S-G Lee, ‘Design and analysis of a new silicon-on-insulator waveguide Michelson interferometer sensor’, *Opt. Quant. Elect.*, **36**, 2004, 309–320.
- [90] A. K. Kewell, G. T. Reed and F. Namavar, ‘Integrated temperature sensor in Er-doped silicon’, *Sens. Act.*, **A 65**, 1998, 160–164.
- [91] P. R. Stoddart, J. B. Pearce and A. P. Mazzolini, ‘Evanescently coupled dewpoint sensor based on a silicon waveguide’, *Sens. Act.*, **A 128** (2), 2006, 225–229.
- [92] G. N. Pearson and P. E. Jessop, ‘Temperature-insensitive Mach–Zehnder interferometric strain sensor in silicon-on-insulator’, *Proc. SPIE*, **4997**, 2003, 242–249.
- [93] K. Benissa and A. Nathan, ‘IC compatible optomechanical pressure sensors using Mach–Zehnder Interferometry’, *IEEE Trans. Elect. Dev.*, **43** (9), 1996, 1571–1582.
- [94] D. Peters, K. Fischer and J. Muller, ‘Integrated optics based on silicon oxynitride thin films deposited on silicon substrates for sensor applications’, *Sens. Act.*, **A 25–27**, 1991, 425–431.
- [95] A. Llobera, J. A. Plaza, I. Salinas, J. Berganzo, J. Garcia, J. Esteve and C. Dominguez, ‘Technological aspects on the fabrication of silicon-based optical accelerometer with ARROW structures’, *Sens. Act.*, **A 110**, 2004, 395–400.
- [96] P. Mottier and P. Pouteau, ‘Solid state optical gyrometer integrated on silicon’, *Elect. Lett.*, **22** (23), 1997, 1975–1977.
- [97] V. M. N. Passaro and F. Deleonardis, ‘Modeling and design of a novel high-sensitivity electric field silicon-on-insulator sensor based on whispering-gallery-mode resonator’, *IEEE J. sel. Top. Quant. Elect.*, **12** (1), 2006, 124–133.
- [98] R. Siebert and J. Muller, ‘Infrared integrated optical evanescent field sensor for gas analysis Part II. Fabrication’, *Sens. Act.*, **A 119**, 2005, 584–592.
- [99] K. B. Mogensen, J. El-Ali, A. Wolff and J. P. Kutter, ‘Integration of polymer waveguides for optical detection in microfabricated chemical analysis systems’, *Appl. Opt.*, **42** (19), 2003, 4072–4079.
- [100] N. Fabricius, G. Gauglitz and J. Ingenhoff, ‘A gas sensor based on an integrated Mach–Zehnder inetrferomter’, *Sens. Act.*, **B 7**, 1992, 672–676.
- [101] T. Koster and P. Lambeck, ‘An integrated optical platform for absorptive sensing of chemical concentrations using chemo-optical monolayers’, *Meas. Sci. Technol.*, **12**, 2002, 1230–1238.
- [102] K. M. Grace, K. Shrouf, S. Honkanen, P. Ayras, P. Katila, M. Leppihalme, R. G. Johnstone, X. Yang, B. Swanson and N. Peyghambarian, ‘Waveguide Zeeman interferometry for thin-film chemical sensors’, *Elect. Lett.*, **33** (19), 1997, 1651–1652.
- [103] G. Gauglitz and J. Ingenhoff, ‘Integrated optical sensors for halogenated and non-halogenated hydrocarbons’, *Sens. Act.*, **B 11**, 1993, 207–212.
- [104] C-Y Chao and L. J. Guo, ‘Biochemical sensors based on polymer microrings with sharp asymmetrical resonance’, *Appl. Phys. Lett.*, **83** (8), 2003, 1527–1529.
- [105] T. M. Butler, E. Igata, S. J. Sheard and N. Blakie, ‘Integrated optical Bragg-grating-based chemical sensor on a curved input edge waveguide structure’, *Opt. Lett.*, **24** (8), 1999, 525–527.
- [106] W. Lukosz and K. Teinfenthaler, ‘Sensitivity of integrated optical grating and prism couplers as (Bio)Chemical sensors’, *Sens. Act.*, **15**, 1988, 273–284.
- [107] R. R. Smardzewski, ‘Multi-element optical waveguide sensor: general concept and design’, *Talanta*, **35** (2), 1988, 95–101.
- [108] A. N. Sloper, J. K. Deacon and M. T. Flanagan, ‘A Planar Indium Phosphate Monomode Waveguide Evanescent Field Immunosensor’, *Sens. Act.*, **B1**, 1990, 589–591.

- [109] T. E. Hirschfield, ‘Total reflection fluorescence’, *Can. J. Spectrosc.*, **10**, 1965, 128.
- [110] S. Blair and Y. Chen, ‘Resonant-enhanced evanescent-wave fluorescence biosensing with cylindrical optical cavities’, *Appl. Opt.*, **40** (4), 2001, 570–582.
- [111] M. Loncar, A. Scherer and Y. Qiu, ‘Photonic crystal laser sources for chemical detection’, *Appl. Phys. Lett.*, **82** (26), 2003, 4648–4650.
- [112] E. Chow, A. Grot, L. W. Mirkarimi, M. Sigalas and G. Girolami, ‘Ultracompact biochemical sensor built with two-dimensional photonic crystal microcavity’, *Opt. Lett.*, **29** (10), 2004, 1093–1095.
- [113] B. Schmidt, V. Almeida, C. Manolatou, S. Preble and M. Lipson, ‘Nanocavity in a silicon waveguide for ultrasensitive nanoparticle detection’, *Appl. Phys. Lett.*, **85** (21), 2004, 4854–4856.
- [114] D. Psaltis, S. R. Quake and C. Yang, ‘Developing optofluidic technology through the fusion of microfluidics and optics’, *Nature*, **442**, 2006, 381–386.
- [115] E. T. Lagally, C. A. Emrich and R. A. Mathies, ‘Fully integrated PCR-capillary electrophoresis microsystem for DNA analysis’, *Lab chip*, **1**, 2001, 102–107.
- [116] J. W. Hong, V. Studer, G. Hang, W. F. Anderson and S. R. Quake, ‘A nanoliter-scale nucleic acid processor with parallel architecture’, *Nature Biotech.*, **22**, 2004, 435–439.
- [117] A. Y. Fu, C. Spence, A. Scherer, F. H. Arnold and S. R. Quake, ‘A microfabricated fluorescence-activated cell sorter’, *Nature Biotech.*, **17**, 1999, 1109–1111.
- [118] E. H. Sargent, ‘Infrared quantum dots’, *Adv. Mat.*, **17**, 2005, 515–522.
- [119] A. Y. Fu, C. Spence, A. Scherer, F. H. Arnold and S.R. Quake, ‘A microfabricated fluorescence-activated cell sorter’, *Nature Biotech.*, **17**, 1999, 1109–1111.
- [120] P. Yager, T. Edwards, E. Fu, K. Helton, K. Nelson, M. R. Tam and B. H. Weigl, ‘Microfluidic diagnostic technologies for global public health’, *Nature*, **442**, 2006, 412–418.
- [121] S. Cesaro-Tedic, G. Dernick, D. Juncker, G. Buurman, H. Kropshofer, B. Michel, C. Fattinger and E. Delamarche, ‘High sensitivity miniaturized immunoassays for tumor necrosis factor  $\alpha$  using microfluidic systems’, *Lab Chip.*, **4**, 2004, 563–569.

# Index

- Absorption 97, 149  
Absorption coefficient 105  
Absorption edge 97, 150  
Aluminum Arsenide 193  
Analogue to digital converter (A/D) 84  
Anti-Resonant Reflecting Optical Waveguide (ARROW) 41  
Anti-Stokes Waves 162, 170  
Anti-reflection coating 31–32  
Application Specific Photonic Integrated Circuit (ASPIC) 86  
Applications 297–320  
Asymmetric waveguide 16  
Avalanche breakdown 199
- Band structure 147–148, 192  
Bandgap energy 191–193  
Bandgap, direct 148, 192  
Bandgap, indirect 148, 192  
Bandwidth 106, 115–116, 118–120, 124–125, 129, 138, 141, 221–225  
Beam Propagation Method (BPM) 25, 234, 236, 240  
Beat Length 238  
Beer Lambert Law 316  
Bragg Grating 112, 117, 250–262, 317  
Bragg reflections 60  
Bragg Reflector 41, 42  
Buried Waveguides 23
- Carrier concentration 105, 125, 127  
Carrier Lifetime 165–166, 173  
Chemical Vapour deposition (CVD) 34, 43, 67, 70, 213, 275, 279–280
- CMOS 2–12, 44, 67, 70, 72, 80, 96, 115, 118–119, 135, 139, 147, 182, 191, 206–207, 215, 225, 269–285, 290–291, 294  
CMOS Process selection 272–275  
Code Division Multiple Access (CDMA) 4–5  
Communications 4, 6, 28, 36, 97, 101, 109, 277, 279, 289, 297–307  
Conformal transformation method 230  
Coupled Mode Theory (CMT) 234, 236, 247, 251, 254  
Coupling efficiency 32, 35, 77, 178, 214–215, 235, 260–262, 279, 315  
Coupling Loss 34, 173, 217
- Deep reactive ion etching (DRIE) 31, 71  
Defect Engineering 191, 220  
Depletion 2, 7, 98–99, 101, 105, 108, 110, 125, 131, 164–165, 167–168, 195–197, 199–203, 217, 272  
Depletion Region 167–168, 195–196, 198–203, 217, 272  
Detector 1, 3, 6, 9, 59, 80, 192–225, 258–259, 282–285, 291, 300–302, 306 (229 occurrences)  
Detector Noise 204–205  
Devices 1, 3, 10, 24, 38, 44, 47–48, 60, 62–64, 74, 80, 96, 106, 117, 120, 147, 153, 157, 181–182, 207, 209, 229–264, 282–284 (304 occurrences)  
DIFET 108  
Digital Optical Switch (DOS) 244  
Diode Equation 199  
Directional Coupler 3, 27, 34, 96, 233–237, 308

- Dispersion 11, 40, 49, 51–53, 56–57, 61, 67, 75, 77, 79–80, 82–84, 99, 101, 104, 106, 108, 113, 121, 128, 130, 134–135, 138, 171, 244, 247, 249, 253, 256–257, 263, 302, 306, 316, 318
- Drude-Lorenz equation 97, 101
- Dual Grating Assisted Directional Coupler (DGADC) 34–35
- Effective Carrier Lifetime 168
- Effective Refractive Index, N 19, 32, 57, 108, 238, 245, 251–252, 254 256, 263
- Electro-absorption 2, 135, 301, 306
- Electroluminescence 153
- Electron Beam Lithography 61–62
- Electronic and Photonic Integrated Circuit (EPIC) 1, 7–8, 285, 292
- Electro-optical switches 107
- Epitaxy 9, 214, 280–281
- Erbium Doping 80, 153, 156–159
- Etching 30–31, 33, 37–39, 43, 49, 63–67, 70–72, 85, 155, 214, 232, 237, 241–242, 251, 253–254, 258, 262, 273, 275, 278
- Evanescence field 31, 260, 316–318
- Eye Diagram 128–130, 283–284, 289, 301, 313
- Fabrication 60–72
- Fabry-Perot Resonator 229, 255–259, 318
- Fast Fourier Transform 59
- Fermi-Dirac Function 149, 197
- Fibre To The Home (FTTH) 302
- Finesse 74, 257–259, 261
- Finite Difference Time Domain (FDTD) 22, 53–59, 77, 79–82, 214, 234, 247
- Finite Element Method (FEM) 22, 25, 54, 247
- Floquet's Theorem 51
- Fourier Transform 59, 81, 104, 216
- Franz-Keldysh Effect 96, 98–101, 138
- Free Carrier Absorption (FCA) 97, 152, 158–159, 165–166, 172–173, 180, 186, 308–309
- Free Spectral Range (FSR) 249, 257
- Free standing waveguides 36–40, 316
- Fully depleted transistor technology 273–274
- Gallium Arsenide 95, 147, 157, 193
- Germanium 3–5, 9, 15, 135, 147, 178, 193, 203, 209, 213–217, 272, 279–280, 285
- GeSi 138, 140, 178–179
- Grating coupler 35, 317
- Grey scale 31–32
- Group IV Photonics 9, 213, 215
- Heisenberg Uncertainty Principle 153
- Heterogeneous Integration 4, 159, 207
- Hollow Waveguides 36, 40–43
- Holographic Lithography 62–63, 68
- Hybrid Integration 5, 159–160, 191, 269–270
- IMPATT diode 110
- Implants and activation 277
- Indium Arsenide 193
- Indium Phosphide 95, 147, 193
- Inductively coupled plasma (ICP) 33, 85
- Infra Red (IR) 43
- Insertion loss 34–35, 75, 112, 131, 223, 277, 285, 290, 301
- Integrated RF Oscillator 289–292
- Integration 2–7, 9–12, 15, 54, 67–68, 80, 95–96, 104, 114, 117, 128, 135, 138, 159–160, 191, 206–207, 209, 212–217, 269–294
- Integration, back end 271–272
- Integration, front end 271
- Interconnects 4, 6, 67, 72, 80, 141, 214, 297–307
- Inverted tapers 32–34
- I-V Characteristics 184–185
- Kerr Effect 96, 98–101, 139, 170
- Kramers-Kronig relation 101, 104
- Lab-on-a-chip 4, 319–320
- Layer by layer Lithography 68 – 69
- Light Emitting Diode (LED) 95, 157
- Lithium Niobate 95, 99, 301, 312
- Local Oscillator 289
- $L_\pi V_\pi$  133
- Mach-Zehnder Interferometer (MZI) 75, 109, 112, 115, 118, 120, 125, 127, 129, 137, 139, 245–250, 285, 290, 306, 315
- Maxwell's Equations 15–17, 50–51, 234
- Microcavities 49, 73–74, 182
- Micro-Electro-Mechanical-Systems (MEMS) 147, 314
- Mid Wave IR 180–182

- Minimum bend radius 232  
Minority Carrier Lifetime 110, 194  
Mode Converter 31–32  
Mode Field Patterns 33  
Mode mismatch loss 34  
Modulation depth 106, 108, 112, 118, 125, 128, 135–136, 141, 252, 259  
Modulation Mechanisms 97–98, 249  
Modulator 6–9, 28, 49, 75, 95–142, 178, 207, 224, 241, 249, 258, 263, 270, 273, 286, 301, 304, 306 (289 occurrences)  
Monolithic Integration 9, 221, 270–271  
Moore’s Law 10, 80, 183, 299–300, 303  
MOS Capacitor 119, 125, 249  
MOSFET 5, 108  
Multi-mode 8, 21, 27–28, 30, 229, 277  
Multimode Interference Coupler (MMI) 237–242  
MZI, Asymmetric 245  
MZI, relative power transmission 247  
MZI, Symmetric 245  
MZI, unbalanced 245  
  
Nano-cavities 57–60  
Nanophotonics 5, 10–12  
Novel Waveguide structures 15  
  
Octofluorocyclobutane 65  
Optical add-drop filter/multiplexer 75  
Optical Amplifier 3, 96, 147, 178, 183, 290, 297, 307–308  
Optical Bus 80–81  
Optical Gain 148–153  
Optical Interconnects 4, 6, 72, 80, 141, 214, 264, 297, 299–300, 303  
Optical taps 242  
Opto-Electronic Integrated Circuit (OEIC) 1–12  
  
Partially depleted transistor technology 274  
Perfectly Matched Layer (PML) 55, 230  
Phase mismatch 170–171, 174  
Phase modulator 108–110, 114, 116, 118, 121  
Phase Shift 106–107, 113, 116, 120–121, 125–128, 133–135, 235, 243, 246–247, 249, 292, 301, 318  
Phonon 11, 48, 148–150, 162–163, 170–173, 178–179, 192–194  
  
Photodetector 138, 140, 147, 168, 180, 183, 186, 192–225, 300–302, 306(108 occurrences)  
Photodiode 200–202  
Photolithography 31, 43, 61, 68–70, 241  
Photonic Bandgap 47–60, 312, 318  
Photonic Crystals (PhC) 11, 47–85, 135–136 (147 occurrences)  
Photonic Crystals, doping 57  
Photonic Integrated Circuit (PIC) 1, 7, 35, 72–75, 237  
p-i-n diode 111–112, 118, 124, 135, 137–138, 177, 201–202, 217, 259, 263, 277, 308–309, 312  
p-i-n modulator 113, 116  
Planar Waveguides 15–22  
Plane Wave Expansion Method (PWM) 51–52, 54, 80  
Plasma Dispersion Effect 99, 101, 106, 108, 113, 121, 128, 134, 138, 244, 247, 249, 263, 316  
Plasmo-electronic integrated circuit (PEIC) 12  
Plasmon 11–12, 36  
Plasmonic OEIC (POIEC) 12  
PN Junction 4  
p-n junction 4, 125, 156–157, 165–169, 177–178, 183, 195–200  
Pockels Effect 95, 98–99  
Polarisation Dependent Loss (PDL) 119  
Polarisation Independence (PI) 25–26, 30, 131, 133  
Polarisation, TE 18–19, 24–29, 34, 39, 119, 134, 170, 216, 249–250, 262  
Polarisation, TM 18–19, 24–29, 34, 39, 119, 134, 170, 216, 249–250, 262  
Polymethylmethacrylate (PMMA) 66  
Polysilicon 118, 120, 275–276  
Population inversion 152–153  
Poynting Vector 21  
Propagation constant  $\beta$  18–21, 24, 30, 135, 230, 234, 236, 238, 251–252  
Propagation Loss 2, 24, 30, 36, 39–43, 73, 112, 165, 173, 245, 259, 275–277, 315  
Proton Beam Writing 37–38  
  
Quality factor (Q-Factor) 36, 59, 74, 159  
Quantum Confined Stark Effect (QCSE) 135

- Quantum Confinement 10, 153–156  
 Quantum Efficiency 202–203
- Radiation loss 72–74, 229–230, 232, 236  
 Radio-over-Fibre (ROF) 304–307  
 Raman Amplifier 4, 162, 166, 170, 178, 182, 311  
 Raman Gain 124, 161, 164–166, 170–173, 182–184, 186, 307–308, 312  
 Raman Gain Coefficient 164, 307–308  
 Raman Laser 4–5, 124, 162, 166, 169, 175–178, 180–183, 185, 309  
 Raman Matrices 163  
 Receiver 4, 7, 118–119, 206, 270–271, 283, 285, 288–289, 300–302  
 Refractive index, 16–19, 27–28, 32, 34–35, 40–41  
 Response Time 203–204  
 Responsivity 203  
 RF Over Fibre 305–306  
 Rib Waveguides 23–28  
 Ring Resonator 259–263  
 Rise time 113, 124, 134  
 Roughness, sidewall 30, 39–40, 229, 237, 258, 262
- Scanning Electron Microscope (SEM) 61, 120  
 Schrödinger equation, 50  
 Self imaging 237, 239  
 Sensing 4, 36, 40, 180, 256, 259, 297, 314–320  
 Sensors, Biochemical 317–319  
 Sensors, Chemical 316–317  
 Sensors, Physical 314–316  
 Shockley –Read-Hall Recombination (SRH) 193–194  
 Silicon Lasers 147–186  
 Silicon nanocrystals 153, 157  
 Silicon on Insulator (SOI) 15, 23, 25, 32–33, 40, 64, 139, 163, 210, 240, 274–276, 288  
 Silicon on Sapphire (SOS) 9, 15, 64, 182  
 Silicon Rich Oxide (SRO) 153, 154f  
 Single mode condition (SMC) 23–26, 28f, 29, 29f, 31  
 Spintronics 11  
 Splitter 3–4, 49, 75, 82–84, 139, 237–239, 241–246, 250, 263, 302
- Stimulated Raman Scattering (SRS) 160–162, 164, 172, 186, 307  
 Stokes Waves 162, 170–171  
 Strained Silicon 135, 137, 279  
 Strip waveguides 9, 15, 23, 28–31, 36, 44, 122, 254, 261–262  
 Sub Bandgap detection 205–225  
 Switch 2–3, 7, 75, 96, 107–108, 111, 114, 116, 118, 121, 125, 147, 165, 178, 180, 237, 241, 244, 248, 250, 314
- Tapers 31–32, 173, 215, 282  
 TE-TM converters 3  
 Thermo-optic coefficient 105  
 Thermo-optic effect 105  
 Time Division Multiplexing (TDM) 5  
 Total Internal Reflection (TIR) 54, 60, 75  
 Transceiver 5, 7–8, 160, 206, 274, 285–289  
 Transient 113–114, 134, 201  
 Transition loss 229–232, 236  
 Transmission loss 73, 232, 240, 252–253, 259, 261–262  
 Two Photon Absorption (TPA) 164, 167, 172, 180–181, 186, 308–309
- Unidirectional Emitter 78–80
- Variable Optical Attenuator (VOA) 4, 221, 241  
 Vertical taper 31–32
- Wave Equation 16–19  
 Waveguide Bend 229–233  
 Waveguides 2, 4–5, 7, 9–12, 15–44, 58–60, 97, 107–108, 122, 229, 232–233, 241, 253–254, 256, 261–263  
 Wavelength conversion 161, 170, 175, 178, 302, 312–313  
 Wavelength Division Multiplexing (WDM) 5, 86, 258  
 Weighted splitting 242
- Y-junction 29, 242–245
- Zener Breakdown 199  
 Zero Birefringence condition 25–26