

Heterogeneous Optoelectronics Integration

Heterogeneous Optoelectronics Integration

Elias Towe

Editor



SPIE PRESS

A Publication of SPIE—The International Society for Optical Engineering
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Heterogeneous optoelectronics integration/ [edited by] Elias Towe.

p. cm.

Includes bibliographical references and index.

ISBN 0-8194-3571-6

1. Integration optics—Congresses. 2. Optoelectronic devices—Design and construction—Congresses. 3. Integrated circuits—Very large scale integration—Congresses. I. Towe, Elias. II. SPIE—the International Society for Optical Engineering. III. Series.

TA1660.H48 2000

621.36'93—dc21

00-030808

CIP

Published by

SPIE—The International Society for Optical Engineering

P.O. Box 10

Bellingham, Washington 98227-0010

Phone: 360/676-3290

Fax: 360/647-1445

Email: spie@spie.org

WWW: www.spie.org

Copyright © 2000 The Society of Photo-Optical Instrumentation Engineers

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means without written permission of the publisher.

Printed in the United States of America.

CONTENTS

PREFACE / ix

CHAPTER 1. MONOLITHIC AND HETEROGENEOUS OPTOELECTRONIC INTEGRATION: THE EPITAXY-ON-ELECTRONICS, SILICON-ON-GALLIUM-ARSENIDE, AND ALIGNED-PILLAR-BONDING TECHNIQUES

Clifton G. Fonstad, Jr. / 3

- 1.1 Introduction / 3
- 1.2 Hybrid, Monolithic, and Psuedomonolithic Integration / 4
- 1.3 Materials Issues in Integration / 7
- 1.4 Gallium Arsenide-on-Silicon / 10
- 1.5 Epitaxy-on-Electronics / 14
- 1.6 Silicon-on-Gallium Arsenide / 27
- 1.7 Aligned-Pillar Bonding / 35
- 1.8 Conclusion and Acknowledgements / 40
- References / 41

CHAPTER 2. VERTICAL-CAVITY SURFACE-EMITTING LASERS SPECIFICALLY DESIGNED FOR INTEGRATION WITH ELECTRONIC CIRCUITS

Leon M. F. Chirovsky, Ashok V. Krishnamoorthy, William S. Hobson, John Lopata, Keith W. Goossen, and L. Arthur D'Asaro / 49

- 2.1 Motivation for the Work / 49
- 2.2 Background Information / 52
- 2.3 Detailed Description of an I²-VCSEL, the First Implementation of the P²I²-VCSEL Design Strategy / 54
- 2.4 Data on the First, Top-Emitting I²-VCSELs / 57
- 2.5 Data on Bottom-Emitting and on Hybrid-Integrated I²-VCSELs / 59
- 2.6 Design Modifications for High-Speed P²I²-VCSELs / 66
- 2.7 Data on High-Speed P²I²-VCSELs and its Analysis / 67
- 2.8 Conclusion / 70
- Acknowledgements / 72
- References / 72

**CHAPTER 3. HETEROGENEOUS INTEGRATION OF VERTICAL-CAVITY
SURFACE-EMITTING LASER ARRAYS TO CMOS
INTEGRATED CIRCUITS**

Carl W. Wilmsen and Rui Pu / 75

- 3.1 Introduction / 75
- 3.2 General Discussion of Possible Methods of Bonding VCSELs to CMOS / 76
- 3.3 Review of Heterogeneous Integration of VCSELs and Electronic Circuits / 79
- 3.4 Overview and Details of Bonding Techniques / 80
- 3.5 Electrical and Optical Characteristics / 89
- 3.6 Thermal Characteristics / 95
- 3.7 Comparison of These Results with Other Methods and Appraisal of These Techniques / 104
- 3.8 Applications / 112
- 3.9 Summary / 115
- Acknowledgements / 116
- References / 117

**CHAPTER 4. HETEROGENEOUS AND MONOLITHIC INTEGRATION
FOR OPTICAL INTERCONNECTS**

Duane A. Loderback, Larry A. Coldren, Ola Sjölund, Eric R. Hegblom,
Eva M. Strzelecka, Shigeru Nakagawa, Syn-Yem Hu, Dan D. Lofgreen,
Ryan L. Naone, Jack Ko / 121

- 4.1 Introduction / 122
- 4.2 Scaling VCSEL Performance: Small, High-Performance VCSELs / 123
- 4.3 Monolithic Refractive Microlenses / 136
- 4.4 Monolithically Integrated VCSELs and Detectors with Microlenses / 141
- 4.5 WDM Pie-Shaped VCSEL Arrays / 168
- 4.6 Summary / 172
- References / 173

CHAPTER 5. LARGE-SCALE HETEROGENEOUS OPTOELECTRONIC INTEGRATION

John A. Trezza, Jeffry Powell, Richard J. Williams, Ronald J. Olson,
Mani Sundaram, Keith Kang / 181

- 5.1 Introduction to Very-Large-Scale Heterogeneous Integration / 181
- 5.2 Devices and Device-Integration Techniques / 182
- 5.3 Applications / 198
- 5.4 Summary / 221
- References / 222

CHAPTER 6. MOLDED INTEGRATION OF OPTICAL, OPTOELECTRONIC AND ELECTRONIC COMPONENTS

Volkan H. Ozguz and John C. Carson / 225

- 6.1 Introduction / 225
- 6.2 Description of the Developed Molding Technique / 226
- 6.3 Application to Optoelectronic Integration / 232
- 6.4 Future Development, System Applications and Conclusions / 245
- Acknowledgements / 247
- References / 247

CHAPTER 7. SCALABLE MICRO-OPTICAL MODULES FOR SHORT DISTANCE VLSI-PHOTONIC INTERCONNECTIONS

Hugo Thienpont, Valerie Baukens, Christof Debaes, Heidi Ottevaere, Pedro
Vynck, Patrik Tuteleers, Guy Verschaffelt, Bart Volckaerts, Alex Hermanne,
Mike Hanney, Marnik Brunfaut, Jan Van Campenhout, Irina Veretennicoff / 251

- 7.1 Introduction and Rationale / 251
- 7.2 The Concept / 255
- 7.3 Design of the Optical Bridge / 259
- 7.4 Fabrication with Deep Proton Lithography / 260
- 7.5 Proof-of-Principle Demonstration / 266
- 7.6 Mass-Fabrication and Tolerances / 268
- 7.7 Conclusion / 271
- Acknowledgements / 273
- References / 274

PREFACE

For some time now, the performance of electronic circuits based on silicon very large-scale-integration (VLSI) technology has doubled every eighteen months. Each reduction in the typical device feature size of most VLSI circuits has meant that more gates could be fabricated in a unit area of a chip. This, in turn, has implied an increase in the internal computational bandwidth of the chips. Connecting the VLSI chips to one another with wire interconnects is inefficient because the aggregate communication bandwidth of the wire links is only a small fraction of the internal computational bandwidth of the chips. To obtain the greatest benefit from the internal chip bandwidth, the chip-to-chip and board-to-board communication bandwidths should be commensurate with, or at least be a reasonable fraction of the internal computational bandwidth of the chips. It is now evident that wire interconnections are not likely to meet the chip-to-chip and board-to-board communication needs of future high-performance information systems. A promising alternative scheme that is being explored by a number of groups is the use of optics at the chip-to-chip and board-to-board levels of the interconnection hierarchy. At this level of the hierarchy, however, there are a number of problems that must be resolved before optics can penetrate down to the chip-to-chip level. One of the major problems preventing the use of optics at this hierarchy is the incompatibility of silicon VLSI electronics technology with photonics technology. Photonics technology is largely based on group III and V compound semiconductors. Monolithic integration of the two technologies for low-cost, manufacturable, and practical chips appears to be out of the question. Because of the clear and obvious advantages of using optics in long-haul communications, a number of efforts are underway to extend these benefits to the chip-to-chip and board-to-board levels. The general hierarchy of optical interconnections is shown in Fig. 1.

At the present state of development, the use of optics in VLSI silicon electronics at the chip-to-chip level is likely to be implemented only in extremely high-performance systems, where the electrical solution is inadequate. At the board-to-board level, there appears to be no insurmountable difficulty in using optics. In fact, there are already several manufacturers seriously considering incorporating optics in their products for board-to-board interconnections. In addition to generic point-to-point links, optics is likely to be used in system clock distribution to alleviate skew and jitter problems in high-frequency systems.

The chapters in this book describe several approaches to developing the basic technologies necessary for optical interconnections. The approaches being pursued are hybrid in nature; they are more generally known as heterogeneous opto-

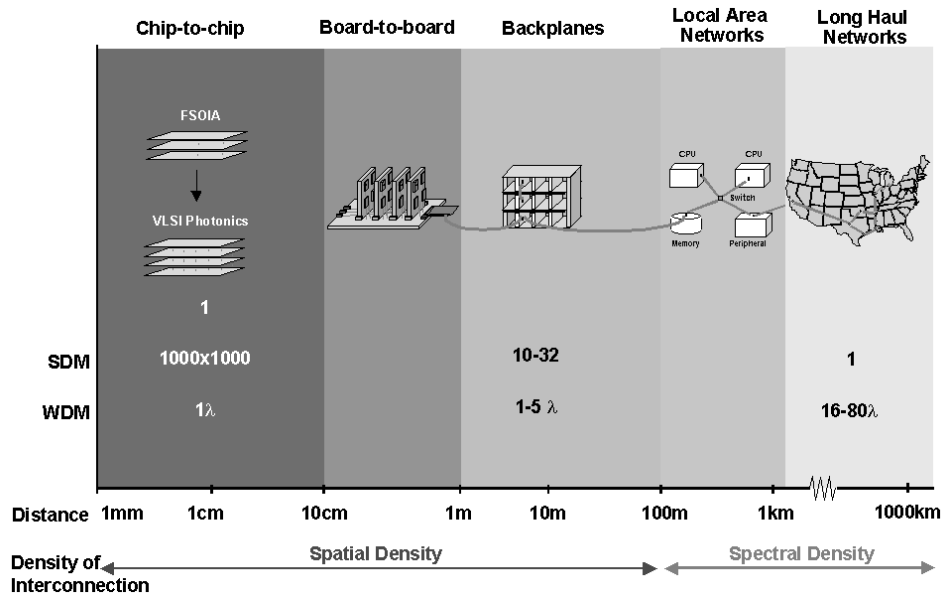


Figure 1. Hierarchy of optical interconnections.

electronic integration technologies. The premise of heterogeneous integration is that for the foreseeable future, silicon electronics will continue to be the dominant technology for computing, and photonics will dominate communications. Underlying this premise is a tacit acknowledgment of the economic reality that investments in billions of dollars have been made to build an extensive worldwide infrastructure for silicon-processing technology. Under these circumstances, it makes sense to explore techniques that harness the benefits of silicon electronics and III-V optoelectronics in certain high-performance applications.

The first chapter of the book, by Fonstad, describes three methods for integrating electronics with optoelectronics. The first of the methods discusses how optoelectronic devices, such as light-emitting diodes and photodetectors, can be integrated onto fairly complex electronic chips fabricated on GaAs substrates. The idea of the approach is to grow optoelectronic device structures in trenches etched in processed GaAs wafers with integrated electronic circuits on them. The second approach is to produce fused Si-on-GaAs wafers which can be used for the epitaxy of optoelectronic devices by opening windows to the GaAs substrates; these same wafers can then be used to produce CMOS-based chips. The third method is based on bonding of processed or partially processed optoelectronic devices onto selected sites on a processed silicon wafer.

The second chapter, by Chirovsky et al., describes a unique vertical-cavity surface-emitting laser (VCSEL) specifically designed for ease of integration with CMOS electronics. The p- and n-ohmic contacts to the device are intracavity con-

tacts. This is in contrast to conventional VCSEL devices, where the contacts are fabricated on the top and bottom distributed-Bragg-reflector (DBR) mirrors. The current injected by the contacts is guided through an aperture defined by a shallow ion-implant region in the device cavity. In addition, this generation of VCSELs has a dielectric top mirror deposited after growth. Before the dielectric mirror stack is deposited, an index guide layer for beam guidance is first deposited. The authors have flip-chip bonded a 16×16 array of the devices to CMOS driving electronics. The flip-chip bonding is facilitated by the intracavity contacts, which are essentially coplanar.

The third chapter of the book is by Wilmsen and Pu. This chapter discusses methods of bonding VCSELs to CMOS electronics. The question of whether the substrate should be removed from the VCSELs is addressed. Other issues discussed in this chapter include the common fabrication sequences used. The authors end the chapter by discussing a potential application of VCSELs in a data-communication switch.

The fourth chapter is by Louderback et al. The authors of this chapter describe the design, fabrication, and characterization of high-performance VCSELs. They discuss the trade-offs made and scaling rules used to achieve single-mode, low-threshold-current operation of these devices with high wall-plug efficiencies. For free-space optical interconnections, the authors describe a scheme for integrating microlenses onto the VCSEL. The device arrays with microlenses also include integral resonant-cavity detectors. Finally, the authors discuss VCSEL-based short-distance multimode fiber links.

In Chapter 5, Trezza et al. discuss several integration technologies. Of the technologies they describe, flip-chip bonding is arguably the most advanced for bonding extremely large chip sizes to other chips or to substrates. They use this technology to bond GaAs optoelectronic device arrays onto silicon CMOS electronics. As an example, they describe the fabrication of large infrared focal-plane arrays for imaging. These arrays are based on GaAs/(Al,Ga)As quantum-well infrared photodetectors flip-chip bonded to silicon-readout CMOS electronics.

Chapter 6, by Ozguz and Carson, discusses integration of chips onto functional three-dimensional cubes. Their integration approach embeds electronic and optoelectronic chips in a transparent molding compound, allowing the incorporation of optical input/output elements for interconnections. Using this technology, the authors have been able to integrate molded VCSELs and detectors with appropriate control electronics; they have been able to demonstrate operation at 1 Gb/s.

The seventh and final chapter of the book is by Thienpont et al. These authors discuss the fabrication of multichannel free-space interconnection modules for inter- and intra-MCM interconnection distances. A single channel in the interconnection scheme discussed in this chapter has a transmission bit rate of 622 Mb/s. With 16 channels, this scheme is capable of operating at aggregate interconnection

rates of up to 1 Tb/s cm². The basic channel consists of a transmitter and a receiver connected by an optical bridge guide, which links the transmitter to the receiver.

The chapters in this book describe the significant progress has been made in the development of optoelectronic devices, micro-optics, and integration technologies for optical interconnection modules. One of the key factors that is accelerating experimentation with optical interconnection technologies within information processing systems is the performance-limiting effects of conventional wire interconnects. The recognition that optics offers a potentially viable interconnection solution, and the rapid advances being made in the integration of optoelectronic devices with silicon CMOS technology, are creating the right mix of circumstances for the emergence of optics as an important enabling technology.

The preparation of this book without the combined efforts of many individuals would not have been possible. I would like to thank all the authors and co-authors for contributing their work to this volume. And finally, I want to acknowledge the support and assistance of Ms. Claire T. Heller during the preparation of this book.

Elias Towe

Arlington, Virginia

June 2000

Monolithic and Heterogeneous Optoelectronic Integration: The Epitaxy-on-Electronics, Silicon-on-Gallium-Arsenide, and Aligned-Pillar-Bonding Techniques

Clifton G. Fonstad, Jr.
Massachusetts Institute of Technology

1.1 INTRODUCTION

Before optoelectronic integrated circuits and systems can be produced in high volumes and at low cost it will be necessary to develop for them monolithic integration techniques which can be executed on full wafers and in multiwafer lots. Monolithic processes are also essential to achieving the highest speed and performance from optoelectronic integrated circuits (OEICs) and to insuring robustness and high mechanical reliability. Hybrid techniques can satisfy many near-term needs for OEICs, but they involve a great deal of piecework and assembly, thereby increasing costs, introducing performance-limiting parasitics, and decreasing ruggedness. It is monolithic, full-wafer, batch-level processing of OEICs that must be the ultimate goal of any program promoting the commercial use of optoelectronics to solve problems in computation, signal processing, data analysis, and sensing. This is the hypothesis of our program and the focus of this review.

The review begins in Section 1.2 with a discussion expanding upon the statements made in the preceding paragraph and making the arguments in favor of monolithic and pseudomonolithic integration. After the importance of these technologies to the future commercial success of OEICs is established, Section 1.3 continues with a review of the properties of silicon and of the common III-V materials used in optoelectronic devices, and with a discussion of the issues which make the integration of these materials difficult.

The history of gallium arsenide epitaxy on silicon (GaAs-on-Si) is then briefly reviewed in Section 1.4. This section concludes with a discussion of promising etch-release techniques that were pursued briefly by several groups just as support for GaAs-on-Si research was waning and the approach was abandoned.

In Section 1.5 the epitaxy-on-electronics (EoE) technique is described. EoE has been used to produce the most complex monolithic OEICs yet fabricated, and is compatible with full-wafer, batch-level processing. EoE builds on gallium arsenide integrated circuits, rather than silicon ICs; it is thus a homogeneous integration technique and it leaves unsolved the problem of using Si-CMOS in OEICs.

To extend EoE to heterogeneous integration with silicon electronics, the silicon-on-gallium arsenide (SonG) process is being developed. SonG, which is described and reviewed in Section 1.6, provides substrates for EoE integration on silicon CMOS, i.e., for true monolithic heterogeneous optoelectronic integration. SonG builds on recent developments in silicon on insulator (SOI) technology, wafer bonding and micro-electro-mechanical systems (MEMS) research, and the epitaxy-on-electronics (EoE) program to integrate state-of-the-art CMOS, GaAs direct coupled FET logic (DCFL), and optoelectronic devices monolithically on the same substrate. It is an exciting recent development that promises to have an impact far beyond optoelectronic integration.

A final integration technology, aligned pillar bonding (APB), will be described in Section 1.7. APB is a complementary technique which overcomes many of the constraints of EoE processing and provides an alternate way to integrate optoelectronic device heterostructures on SonG substrates and GaAs IC wafers. APB can also be used to integrate optoelectronic devices with silicon-on-sapphire (SOS) wafers. It is a pseudomonolithic technique which combines all of the best features of monolithic integration with the flexibility and universality of hybrid integration.

The review concludes in Section 1.8 with a brief summary and look to the future and with an acknowledgement of the many persons who have contributed to this work.

1.2 HYBRID, MONOLITHIC, AND PSEUDOMONOLITHIC INTEGRATION

1.2.1 THE EVOLUTION OF ELECTRONIC INTEGRATION

The push to miniaturize electronic components and to package them in ever smaller, more rugged and economical ways began when vacuum tubes still ruled the electronics field, and well before the successful demonstration of silicon bipolar transistors and MOSFETs [1]. With the advent of the transistor, however, hybrid integration truly began to develop in the 1950s and take form as we know it today. By the early 1960s solder-ball flip-chip mounting of devices on ceramic substrates was being used commercially [2], a good example being the IBM 360 of 1964, the first true computer “family” in which this technique was used to mount the individual transistors and small scale integrated (SSI) circuits. The author’s first summer job in engineering was at an small IBM laboratory in Rochester, MN, during the summer of 1964 when the 360 family of computers

was “rolled out” to much fanfare, and the solder-ball packaging technique was an important feature of the family. This technique has been improved and refined over the years and is now known as the C4 Process.

The limitations of hybrid integration were recognized even as it was being developed, however, and this led to the development of monolithic integration techniques in which most, and ideally all, of the devices in a circuit could be formed on a single substrate, and in a manner that is compatible with simultaneously processing large batches of full wafers, each containing numerous individual circuit chips. The economic advantages of amortizing the cost of a complex process over the many chips fabricated simultaneously in this way resulted in a dramatic decrease in the cost of consumer and industrial electronics and a corresponding increase in the utilization of integrated electronics in products, which in turn drove further innovation and growth in the industry. The resulting Moore’s Law evolution of integrated circuits is legend [3].

1.2.2 OPTOELECTRONIC DEVICES AND THEIR INTEGRATION

Optoelectronic devices have a long history, also. One can even argue convincingly that some of the first semiconductor devices were optoelectronic photo-detectors, and photoconductive photodetectors were widely used in photography and other sensor applications well before the invention of the transistor in 1948 [4]. The potential value of using optoelectronic devices to augment electronic devices was also recognized early and proposals for optoelectronic integrated circuits can be found in the mid- and late 1970s [5, 6].

The commercial impact of optoelectronic integrated circuits has been very modest, however, and OEICs have yet to “take off” economically. Even discounting the fact that many of the initial visions of how pervasive the impact OEICs would be were unrealistic, and that early attention focused unwisely at replacing, rather than complementing, silicon with gallium arsenide, and then electronics with optics, one still finds that the impact of OEICs has been far less than it might have been. One can argue that one of the reasons for this was the lack of an attractive vertically-emitting laser, but this problem was solved by the invention of the vertical cavity surface emitting laser (VCSEL) around 1990. A more important reason, and perhaps the primary reason for the modest impact of OEICs, has been the lack of any sophisticated optoelectronic integration technology, particularly technology that could build on the enormous investments already made in developing and processing commercial electronic integrated circuit processes, be they in silicon or GaAs. The situation is slowly changing with the development of solder bump bonding techniques for integrating VCSEL and detector arrays with silicon CMOS chips [7], but this process is only a relatively modest advancement over the flip-chip process of the early 1960s and it is unclear how far it will take OEICs commercially. What are needed are additional integration technologies that bring the

economic and performance advantages of monolithic integration to the production of OEICs, and that build upon existing commercial integrated circuit foundations.

The difficulty of making OEICs, even in the laboratory, much less in any readily accessible foundry or similar commercial setting, has also limited their impact by denying many researchers access to them, researchers who might in turn develop new applications for OEICs. There is a feedback mechanism at work in the development of many fields wherein access to technology for experimentation leads to new inventions and applications, which in turn drive further development of the technology and further growth of the industry, and so on. OEICs have not gotten rolling, as it were, in this respect either.

1.2.3 LESSONS FROM MEMS

One way of seeing the significance of building upon existing technologies, especially the silicon IC processing industry, is to look at the example of micro-electro-mechanical systems, MEMS. This field appeared much more recently than the field of optoelectronics, but has advanced considerably more rapidly. The presence of several sophisticated MEMS ICs in the airbag deployment systems of all new cars should be cause for thought for OEIC champions.

The first lesson from MEMS for OEIC researchers concerns the advantages of building upon existing commercial processes and capabilities. MEMS builds directly on silicon and silicon processes, taking them as a starting point and modifying them as little as possible. The compatibility of the materials and fabrication techniques used in MEMS structures with silicon integrated circuit materials and processes has allowed the rapid development, and commercial adoption, of fabrication techniques which bring to MEMS integrated circuits all of the performance and economical advantages of monolithic integration and full-wafer, batch processing that are enjoyed by the silicon IC industry.

In contrast, optoelectronic devices are difficult, and often impossible, to produce in silicon and are best made in gallium arsenide and indium phosphide-based III-V materials systems. These materials, as we shall see in Section 1.3, are in important ways incompatible with silicon, and exploiting and building upon the IC industry's installed technology base is much more difficult in the case of OEICs than it is for MEMS. Historically what has been done instead has been to attempt to build a complete OEIC technology from the substrate up. Many versions of III-V based OEIC process sequences have been developed over the years [8, 9], but simultaneously developing a high-density, high-performance electronics technology, and a technology for high-performance optoelectronic devices such as laser diodes is an enormous undertaking. As a consequence, the levels of complexity, density, and performance that have been achieved in the electronics have been relatively modest: medium-scale integration (MSI), at best, whereas what most applications really need is state-of-the-art very-large-scale integration (VLSI). The

lesson is that even though building OEICs on a commercial VLSI foundation is difficult, it is a very important thing to do. The gallium arsenide-on-silicon efforts of the mid- to late-1980s which will be described in Section 1.4 were an attempt to apply this lesson, and the development of the epitaxy-on-electronics and silicon-on-gallium arsenide integration technologies to be described in Sections 1.5 and 1.6, respectively, is a direct response to it also.

A second important lesson of MEMS for OEIC researchers is that traditional pure monolithic integration is not essential, and that pseudomonolithic integration is acceptable. By “pseudomonolithic” we mean integration processes which involve monolithic processing of separate wafers and then the combining of two or more such wafers using a bonding technique, followed in most cases by additional monolithic processing and testing of this new “composite wafer” prior to dividing it into individual integrated circuit chips. In using the term pseudomonolithic we seek to avoid a semantic battle over whether anything involving bonding can be considered monolithic (i.e., whether bonding is an exclusively hybrid assembly technique), and to focus on the key issue: that the primary objective is to achieve the packing density, ruggedness, and processing economics of traditional monolithic processing by doing as much processing as possible at the wafer level in batches, and to not handle individual circuits until the final stage of packaging. Widely used in MEMS IC processing, pseudomonolithic integration can be a particularly powerful technique for optoelectronic integration, also, as we shall see in when we discuss aligned pillar bonding in Section 1.7. In this regard it is important to note that precision wafer alignment and bonding techniques and equipment have been developed in the MEMS field which allow pseudomonolithic integration to be done with much the same alignment tolerances and device densities as found in traditional monolithic integrated circuits. (The densities can actually be higher if one takes into account the three-dimensional capabilities of pseudomonolithic integration.)

1.3 MATERIALS ISSUES IN INTEGRATION

1.3.1 NARROWING THE FOCUS

There are many materials properties which must be taken into account in heterogeneous integration of semiconductors, but there are two properties which are far more important than the rest: the crystal lattice structure and the coefficient of thermal expansion. Of these, the second is the most problematic because its impact is felt whether one is combining the materials by epitaxy (for which the crystal lattice structure is also very important) or by bonding (in which case crystal structure is unimportant).

1.3.2 THERMAL EXPANSION

The coefficients of thermal expansion (CTEs) of several important materials of interest in optoelectronic integrated circuits are summarized in Table 1.1. Concentrating for the moment on silicon and gallium arsenide, one notices that there is a factor of almost 3 difference in the coefficients of thermal expansion. This has important consequences for both epitaxy and bonding, as we have said.

In the case of epitaxy, imagine growing a layer of silicon epitaxially on a gallium arsenide substrate, or vice versa. Typical epitaxy temperatures are 600°C and 750°C for GaAs by molecular beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD), respectively, and 1000°C, or more, for Si by CVD. Upon cooling to room temperature, the epitaxial layer will be under severe stress (compressive in the case of Si on GaAs, and tensile in the case of GaAs on Si). A layer that is much more than a hundred nanometers thick will crack and there will be many defects generated elsewhere within it. It is impossible to quantify the effects more without picking a specific layer thickness and growth temperature, but the picture is clear: Unless the epitaxy can be done at very near room temperature (which it generally cannot), the epitaxial layer will be under severe stress at room temperature, and very likely will contain a large number of stress-induced defects unless it is very thin. This was what limited gallium arsenide-on-silicon research, as we shall see in Section 1.4, because the gallium arsenide-based heterostructures which one had to grow on silicon for the desired optoelectronic devices had to be many hundreds of nanometers thick, often several microns thick. (Even if the epitaxy could somehow be done near room temperature, the problem would not be totally solved unless all of the subsequent device and circuit processing could also be done near room temperature.)

For bonding the CTE problem can surface in several ways. One is in much the same way as it does in the case of epitaxy, namely in the form of destructively large stresses in wafers bonded together and heated or cooled several hundred

Material	Coefficient of Thermal Expansion
Silicon, Si	2.6×10^{-6}
Germanium	5.9×10^{-6}
Gallium arsenide, GaAs	6.86×10^{-6}
Indium phosphide, InP	4.75×10^{-6}
Sapphire, Al_2O_3	5.2×10^{-6}
Silicon dioxide, SiO_2	5×10^{-7}

Table 1.1. The coefficients of thermal expansion near and above room temperature for several materials of importance in optoelectronic integrated circuits [10, 11].

degrees from the bonding temperature. For example, if a pair of typical thickness, four-inch-diameter silicon and gallium arsenide wafers is hydrophilically bonded, or even simply epoxied, together at room temperature and the temperature of the bonded pair is then increased, there will be significant stressing and bowing of the pair until one or both of the wafers breaks at approximately 325°C [12].

Another bonding context in which the CTE difference surfaces is during aligned bonding. In this case, if a patterned GaAs wafer is to be aligned and bonded to a correspondingly patterned Si wafer, for example, the CTE difference will have a significant deleterious impact on the relative alignment of the patterns if the bonding is done at anything other than room temperature. Consider, for illustrative purposes, aligned bonding of two six-inch wafers, the smallest diameter one is likely to encounter in industry. If the patterns on the two wafers are aligned at room temperature and their temperature is raised by 100°C, the diameter of the GaAs wafer will be found to have increased by 65 microns more than that of the Si wafer, and the patterns near the edges of the wafers will be misaligned over 30 μm (assuming alignment is maintained in the center). Most bonding requires temperatures of several hundred degrees, so there is clearly a problem. The only solution would be room temperature bonding.

If in the above example the bonding could be done at room temperature, one still faces the problem of also doing all subsequent processing at room temperature. A way around this problem is to thin one of the bonded pair of wafers (usually the one with the optoelectronic devices on it) and to then etch it into many smaller pieces, each one of which is bonded to the other substrate but is otherwise physically isolated from its mates. This does not eliminate the stress experienced by each individual piece, but it does reduce the overall stress on the remaining whole substrate and it does alleviate much of the problem. Nonetheless, it is still necessary to limit the temperature excursions or risk damaging the bonded pieces. Even if the stress in each individual piece is not enough to cause catastrophic damage, it may be sufficient to significantly reduce the life expectancy of the bonded device. It is well known, for example, that in-plane lasers will degrade rapidly, even under nominal operating conditions, if they are not mounted in packages which match their thermal expansion coefficient and with a solder which itself does not introduce stress [13, 14]. This is particularly true of GaAs-based in-plane devices mounted junction-side down. The situation appears to be less severe for GaAs-based vertical cavity surface emitting lasers (VCSELs) because the distributed Bragg reflector mirror stacks tend to be several microns thick, which places the junction further away from the bond, but it is still a cause for concern because reliability is always an important issue and anything which impacts it negatively should be avoided. InP-based devices, on the other hand, tend to show much less stress-related degradation than do GaAs-based devices [15].

The only way to totally avoid CTE mismatch and the problems such as stress-induced degradation that it creates, is to grow, bond, or mount the stress-sensitive devices (typically the optoelectronic devices and in particular the laser diodes) on substrates to which they are CTE-matched. Looking at Table 1.1, we see that for GaAs-based devices that means using gallium arsenide, germanium, and/or sapphire substrates. For InP-based devices the only choice is indium phosphide itself (fortunately, as mentioned in the preceding paragraph, InP-based devices tend to be much less sensitive to stress). Using substrates to which the optoelectronic devices are CTE-matched is a major theme in the integration technologies discussed Sections 1.5, 1.6, and 1.7 of this review.

1.4 GALLIUM ARSENIDE-ON-SILICON

1.4.1 THE INITIAL SUCCESS AND PROMISE

For many years strict lattice-matching was considered to be a prerequisite for successful epitaxy, and given that the epitaxy technique of record for much of that time was liquid phase epitaxy (LPE), this constraint made a good deal of sense. LPE, being a nearly-equilibrium technique, produces very poor results unless the degree of lattice matching is high.

With the emergence and increased availability of molecular beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD), both of which are more nonequilibrium growth techniques, researchers began to explore breaking the lattice-matched barrier. One direction this work took was to study thin, highly strained layers, often termed “pseudomorphic strained layers,” and the result was such important devices as the pseudomorphic high electron mobility transistor, or PHEMT [16], and the strained quantum well laser [17].

Another direction this work took was to study the growth of thicker mismatched layers which were thick enough that they “relaxed” to their intrinsic lattice constant, and one of the results of this work was the effort on GaAs-on-Si. Much to the surprise of many, it was discovered that quite good quality gallium arsenide could be grown directly on silicon substrates [18] in spite of the large lattice constant difference (5.43 Å for Si versus 5.65 Å for GaAs [10]). The quality of the GaAs was much better than most people had dreamt would be possible and when it was further shown that respectable LEDs and lasers could be produced on silicon [19], the message seemed clear: it would only be a matter of time before OEICs with laser diodes grown on Si ICs were pervasive. It was quite natural, then, that the first attempts to monolithically integrate III-V optoelectronic devices with silicon integrated circuits focused on growing gallium arsenide and gallium aluminum arsenide heterostructures epitaxially on silicon substrates.

1.4.2 REALITY AND FRUSTRATION

Unfortunately, the initial enthusiasm for GaAs-on-Si soon turned to frustration as it became clear that although the quality of the GaAs grown on Si was much better than anyone had thought possible, it was still far away from the level that is needed to produce long-life, reliable laser diodes. That is not to say that there were not many significant advances made and much excellent science done, but rather that there are intrinsic problems in this materials combination which went largely unaddressed and which meant that all of the excellent results achieved were necessary, but not sufficient, to achieve the desired goal. Several excellent reviews of the many successes of researchers working on GaAs-on-Si can be found in the literature [20]; in the end, however, the fact that GaAs-on-Si laser diodes continued to fail after only a few hours of operation at best, in spite of the high level of resources and research funding devoted to the effort, led to a general disillusionment with GaAs-on-Si, and an abandonment of this avenue of optoelectronic integration.

1.4.3 DEALING WITH THERMAL EXPANSION IN GaAs-ON-Si

The problem with GaAs-on-Si was not the lattice mismatch between GaAs and Si, but rather the coefficient of thermal expansion mismatch. Techniques were developed to grow very-high-quality GaAs on silicon, or at least the quality was very high as the layers were being grown and before they were cooled to room temperature [21]. The problem was that because of the CTE mismatch, first, many defects were generated upon cooling to room temperature from the growth temperature, and, second, the gallium arsenide layers were under severe tensile stress even after some of the strain was released by forming defects. Both of these factors contributed to unacceptably short device lifetimes.

Even as interest in GaAs-on-Si was waning on the part of the funding agencies, several research groups, one each in Japan [22], Europe [23], and the United States [24], began taking notice of the CTE problem and began developing techniques to overcome it. In each case the approach taken was to release the epitaxial gallium arsenide from its rigid attachment to the silicon substrate, leaving it attached only by a very pliable bond or retainer. The epitaxial structure would then tend to relax because it was no longer tightly constrained by the silicon substrate, and it could be annealed to remove most, if not all, of the defects created in it during the cooldown process. In one instance, the released epilayer was used as a substrate for more epitaxial grown [23].

An example of the type of process pursued in our group is shown in Fig. 1.1. An AlAs layer was included in the structure just past the initial, high-defect buffer layers, and before the device layers. After the originally contiguous epitaxial growth was etched into small, isolated device mesas, a photoresist retaining structure was patterned between the mesa and the newly exposed substrate areas. The AlAs layer was then etched away using hydrofluoric acid, thereby releasing the

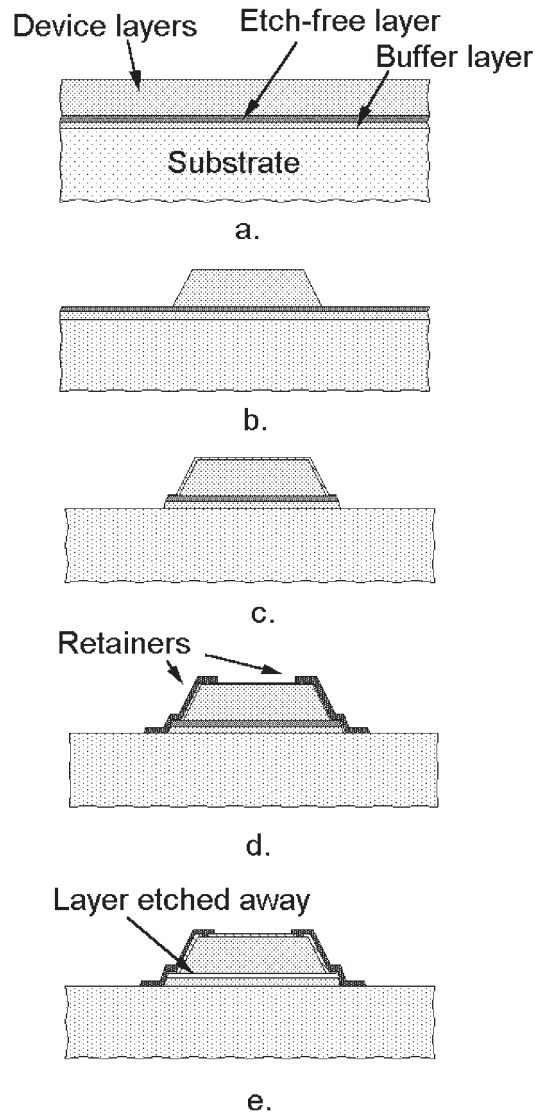


Figure 1.1. A series of sketches illustrating the use of an AlAs etch-free layer to separate a GaAlAs heterostructure from its silicon substrate to relieve strain induced by the differences in thermal expansion coefficients between Si and GaAs: (a) The original epitaxial heterostructure showing the nucleation layer, the sacrificial etch-free layer, and the device heterostructure. (b) The epitaxial heterostructure etched into device-size mesas going down to the AlAs layer. (c) With a coating applied to protect the device layers and second etch performed to expose the etch-free layer. (d) Application of the metal leads and/or a photoresist stay patterned between each mesa and the Si substrate. (e) After removal of the etch-free layer. At this point the heterostructure may be allowed to settle back onto the surface, or it can be left suspended above the surface. In either case, once freed from the substrate the device heterostructure is weakly enough held that it can be annealed to reduce the TCE-induced defects and stresses, yet it is still held in position sufficiently well that further device processing can be done on it.

GaAs-based heterostructure from the silicon substrate. The photoresist retainer was often found to be unnecessary because once the AlAs was etched away the mesa “settled” down on the substrate and was held in place by forces that were strong enough to keep it in place during subsequent processing, yet weak enough so that there was no CTE-induced strain in the mesa [24, 25].

The level of strain in the freed mesa could be monitored by studying the photoluminescence spectrum from the structure, as seen in Fig. 1.2. The four traces in this figure have been shifted vertically so they do not overlap, and the uppermost trace shows the spectrum from the as-grown GaAs-on-Si layer. Comparing this spectrum with the lowermost trace, which is the spectrum of a layer grown simultaneously on a GaAs substrate, we see that there is considerable tensile strain in the layer and the peak is shifted significantly to higher energies. Immediately after the AlAs layer is etched away, the strain in the layer is found to be largely removed since the main peak shifts to nearly line up with that in the layer grown on GaAs. There is, however, a broad deep peak indicative of residual damage. A subsequent 1 sec rapid thermal anneal of the “freed” (yet weakly held) layer to 600°C results in the disappearance of the deep peak and yields a spectrum which is identical to that of the layer grown on GaAs, indicating that all of the strain has been removed and many of the defects have been annealed out. In contrast, if an identical GaAs-on-Si layer is annealed without etching away the AlAs, one finds that the amount of stress actually increases as the annealing temperature is raised more and more above the growth temperature because more and more strain is being “set” into the layer [24, 25].

The same AlAs release layer technique was used to fabricate strain-free quantum well laser diodes on silicon substrates [26]. Since GaAs-on-Si lasers are normally under tensile strain, they lase in a TM mode, in spite of the fact that the cavity favors the TE mode, because the strain-shifted bands make the TM modes much higher gain. Once the strain is released by etching away the AlAs layer and annealing the released device heterostructure, however, the laser lases in the TE mode, as it should if it is relaxed, confirming again that the epilayers are no longer constrained by the Si substrate and are no longer subject to CTE-induced stresses.

1.4.4 GaAs-ON-Si REBIRTH

The successful demonstration that an AlAs etch release layer can free a GaAs-on-Si heterostructure from its CTE-mismatching Si substrate, while maintaining the monolithicity of the structure, offers considerable promise for the successful monolithic integration of reliable laser diodes on Si integrated circuits. The combination of this technique with modern techniques for gettering and pinning defects may well solve the problems that have plagued GaAs-on-Si integration. As noted earlier, however, by the time these results were being demonstrated, the fervor over GaAs-on-Si had been replaced by disillusionment and the attention of

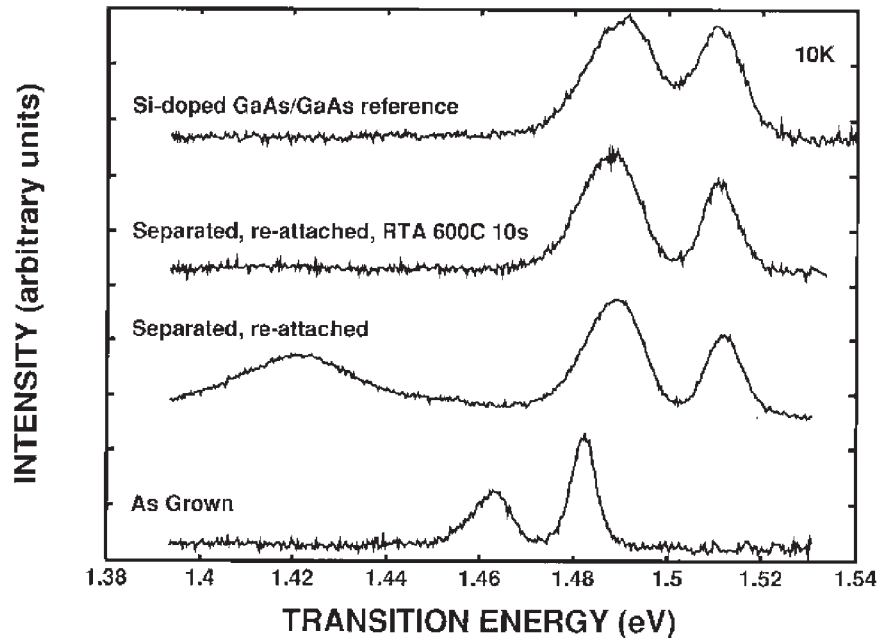


Figure 1.2. Photoluminescence spectra taken at 8 K from a GaAs-on-Si epilayer grown with an etch-free layer at several stages in its processing: Lower trace—As-grown. Second trace from bottom—After removal of the etch-free layer. Third trace from bottom—After a 10 s rapid thermal anneal to 600°C. After this anneal the spectrum is seen to compare favorably with that of a control structure grown simultaneously on a GaAs substrate, shown in the upper trace.

researchers and funding providers had turned to other topics of equivalent importance, such as vertical emitting devices. Now, a decade and more later, with the successful development of VCSELs and in light of the continuing need for monolithic Si-based OEICs, it will perhaps soon be time to revisit etch release techniques and to reassess GaAs-on-Si. At the same time, however, GaAs-on-Si is no longer the only option for VLSI-level OEICs and there are several viable alternative technologies available; these are the subjects of the next three sections.

1.5 EPITAXY-ON-ELECTRONICS

1.5.1 THE BASIC CONCEPT

In the latter half of the 1980s, when monolithic optoelectronic integration efforts were focused on growing gallium arsenide on silicon, the gallium arsenide integrated circuit industry was quickly advancing and growing in sophistication; by 1990 it was capable of producing very large scale integrated circuits that competed very favorably in several areas with silicon ICs. It also became possible for

researchers to get access to GaAs foundry services through the MOSIS service [27]; in particular MOSIS offered Vitesse Semiconductor Corporation's [28] digital H-GaAs MESFET process. These events were noted by our group in our search for an optoelectronics-friendly VLSI alternative to silicon, and gallium arsenide VLSI presented a possible solution [29, 30]. Similar observations were made by a group at CalTech, which soon began a collaboration with our group [31], and by a group at AT&T Bell Laboratories [32] (now Lucent), which did some important early work, but chose not to continue in this direction.

The idea, just as it had been with GaAs-on-Si, is to grow optoelectronic device heterostructures epitaxially on processed integrated circuit wafers on which the dielectric layers insulating the levels of interconnect metal and passivating the devices have been removed from the regions in which devices are to be integrated, so that the underlying substrate is exposed. Device heterostructures can then be grown in these dielectric growth windows (DGWs) and subsequently processed into devices. In the process, the devices are connected electrically to the pre-existing electronic circuitry. This process, which has been termed epitaxy-on-electronics (EoE), is illustrated in Fig. 1.3; it has been described thoroughly in the literature [33, 34] and in several graduate theses [30, 35, 36, 37]. It will only be reviewed in more general terms here.

Among the features of note in the EoE process is, first, that by growing in a window cut through the several-micron-thick dielectric layer stack, and adjusting the thickness of the epitaxial structure to match the depth of the opening, a very planar structure can be maintained, after any polycrystalline deposits on the dielectric surface outside the windows are removed. This planarity greatly facilitates the processing of the optoelectronic devices and their final electrical interconnection to the preprocessed electronic circuitry on the chips. Second, by growing after the back-end processing of the IC wafer has been completed, the optoelectronic devices are not overcoated with the thick dielectric stack covering most of the wafer surface. Such a coating would subject the underlying optoelectronic devices to a great deal of stress and would undoubtedly reduce their life expectancy significantly. Third, note that the contact to the substrate side of the device can be made through an ion-implanted region in the bottom of the growth window. This doped pattern is produced when the source and drain regions are implanted, and contact is made to it in the same manner that the source and drain are contacted. The final point to be made about the process is that just as it is a significant advantage to be building upon a commercial integrated circuits technology and not having to develop any piece of that technology, it is advantageous if existing, proven optoelectronic device structures can be used as much as possible, adapting them to the EoE environment with as few changes from the standard process sequence as possible. The more that can be imported from and built upon the work of other researchers, the more resources and energy are available to solve problems unique to the EoE environment.

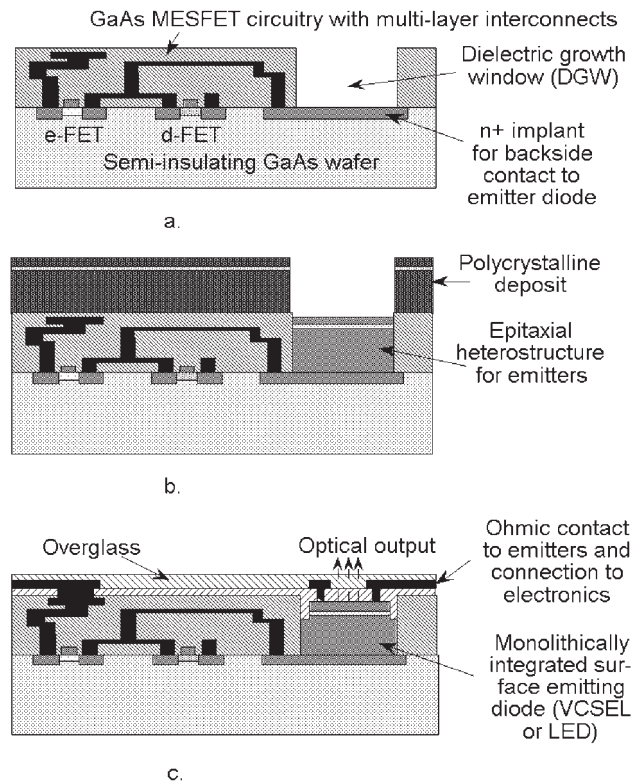


Figure 1.3. Key points in the EoE process: (a) the processed GaAs IC wafer as received from the manufacturer with growth wells and completed electronics; (b) after epitaxial layers growth and prior to removal of the polycrystalline deposits; (c) optoelectronic device processing and monolithic integration with GaAs e/d-mode electronics completed.

1.5.2 GALLIUM ARSENIDE ICs AND VLSI

MOSFETs like those used in silicon ICs are not available in gallium arsenide and thus GaAs integrated circuits (ICs) rely instead on metal semiconductor field effect transistors (MESFETs), heterojunction insulator gate FETs (HIGFETs), high electron mobility transistors (HEMTs), or heterojunction bipolar transistors (HBTs). A variety of processes are used commercially to fabricate GaAs integrated circuits, but not all GaAs ICs are suitable for EoE integration. The most fundamental problem with the circuits which are not suitable is that they contain gold-based ohmic contacts and interconnect lines, and/or they do not have refractory metal gates, and as a consequence they cannot withstand the high-temperature epitaxy step in the EoE process. To be EoE-compatible the circuits must be fabricated with refractory-metal gates and ohmic contacts, and must use a gold-free interconnect metallurgy, as will be discussed in the next subsection.

Many of the factors which lead to degradation of the EoE-incompatible circuits at high temperature may also limit device reliability, especially under ex-

treme operating conditions, so it is not entirely a coincidence that some GaAs processes have evolved away from gold-based metallurgies and nonrefractory metal gates and contacts, and toward EoE-compatible processes. Fabricating self-aligned structures typically involves high-temperature processing with various metals in place also, and this is another factor contributing to the emergence of EoE-compatible processes. There are now a number of manufacturers whose processes are EoE compatible (i.e. able to withstand additional high-temperature processing).

The circuits which have been used in the development of the EoE process are fabricated by Vitesse Semiconductor [28] because they are EoE compatible and because they were available through the MOSIS service [27], which used Vitesse as its GaAs IC foundry. The logic family used in these circuits is direct-coupled FET logic (DCFL), which is similar to silicon NMOS in that the switch device in a basic inverter is an enhancement mode transistor and the pull-up device is a depletion mode transistor.

An important difference between NMOS and DCFL is that the MESFET gates conduct when forward biased more than about 0.6 volts. This limits the logic high-output level in DCFL to this value, rather than the supply voltage as in NMOS, but this is in general more of an advantage than a problem. That MESFET gates conduct when forward biased has the desirable effect of reducing the noise on the power supply buss due to gate switching. This makes DCFL logic in general quieter in this regard than NMOS and CMOS. An important negative consequence of the Schottky-barrier gate of MESFETs is that charge can not be stored on a device gate as it can be with a MOSFET. This means that dynamic memory is not practical and that MESFET memory employs flip-flop-based static memory cells, which require more transistors, occupy more die area, and consume more power than single-transistor dynamic memory cells. This in turn means that large caches and RAM arrays are not commonly found in GaAs digital logic designs.

Complementary logic, such as a MESFET analog to CMOS, is also generally not available in GaAs. There is, however, one complementary family, known as CFET, which is based on heterojunction insulating gate field effect transistors, HIGFETs [38]. HIGFETs are made using an undoped, semi-insulating AlGaAs epilayer as the gate “insulator” in enhancement mode FETs. This structure is more complicated to fabricate than a normal MESFET because it involves epitaxial layers. It also suffers from the fact that the hole mobility in GaAs (and in III-Vs generally) is markedly less than the electron mobility. Nonetheless, CFET gates, like CMOS gates, do not dissipate standby power, and retain the speed performance advantages of GaAs. CFET ICs have been fabricated on a prototype production line using refractory metal gates and a gold-free back end process [39] so they may become a viable foundation for EoE integration; they are not, however, currently available commercially.

Returning to the specific MESFET DCFL IC technologies used in the present work, the most recent generation employed was H-GaAs IV. H-GaAs IV provides four levels of metal interconnect, ohmic contact to the lightly doped p-type layer in which the n-channel MESFETs are fabricated, and a deep, heavy n-type isolation implant with which the p-layer can be segmented into separate islands or wells. The latter two features (the implant to partition the p-layer, and the ohmic contacts to the separate p-type regions) were not available in the H-GaAs III process which was used in many of the EoE OEIC designs, and this limited the performance of those circuits, as will be pointed out later.

1.5.3 THE THERMAL STABILITY OF ELECTRONICS

Epitaxy is a high-temperature process so the impact of temperature on the integrated circuits is an important factor in defining the epitaxy-on-electronics process. Studies of the impact of five-hour periods at temperatures of 400°C and above on the MESFET characteristics and a variety of test structures showed that the resistance of the interconnect metal lines on the chips increased at temperatures much above 475°C [40]. At still higher temperatures the ohmic contacts and gate barriers change, but the weak link is clearly the interconnect metallurgy. The changes what occur are summarized in Fig. 1.4.

The changes in the interconnect lines were traced to a reaction between the aluminum which forms the bulk of the interconnect pattern, and tungsten nitride films on the upper and lower surfaces of the aluminum, which promote adhesion and facilitate photolithography. The reaction between Al and WN_x is a temperature-activated process with an activation energy for the WN_x alloy used in the Vitesse process of 3.5 to 4 eV [40]. Consequently there is no fixed upper temperature bound, but rather a boundary in the time-temperature space which should not be crossed. To ensure that the changes during a typical six- to seven-hour epitaxy sequence will have negligible impact on the circuits, it was determined that the temperature should be limited to 475°C.

By adjusting the nitride fraction of the WN_x layers it is possible to keep the resistance of the lines stable to higher temperatures, but not enough to make a significant difference. Furthermore, the WN_x composition had earlier been optimized with respect to the overall IC process, so that changing the interconnect metals solely for the convenience of the EoE process was not practical. Nor was it critical because the device epitaxy can be done within the permissible time-temperature window, as will be discussed in the next subsection.

1.5.4 THE EOE PROCESS

The epitaxy-on-electronics, or EoE, process was illustrated in Fig. 1.3. It begins with the design of the underlying integrated circuit chip. In addition to the obvious electronic circuitry, chips designed for epitaxy-on-electronics must in-

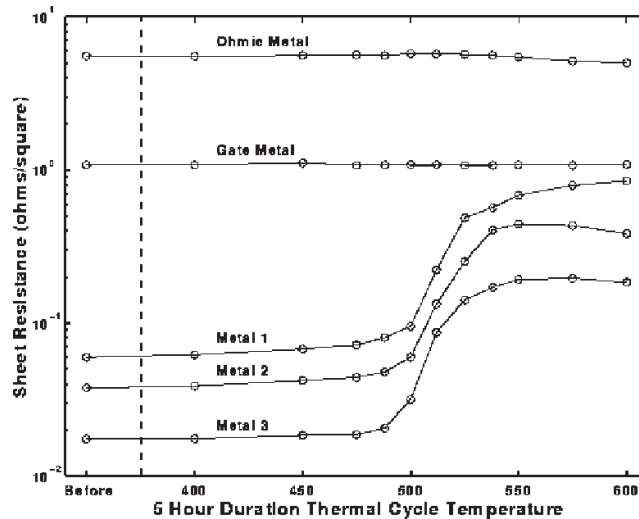


Figure 1.4. Metal layer sheet resistance in H-GaAs III test circuits after five-hour thermal cycles at the temperatures indicated on the horizontal scale (after Ref. 38).

clude dielectric growth windows (DGWs). Growth windows are regions in which the dielectric covering the wafer can be removed, exposing the original substrate surface. The structure of a growth window is illustrated in Fig. 1.5, which in Fig. 1.5(a) shows a cross section through a well on the wafer as received from the IC manufacturer. Features to note include, first, the n⁺ source/drain implant region under the well and the ohmic contact made to it on the side. This implanted region and contact will eventually provide the contact to the n-side of the device. Second, note the layer of metal above the implanted region. This provides an etch stop for the plasma etch used to remove the dielectric layers above it and clear the window to within a few tens of nanometers of the substrate surface. Dry etching all the way to the substrate leaves particles on the surface which lead to defects in the epilayers, so the dry etch is stopped at the metal layer. The metal is then removed with a wet etch, and the remaining dielectric in the well is next also wet-etched away. Photoresist protects the rest of the wafer while the last dielectric in the bottom of the well is removed. This final dielectric is actually a 50-nm-thick pair of silicon dioxide and silicon nitride layers. The former, which is the thicker of the two, etches quickly, but the later can be more difficult to etch. Consequently, a slightly undersized pattern is used to protect the upper portion of the window sidewalls from lateral etching, and a ring of ohmic contact metal surrounding the window is used to block any lateral etching of the lowest SiO₂ layer while the Si₃N₄ is being removed.

The time-temperature bound placed on the epitaxy process by the interconnect metals eliminates metalorganic chemical vapor deposition (MOCVD) from consideration because MOCVD is typically done at temperatures in excess of 700°C.

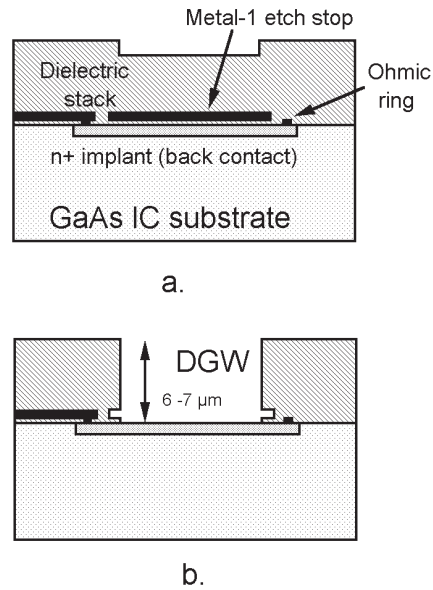


Figure 1.5. A cross-sectional sketch of a dielectric growth window as received from the GaAs IC manufacturer (a) and after cleaning and preparation for epitaxy (b).

Molecular beam epitaxy (MBE) is, however, a viable alternative. It can be done at under 475°C, although doing so does impose some compromise and does require specific capabilities. The first capability needed is a method of preparing the substrate surfaces in situ for epitaxy. Traditionally this is done by heating the substrate to over 580°C to desorb the native oxide that forms on GaAs during preparation and loading of the samples into the growth system. To avoid this high-temperature excursion, atomic hydrogen can be used to remove the surface oxide in situ at as low as 350°C. The atomic hydrogen can be obtained by a variety of methods, but perhaps the least expensive and easiest is to install a hydrogen cracker on the MBE growth chamber [41]. In such a cracker molecular hydrogen passes over a heated filament and is converted to atomic hydrogen, which is in turn directed at the sample surface. The small amount of hydrogen involved can typically be accommodated by a system's existing vacuum pumps and causes no other problems (e.g., embrittlement).

The second capability needed, or at least highly desirable, for EoE MBE is the ability to grow phosphides, specifically indium gallium phosphide, InGaP, for use in the wide bandgap layers of GaAs-based heterostructures. The more commonly used wide bandgap material is aluminum gallium arsenide, AlGaAs, but because of the high affinity of aluminum for oxygen and the deleterious effect of oxygen on its electrical properties, AlGaAs is best grown above 650°C. InGaP, on the other hand, does not suffer this oxygen problem and furthermore it is in fact difficult to grow much above 500°C because of the relatively high vapor pressures of In and P. InGaP

is a relatively new addition to the GaAs-based device picture but it has already gained wide acceptance, being used widely to produce so-called “aluminum-free” lasers [42]; it is well suited for the EoE process. Perhaps the only shortcoming of InGaP is that its index of refraction is closer to that of GaAs than is that of AlGaAs. Consequently, AlGaAs is needed to make distributed Bragg reflector (DBR) stacks for vertical cavity surface emitting lasers (VCSELs) and other resonant cavity devices. AlGaAs grown below 475°C is not a particular problem for n-type DBRs, but can be for p-type DBRs because of unacceptably high series resistance, so EoE VCSELs are best produced with dielectric p-side mirrors [43]. If AlGaAs grown at low temperature is to be used in EoE lasers, it should be kept away from the active region, and similar precautions should be used in other devices.

EoE device heterostructures are designed to be as thick as the growth windows are deep, so the final surface of the epitaxial structure is coplanar with the upper wafer surface. Thin device structures can be raised to this level by first growing a suitably thick buffer region. In the H-GaAs IV process, for example, the growth windows are approximately 6.5 μm deep. Devices that are thicker than this will extend above the surface of the wafer and this makes the postepitaxy processing more difficult. For this reason, for example, a VCSEL structure with a grown lower DBR and a deposited dielectric upper DBR is more suitable for EoE integration than is a VCSEL structure with grown upper and lower DBRs, which would be much thicker.

After epitaxy most of the wafer surface is covered with polycrystalline deposition, because growth windows typically occupy only a small fraction of the area. This deposit must be removed before the heterostructure grown in the windows is processed into devices integrated with the electronics. This is best done with chlorine using ion beam assisted etching (IBAE) or reactive ion etching (RIE) [37]. The single crystal growth in the dielectric growth windows (DGWs) is protected with photoresist during this step using a mask pattern which is 3 μm larger than the DGW. After the RIE, small polycrystalline pieces may remain around the DGW edges; if so these pieces are removed by lightly rubbing the wafer surface with a soft brush of the type used to clean wafers after a chemical mechanical polish (CMP) procedure.

At this stage the wafer surface is essentially planar and the optoelectronic device processing and integration proceeds using standard processing sequences. Any process that has been developed to produce planar or low profile devices can in general readily be adapted to the EoE environment.

Eventually it will be important to study how the DGW size for a given device can be minimized, but in early work very conservative design rules have been followed. The quality of the growth within a DGW is known to be lower near the edges of the well, but an extensive study of this effect, and of the impact of the profile of the DGW sidewalls, has not been made. As a rule of thumb it was

assumed that any defects at the edges extend outward a distance corresponding to the thickness of the structure grown, i.e. the depth of the well. Consequently the DGWs were sized so that the active region of devices fit well within the DGW and away from this border region. It was also found that the best results are obtained if all the DGWs on a given wafer or chip are roughly the same size.

1.5.5 IC PROCESS PHOTODETECTOR OPTIONS

Several structures can be produced during the course of a standard integrated circuit process which are useful photodetectors. This is particularly true of GaAs processes because the optical absorption coefficient of GaAs increases relatively rapidly above the band edge and to a higher value than is the case in silicon. Consequently a substantial amount of absorption occurs within the relatively thin active device regions near the surface and little light penetrates deeply within the wafer.

The most sensitive H-GaAs process based photodetector is the optical field effect transistor, or OPFET. The OPFET is an enhancement-mode MESFET with a large source-drain separation and a narrow, floating gate. It works because of the photovoltaic effect of the light on the back gate diode of the MESFET, which modulates the channel conductance. The OPFET is a sensitive device, responding to as little as $0.1 \mu\text{W}$ of light, but it is also a relatively slow device, with bandwidths typically being 10 to 100 kHz [33, 44].

A much faster H-GaAs photodetector is a metal-semiconductor-metal, MSM, detector made using the gate Schottky diode metal in an interdigitated finger pattern over a channel implant. In the standard H-GaAs process the exposed regions between the fingers are ion implanted with the source/drain implant dose which reduces the sensitivity of the detector significantly. Adding a masking step to the process to block the implant in these regions improves the MSM sensitivity. Another problem with an MSM made with the basic MESFET process is that the shallow channel implant is made in a thicker, lightly p-type layer. The underlying p-n junction is reverse biased and collects many of the photogenerated carriers. In early H-GaAs processes the p-region was left floating and this led to cross talk between the devices on a chip and made the design of detector circuits and, in fact, all optoelectronic integration, more complicated. Beginning with the H-GaAs IV process, p-contacts were added along with a deep n-type isolation implant that can be used to divide the p-region into isolated islands, as was described earlier. This feature dramatically reduces cross talk between devices and circuits; this makes designing OEICs with the H-GaAs technology much easier and yields improved OEIC performance. It does not, however, eliminate the p-n diode underlying the MSM photodetectors and this remains a problem. The performance of these detectors is a complex function of the biases applied to the p-region and the MSM electrodes, and of the detailed geometry of the device and p-island [35].

The problem of the underlying p-n junction can be eliminated completely by adding more steps to the basic H-GaAs process to mask the p-type blanket implant in the region of the MSMs and to make the MSM n-region with an additional, deeper implant. Such a process has been developed by IBM researchers working with Vitesse, but it is not generally available [45]. The work at MIT used circuits fabricated using the standard commercial processes with one mask added to block the source/drain alignment implant between the MSM fingers.

Another way to address the problem of the underlying p-n junction for MSMs is to look to another device. Specifically, the p-contact in the H-GaAs IV process presents the opportunity to make a lateral p-i-n photodiode [35], in which the “problem” junction becomes the operating heart of the detector; this device is currently under study. In this device the lower p-n junction is now much smaller and becomes the active region of the device, rather than a performance-limiting parasitic. The sensitivity and speed of this structure should be much better than that of the MSMs, and its fabrication does not require any additional mask steps, i.e. it can truly be made with the standard H-GaAs process.

1.5.6 EMITTERS AND OTHER EOE DEVICES

For many OEIC applications of the EoE process the photodetectors described in the preceding section are satisfactory and the objective of EoE is solely to integrate light emitting devices, that is, either light-emitting diodes (LEDs) or laser diodes (LDs). Of these, LEDs are the simpler device and they were used in the first demonstrations of the EoE technique so that the resources available could be concentrated on perfecting other aspects of the process.

Details of the LED structure and performance used in this work can be found in the literature [33, 34, 46]. A photomicrograph of a focused ion beam (FIBE) cross section made through an LED integrated on a GaAs chip using the EoE process is shown in Fig. 1.6. It is interesting to note in this picture that the heterostructure fills the dielectric growth window nicely, resulting in a surface planarity somewhat better than that of the original chip, that the interface between the original substrate and the grown heterostructure is not visible at the bottom of the well, and that the contact to the ion implanted n-region at the back of the device and the contact extending from the p-ohmic on the top of the device over to the metal pad provided to the side of the DGW are both visible. The device pictured in this figure was grown in a 50- μm -by-50- μm DGW on OPTOCHIP (see the next subsection).

Less effort has been devoted to integrating laser diodes by EoE, and the nature of the problem has evolved with time to the point where the device structure of choice is actually simpler than it was originally. It can be argued that the time is now right for increased effort toward EoE laser diodes. When the EoE effort began, the options for vertically emitting laser diodes were three: in-plane surface-emitting lasers, IPSELs [47]; grating-coupled surface-emitting lasers, GCSELs

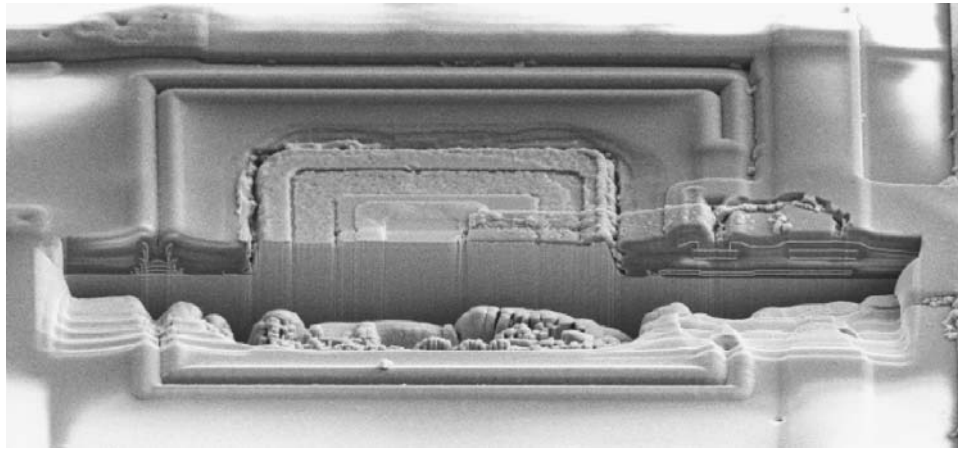


Figure 1.6. A scanning electron microscope image of a focused-ion-beam etched (FIBE) vertical cross-section through a completed EoE-integrated light-emitting diode (LED) on OPTOCHIP. Of particular note is the overall planarity of the structure, the relative smoothness of the epitaxial side-walls, and the lack of any distinguishable interface between the epitaxial heterostructure and the substrate. The top and bottom contacts made from the pre-existing electronics to the LED are also clearly visible. The FIBE cross section and photomicrograph were made by Dr. K. Edinger and Prof. J. Melngailis of the Laboratory for Ion Beam Research and Application at the University of Maryland, College Park, MD.

[48]; and vertical-cavity surface-emitting lasers, VCSELs [49]. Early VCSELs were typically 10 μm or more thick, had operating voltages well in excess of the 2-V GaAs DCFL power supply, and were themselves a topic of intense research. GCSELs also required much research and development to be suitable for OEIC applications. At the same time, a group at MIT Lincoln Laboratory had recently invented a novel IPSEL structure with dry-etched vertical end mirrors and a dry-etched parabolic deflector at the output end to form the beam and direct it vertically [47]. Initial EoE circuits were designed around this laser, and some effort was directed at realizing it [50]. It soon became clear, however, that Al-free lasers would be needed, and this in turn necessitated developing a dry-etch chemistry for use with these heterostructures, before the Al-free IPSEL and its EoE integration could be attempted. It also, of course, necessitated development of growth capabilities and procedures for Al-free lasers and the optimization of those devices. A new nonselective etch chemistry for InGaAs/GaAs/InGaP heterostructures was developed [51], and Al-free IPSELs were produced [52]. Work on the next step, EoE integration of IPSELs with GaAs, is continuing.

Concurrent with the IPSEL work described above, substantial advances were made by VCSEL researchers and the threshold voltages and currents of these dev-

cies have been reduced significantly, and so too has their overall thickness. In particular VCSELs using a deposited dielectric mirror [43] are thin enough to fit satisfactorily in a dielectric growth window, and are an attractive option. EoE-integrated VCSELs are not being actively pursued at this writing, however, because the focus is instead on integrating VCSELs by aligned pillar bonding (APB). It is, however, important to note that APB VCSELs involve the same processing after the heterostructures are bonded in the DGWs as do EoE VCSELs after epitaxy and removal of the polycrystalline deposits. Thus much of the technology developed for APB VCSEL processing will apply for EoE VCSELs as well.

Other devices which have been EoE-integrated on GaAs ICs are resonant tunneling diodes, RTDs [53] and self-electro-optical devices, SEEDs [54].

1.5.7 THE OPTOCHIP PROJECT

While lasers are the most desirable emitters for many OEIC applications, there is also a great deal of interest in integrated LEDs and much can be done with them. Consequently, in 1995 it was decided to offer a research foundry service, the OPTOCHIP Project, which would make the EoE technology, rudimentary at that stage as it was, available to the OEIC applications research community. Eighteen groups submitted proposals to participate and nine groups were selected. Those needing it were given basic training in the design and layout of DCFL digital circuits, and all were given “drop-in” cells for an MSM detector, an OPFET, an LED driver, and a dielectric growth window (DGW) cell for an LED. In addition to a 50-by-50-micron growth opening, the DGW cell also included the n+ back contact implant, the ohmic contact ring, and a light shield surrounding the opening.

Each group participating in OPTOCHIP designed a 2-mm-by-2-mm chip and in May 1996 the designs were collected and combined into a single 7-mm-by-7-mm chip, pictured in Fig. 1.7, which was then submitted to MOSIS for fabrication (the Vitesse H-GaAs III process was used). Processed dies were received in August.

The DGWs were opened on OPTOCHIP dies, and LEDs were grown in them and processed at MIT. The completed OPTOCHIP dies were cut into the individual 2-by-2-mm chips and each of the participating groups received three copies of their chip in May 1997 [33–35]. Several groups did not test their chips, but other groups did do extensive studies on theirs and were able to incorporate them into demonstrations of their technology [55, 56].

The 50-micron DGW size proved to be somewhat small and appeared to adversely effect LED yield. All of the 196 LEDs on each chip functioned, but the efficiency varied widely, meaning that one third of the LEDs worked very well, one third were mediocre, and one third emitted very low levels of light [33]. For comparison, LEDs in larger, 100-by-100-micron, wells in the monitor and test cells consistently gave bright efficient emission. The difficulty is thought to have been due to problems with the well cleaning procedures during the OPTOCHIP

processing, and not to be an intrinsic limit of the EoE process, since in other cases DGWs as small as 10 microns on a side have been used.

Another problem encountered on some OPTOCHIP designs concerned the optical cross talk seen in the H-GaAs III process, as was mentioned earlier. This limits the performance of some of the cell designs. Processing additional OPTOCHIP dies using lessons learned from the first series could have yielded high-brightness LEDs in all DGWs, but would not have dealt with the optical cross-talk issues, so it was felt that the resources available were better spent on other aspects of the EoE process. Future OPTOCHIP Project offerings can use the H-GaAs IV process (or its successor) which will eliminate backgate-related optical cross-talk problems, and will offer VCSELs rather than LEDs, and lateral p-i-n detectors, rather than MSM detectors.

The OEIC chips produced in the OPTOCHIP Project represent the most complicated, dense OEICs ever produced by any technique, and it is noteworthy that the multitude of circuits on OPTOCHIP were produced by relatively novice

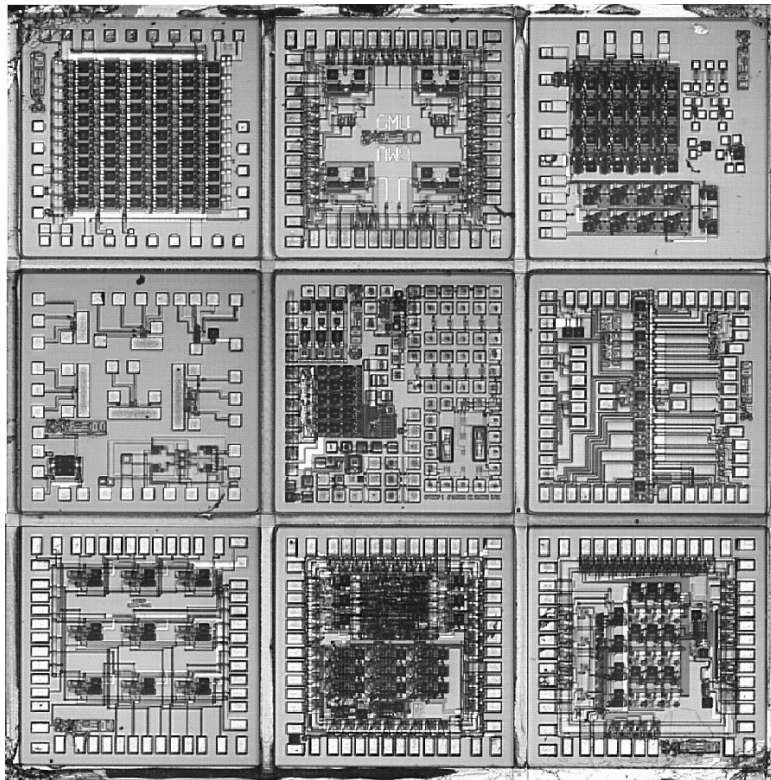


Figure 1.7. A photomicrograph of a completed OPTOCHIP master die. The die is composed of nine 2-by-2-mm OEIC designs. Including the scribe line patterns, the master die measures 7.3 mm on a side (after Ref. 32).

designers using totally standard design tools and standard cell libraries, and were produced in part on a commercial production line, and in part in the laboratory using techniques similar to those used by optoelectronic device manufacturers. The OPTOCHIP Project demonstrated that the EoE process for monolithic optoelectronic integration is accessible to users and is manufacturable; it is real and it is available here and now.

1.6 SILICON-ON-GALLIUM ARSENIDE

1.6.1 LESSONS FROM SOI AND MEMS

In spite of the dramatic advances in GaAs integrated circuit technology in the past decade, the importance of developing optoelectronic integration technologies built around silicon integrated circuits, particularly CMOS, has only increased. Both the pervasiveness of Si, and the difficulty of producing large high-density memory arrays, high-performance microprocessors, etc., in GaAs, mean that silicon integrated circuits, or Si ICs in companion with GaAs ICs, must be available to optoelectronic integrated circuit designers before OEICs will be taken seriously. The flip-chip bump bonding and ball-grid array hybrid assembly techniques described in the other chapters of this volume provide access to Si ICs, but a monolithic solution to this problem is still needed.

The gallium arsenide on silicon work described in Section 1.4, which was the original path followed to silicon-based OEICs, failed because of the large thermal expansion coefficient (TEC) mismatch between Si and the III-Vs, and provided an important lesson on the importance of paying attention to TECs, at least for optoelectronic devices. At the same time, ongoing work on silicon-on-insulator (SOI) technologies can provide an important lesson on the impact of TEC mismatch on Si ICs. In particular, one learns from SOI work that very little silicon is needed to fabricate silicon integrated circuits. Many modern SOI processes (specifically, “fully-depleted” SOI processes) require a silicon layer only a few tens of nanometers thick. Such a thin layer is very tolerant to stress and one can anticipate that it would not be damaged by the TEC mismatch between Si and GaAs. That this is indeed the case is shown by the success of a subclass of SOI circuits, silicon-on-sapphire (SOS) integrated circuits. The TEC of sapphire is comparable to that of gallium arsenide so the silicon in an SOS wafer experiences a level of stress very much like that it would experience if one had a thin layer of silicon grown or bonded on a GaAs substrate. That is to say, Si-on-GaAs can work because only a thin layer of Si is needed to make a Si IC, whereas GaAs-on-Si does not work because GaAs-based optoelectronic device heterostructures must be thick.

The way to form a silicon-on-gallium arsenide (SonG) wafer is found again in work on SOI, as well as in work on three-dimensional integration of silicon

devices, and work on micro-electro-mechanical structures (MEMS). Oxide-to-oxide bonding is widely used in all of these efforts and this work provides a wealth of relevant experience with, and information on, this process that can be used in making SonG wafers. Very importantly, this body of work also demonstrates that wafer bonding is a commercially viable process.

Putting these observations together one realizes that a possible solution to the problem of producing monolithic OEICs incorporating silicon electronics is to build SOI electronics on SOI wafers in which the “I” is semi-insulating gallium arsenide, and to then use these processed SOI wafers as the foundations for epitaxy-on-electronics integration of optoelectronic devices. The relatively thicker, strain-sensitive optoelectronic devices are thereby provided with their optimum substrate, and the very thin, strain-resistant silicon electronic devices are also uncompromised by this arrangement. We call these SOI wafers silicon-on-gallium arsenide, or SonG, wafers.

1.6.2 PRODUCING SONG WAFERS

The process which has been developed to produce silicon-on-gallium arsenide wafers was built upon three-dimensional (3D) integration work done by Professor Dimitri Antoniadis and his students, who have developed a technique of bonding oxidized SOI wafers to oxidized Si wafers and then removing the original substrate of the SOI wafer using the buried oxide layer as an etch stop to leave the Si from the SOI wafer bonded to the other Si wafer [57]. The Antoniadis group worked closely with researchers at MIT Lincoln Laboratory, especially Andrew Loomis, who have developed expertise in low-temperature oxide-to-oxide wafer bonding and in back-side thinning of Si wafers [58]. Both the campus and Lincoln Laboratory groups contributed their knowledge generously toward the development of the SonG process described in the following.

The SonG process, which is illustrated in Fig. 1.8, begins with GaAs wafers and SOI wafers [59, 60]. Four inch diameter wafers were used to be compatible with the processing equipment available. The GaAs wafers used in this work were coated with a PECVD (plasma-enhanced chemical vapor deposition) silicon oxide layer by Vitesse Semiconductor following procedures they have developed and use as one of the initial steps in their H-GaAs processes. The SOI wafers were SIMOX wafers from IBIS Corporation (“SIMOX” stands for Separation by IMplantation of OXygen, which reflects the process used to produce them). The SIMOX wafers were thermally oxidized and then a borophosphosilicate glass (BPSG) was deposited on them at 400°C and densified at 925°C. Both types of wafers were given a chemical mechanical polish (CMP) to remove any hillocks or other surface roughness that might have existed.

The wafers are then thoroughly cleaned, given a brief hydrofluoric (HF) acid etch to produce a hydrophilic surface, and bonded under a slight pressure.

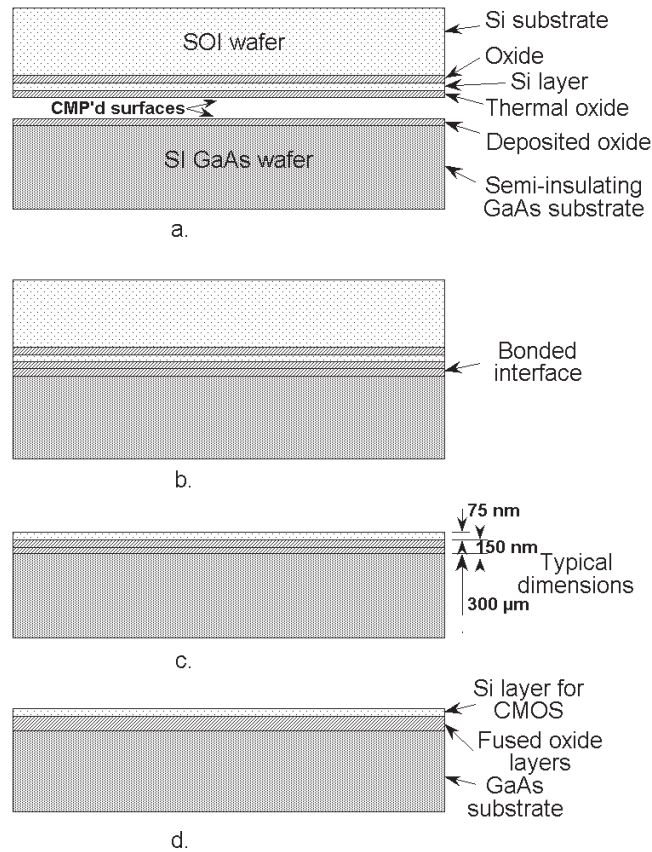


Figure 1.8. The SonG process for producing EoE-compatible SOI wafers: (a) the bulk GaAs wafer and the SIMOX wafer prepared for bonding, and placed face to face; (b) after hydrophillic room temperature bonding; (c) after removal of the SIMOX wafer substrate; (d) after high temperature fusion of the bond to complete the SonG wafer fabrication.

The bond between the wafers is sufficiently strong at this point that they will break if the pair is heated above 335°C because of the TEC difference, but the bond should still be hardened further by heating the pair below this temperature. The optimum anneal cycle has not yet been determined but a cycle of 15 hours at 150°C followed by 15 hours at 250°C is typical of those found to work satisfactorily.

After the postbonding anneal, the substrate of the original SIMOX wafer is removed in two etching steps. First, the bulk of the substrate is removed in a jet-etching apparatus developed at Lincoln Laboratory, leaving only a few tens of microns of Si above the buried oxide layer of the SIMOX wafer. The remaining Si is then etched away cleanly using a selective etch, tetramethylammonium hydroxide (TMAH) at 70°C, which does not etch the buried oxide layer. The resulting SonG wafer can now be annealed at up to 700°C (and, likely, higher, but this is the

highest attempted to date) without negatively effecting the structure. Doing this undoubtedly further strengthens the bond.

A photograph of a 4"-diameter SonG wafer is shown in Fig. 1.9, and in Fig. 1.10 a photomicrograph of a cross section through the top layers on a SonG wafer is shown. All of the layers are visible, but the bond itself is difficult to identify with any precision. Instead, a diffuse line is seen in the region where the bond is expected to be, indicating perhaps that the wafers have indeed been fused together.

There are a variety of other procedures which might also be used to make SonG wafers. Two of particular note are the "Smart Cut" process [61], and plasma-activated bonding [62]. Smart Cut is the trademarked name for an exfoliation process of forming SOI wafers which involves implanting silicon wafers with a heavy dose of hydrogen atoms, bonding them at low temperature to another oxidized silicon wafer, and then heating the pair to a high temperature. The implanted wafer is weakened along the plane coincident with the peak of the implanted ion profile and the splits along this plane. If the wafer is not bonded to another wafer the silicon flakes off, i.e. exfoliates, but if the wafer is bonded to another wafer (as it is in the Smart Cut process), the thin silicon layer splits cleanly free

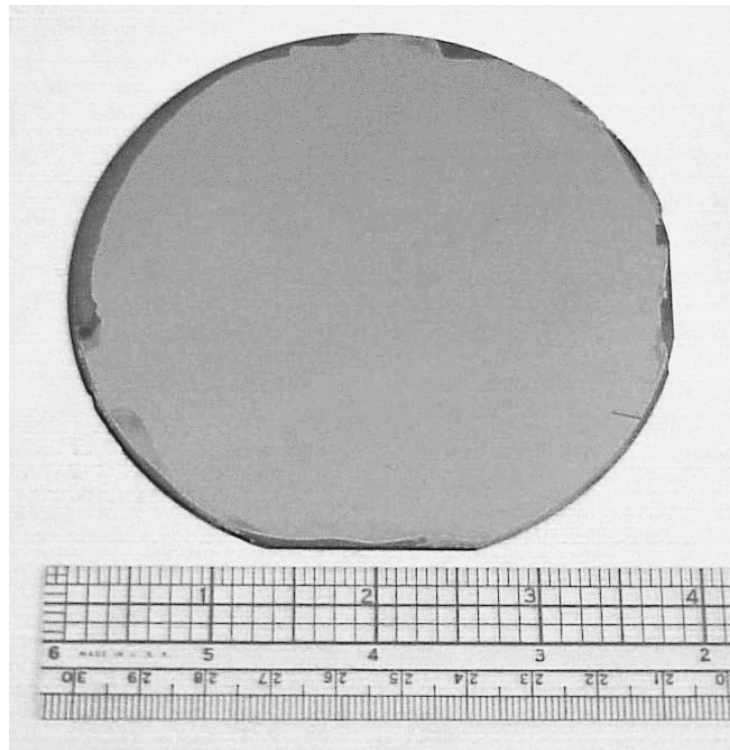


Figure 1.9. A photograph of a four-inch-diameter silicon-on-gallium arsenide (SonG) wafer [53].

from its original substrate and remains bonded to the second wafer. The newly revealed surface is chemically and mechanically polished (CMP'd) to smooth it, and the SOI wafer is ready for use. This process has been demonstrated on a variety of materials and has been used to produce single crystal films of a variety of materials on Si substrates. The temperature at which the exfoliation has been lowered by implanting multiple species, such as boron and hydrogen. One can thus anticipate that it should be possible to use this technique to produce SonG wafers, although this has not yet been attempted.



Figure 1.10. A scanning electron microscope image of a focused-ion-beam etched (FIBE) vertical cross section of a completed SonG wafer showing the various layers in the structure viewed at an angle of 60°. The 130-nm silicon layer and the GaAs substrate are clearly visible, as is the location of the bonding interface between the BPSG layer on the SIMOX wafer and the PECVD oxide on the GaAs wafer [59]. The top of the sample was coated with a layer of tungsten prior to FIBE processing; this layer is also visible. The FIBE cross section and photomicrograph were made by Dr. A. Stanishevsky and Prof. J. Melngailis of the Laboratory for Ion Beam Research and Application at the University of Maryland, College Park, MD.

The plasma-assisted bonding technique [62] is a method to produce very strong room-temperature oxide-oxide bonds. It is being used by Silicon Genesis Corporation [63] along with deep hydrogen implantation to produce SOI wafers commercially. In this process, the details of which are largely proprietary, the oxide surfaces to be bonded are activated in a plasma and the wafers are contacted in situ. The bonds formed this way at room temperature can be sufficiently strong that the top layer of the hydrogen-implanted wafer can be parted from its original substrate, without the need to heat it above room temperature, by directing a high-velocity jet of nitrogen at the edge of the wafer near the peak of the implanted profile [63].

Before proceeding to the next subsection, it is worth anticipating some of the future discussion and noting that while the exfoliation technique can be used to produce SonG wafers, it is not as useful for bonding processed wafers (an objective that will be discussed in Section 1.6.4). This is so because the hydrogen implant must be done after all high-temperature processing, meaning it is likely that the electronics will be fully processed. Any device structures and metal patterns on the processed wafer will locally distort the implanted profile, and prevent formation of a clean exfoliation plane. The plasma-assisted bonding technique, on the other hand, is more likely to be useful even for bonding processed wafers, but even here some caution is in order. In particular, it will be important to determine whether the plasma activation process damages any of the underlying electronic devices and/or circuitry. It is conceivable that electrostatic discharge (ESD) damage may occur, but given the wide use of plasma processing in many back-end processes, this seems unlikely.

Another point to address is that of wafer size. Present GaAs processes are limited to 6"- (150-mm) diameter wafers, whereas 8"- (200-mm) diameter Si wafers are standard in new fabs, and fabs using 12"- (300-mm) diameter wafers exist. An obvious concern is whether the lack of 8"- and 12"-diameter gallium arsenide wafers on which to bond Si layers will be a fatal shortcoming of the SonG process. A way to partially address this concern is to note that the transition from 4"- to 6"-diameter wafers occurred very quickly in the GaAs industry and the same economic pressures that drove that change, and the change from 6" to 8" Si, will undoubtedly drive the development of 8"-diameter GaAs wafers. Success of the SonG process can only hasten that change.

We should also note that the SonG technique can be extended to other semiconductor substrates. Of particular interest to optoelectronic integration is indium phosphide, and the same techniques used to produce SonG substrates could be used to produce silicon-on-indium phosphide, or SonI (pronounced "sunny"), substrates for EoE integration of 1.3- and 1.55- μm devices with Si CMOS.

Finally, we should note that process temperatures would be limited during any CMOS processing of SonG wafers because of the instability of GaAs at elevated temperatures. The backside of the SonG wafer can be coated so no GaAs is exposed, but it is still likely that a reduced-temperature process would be desirable

(with InP the situation would be even more critical). Such processes do exist, but a better solution is to do the silicon processing before bonding, a process discussed in Section 1.6.4.

1.6.3 EPITAXY ON SONG

SonG substrates were developed for the epitaxy-on-electronics process so demonstrating successful epitaxy on them is a key step in their development. This was done recently [64] by opening dielectric growth windows in a piece of SonG material and growing an InGaAs heterostructure in the DGWs following normal EoE procedures (see Section 1.5). The resulting material was then tested using photo- and cathodoluminescence. Comparison was made with material grown simultaneously on a conventional epi-ready gallium arsenide substrate, and it was found that the spectra from the two samples were comparable in structure, but that the intensity of the luminescence from the growth in the windows on the SonG substrate was one third that of the growth on the epi-ready substrate. In as much as the growth involved a relatively thin buffer layer, it is likely that the lower intensity is due to the semi-insulating substrate of the SonG wafer and that using a thicker buffer would result in increased intensity, but further work is needed to clarify this point. The luminescent intensity of the growth on both wafers is consistent with obtaining useful device performance from the material.

1.6.4 CMOS ON GaAs FOR EOE AND 3D INTEGRATION

One can envision replacing the SOI wafers currently used in a commercial reduced-temperature CMOS process with SonG wafers (with their backsides suitably coated to prevent their decomposition), but doing this in practice is likely to involve a number of challenges, not the least of which is convincing a silicon manufacturer to allow gallium arsenide in its production facility. Processing SonG wafers in a research facility is more feasible, but the ideal solution would be to completely avoid processing CMOS on SonG wafers. The answer lies not in bonding virgin SOI wafers to GaAs, but in bonding fully processed CMOS wafers to GaAs; that is, in doing the CMOS processing before, rather than after, the SonG processing. This approach is illustrated schematically in Fig. 1.11.

The adoption of chemical mechanical polishing (CMP) in back-end IC processing in the past several years fits well with the concept of bonding processed IC wafers because CMP-based processes yield wafers that are much more planar than traditional multilevel metal wafers. They are often even laid out according to guidelines designed to enhance planarity. Nonetheless, wafer bonding requires not just a nicely planar surface, but what would better be called a highly planar surface, one comparable to that of a bare silicon wafer before any processing. A modern back-end IC process does not provide this level of planarity, but it does provide an excellent starting point from which work toward achieving this level of planarity.

Some research toward three-dimensional integration of silicon ICs involves low temperature oxide bonding of processed wafers and developing techniques to form small area interconnecting electrical vias between the layers. All of this work provides useful guidance and some solutions for the present effort, but there is a great deal yet to be done. The planarization of processed CMOS and CMOS SOI wafers for bonding using a variety of deposited and spun-on dielectrics in combination with interlevel CMP is currently being studied.

The logical extension of this work bonding processed CMOS wafers on GaAs wafers is to also use processed GaAs IC wafers. That is, to do three-dimensional multiprocess integration of Si-CMOS, III-V electronics, and optoelectronics by

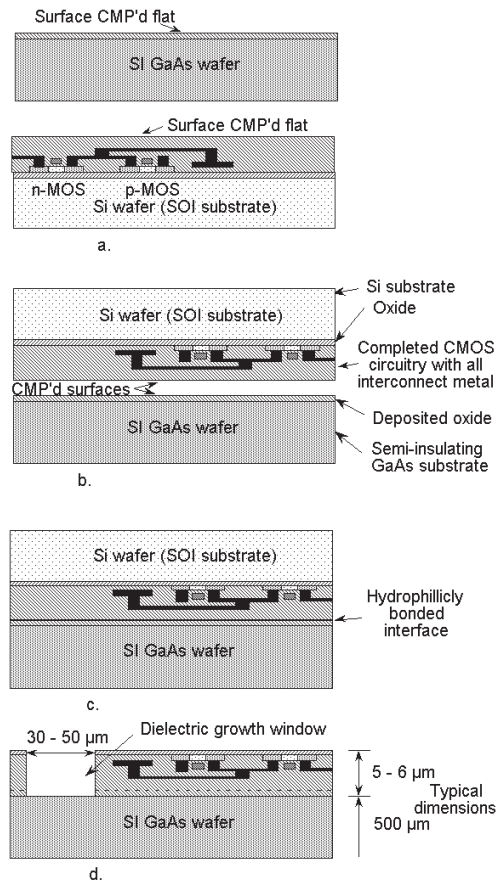


Figure 1.11. The SonG process executed with preprocessed SOI CMOS wafers: (a) the bulk GaAs wafer and the processed SOI CMOS wafer prepared for bonding, and (b) placed face to face immediately prior to bonding; (c) after hydrophillic room-temperature bonding and prior to removal of the CMOS wafer substrate and high-temperature fusion of the bond; (d) after substrate removal, bond fusion, and preparation of windows for EoE or APB processing as shown in Figs. 1.3 and 1.12, respectively.

bonding processed Si wafers and processed III-V IC wafers, interconnecting the circuitry as appropriate with vias, and adding monolithically integrated optoelectronic devices.

In many situations, the three-dimensional integration of III-V and Si electronics will be a useful end in itself, and optoelectronic integration will not be a part of the design. The ability to add memory and dense (albeit, moderate speed) Si digital signal processing capabilities to multigigabit GaAs and InP MESFET, HEMT, and HBT-based micro- and millimeter wave integrated circuits (MMICs) and digital circuits will in itself be very significant and empowering. The resulting “intelligent” MMICs and “enhanced” GaAs digital ICs are sure to find many applications. Adding optoelectronic devices will be desirable in some instances, while in others doing so will be unnecessary.

1.7 ALIGNED-PILLAR BONDING

1.7.1 THE APB CONCEPT: MERGING EOE AND BONDING

The epitaxial growth step that is intrinsic to the epitaxy-on-electronics (EoE) process makes EoE a fully monolithic process and has numerous advantages, as we have seen, but it comes with some costs. Specifically, the manufacturer or researcher using the EoE process must have epitaxy capabilities, including the abilities to do atomic hydrogen oxide removal and to grow phosphides. Second, they must have developed and be able to reliably control and reproduce the sophisticated growth sequences needed for the advanced device structures of interest, and they must be able to execute them with high yield. Furthermore, those growth sequences must be EoE compatible, meaning that they must occur entirely below 475°C, and be suitable for use on ion-implanted, semi-insulating substrates. If there was a way other than direct epitaxy to get optoelectronic device heterostructures on processed electronic wafers, while still retaining the planarity and monolithic aspects of the EoE process, one could avoid these “costs” and get the best of both worlds. (It should be noted that for some groups, the “costs” of epitaxy referred to are negligible; for others they are very major.)

Aligned wafer bonding, which has been brought to a high state of development by the MEMS community, may offer the solution and provide an attractive alternative to epitaxy. This is the thought behind the aligned pillar bonding (APB) process. The idea is to grow the optoelectronic heterostructure separately on its optimum substrate and under optimum conditions. This can be done either by the organization doing the APB integration, or by a collaborator or contract foundry. The epitaxial wafer surface is then etched to form heterostructure pillars in a pattern mirroring that of the dielectric growth windows (DGWs) on the electronics wafer (i.e., the wafer on which one would do epitaxy in the EoE process sequence).

The two wafers are then brought face-to-face, aligned, and bonded, as illustrated in Fig. 1.12. The substrate of the wafer on which the heterostructures were grown is then removed, leaving optoelectronic device heterostructures filling the DGWs just as if they had been grown there directly. From this point on, the fabrication continues just as in the EoE process.

Aligned pillar bonding appears at first to be similar to processes developed recently at Lucent [7], Georgia Tech [65], and Colorado State [66], but it differs in several important ways. In the work referred to, processed device heterostructure chips are bonded on top of CMOS chips (typically via metal pads rather than solder balls), and the substrate under the heterostructure devices is then removed leaving separated devices bonded over the surface of the CMOS chip. In APB,

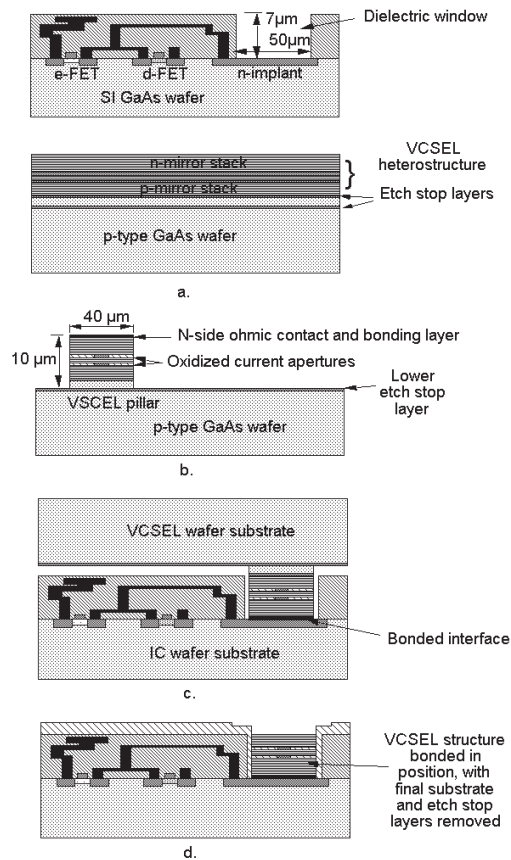


Figure 1.12. The APB process: (a) the processed IC wafer and p-side down VCSEL wafer as received from the manufacturer and grower, respectively; (b) the p-side down VCSEL wafer with pillars etched to match the windows on the IC wafer; (c) after bonding of the VCSEL and IC wafer; (d) after removal of the substrate of the VCSEL wafer leaving VCSEL heterostructures bonded in the windows.

unprocessed, or only partially processed, heterostructures are bonded, and a major fraction of the device processing is completed after bonding and substrate removal. Second, in APB the heterostructures are bonded in recesses in the processed electronics wafer surface (i.e., in the DGWs) so that their upper surface is subsequently coplanar with the upper surface of the processed wafer. Finally, using MEMS precision bonding equipment, APB can be extended to bonding small dimension pillars with very small alignment tolerances.

APB shares several features with the work of the three groups mentioned in the preceding paragraph. The epitaxy can be done on an optimal substrate under optimum conditions, and by a group, or in a foundry, experienced in the device in question, i.e., it need not be grown by the group doing the integration. Also, a heterostructure can be evaluated and rejected if it is unsatisfactory, before proceeding with bonding and further processing; the same, of course, is true of the electronics wafer. And, finally, removing the substrate to decouple the individual device islands or pillars in this way can potentially relieve some of the thermal expansion coefficient mismatch concerns. This is true, however, only if the bonding can be done very near room temperature (or if, as is the case in the references cited, processing is restricted to individual chips). That is to say, TEC differences will effect wafer-to-wafer alignment during bonding and so they are very important at that stage. Once the wafers are bonded and substrate has been removed, however, the TEC differences will only impact the devices individually since they are no longer all rigidly coupled through their original substrate across the entire wafer. Some devices can tolerate the levels of stress they will experience in this situation. Thick VCSELs, for example, wherein the active layer is 5 to 6 microns away from the bonded interface, may be insensitive to strain at that interface. The options, then, if one is interested in full-wafer integration and batch processing, are to do the bonding at room temperature, or to use TCE-matched wafer combinations. We begin doing the latter, while seeking techniques to do room temperature assembly. TCE-matched VLSI wafers that can be used with heterostructures grown on GaAs are GaAs VLSI wafers, silicon-on-GaAs (SonG) VLSI wafers, and silicon-on-sapphire (SOS) VLSI wafers. Significantly, the first and last are both available commercially. For heterostructures grown on InP, the APB option is silicon-on-indium phosphide (SonI) VLSI.

1.7.2 BONDING OPTIONS AND OTHER PROCESSING ISSUES

The basic choice one faces in developing the APB process concerns the bonding process. Oxide-to-oxide bonding such as used in the SonG process is not attractive because it is in general desirable to use a technique that forms a conducting bond so the bond can be used to form one or more of the contacts to the device. Direct semiconductor-to-semiconductor bonding can satisfy this requirement but in general involves prohibitively high temperatures [67]. This leads one

to consider metal-to-metal and metal-to-semiconductor bonding, both of which can provide low resistance ohmic contacts and can be done at lower temperatures [68]. Work has focused on the latter, metal-to-semiconductor bonding, because it can be done at low temperatures and because it involves processing (i.e., metallizing, patterning, etc.) only one of the surfaces involved. It can be anticipated, however, that metal-to-metal bonds might ultimately enable room-temperature assembly once a suitable bonding agent (conducting epoxy, liquid eutectic solder, etc.) is identified. This agent must be able to, first, hold the structure together until the substrate is removed to leave the heterostructures bonded in the DGWs, and, second, be subsequently hardenable, if necessary, by a higher temperature step.

It should also be noted that both of the bonding techniques involving metals (i.e., semiconductor-to-metal and metal-to-metal) give one the option of APB'ing heterostructures to metal pads formed in upper-level interconnect layers (as long as the heterostructures do not need to be so thick that they protrude an unacceptable distance above the wafer surface). This in turn means that there can be devices underneath these metal pads, leading to higher device density and true 3D integration.

Another constraint on the bonding layers is that they be as thin as possible, while also being sufficiently malleable and thick to compensate for any surface irregularities (which will primarily be long range deviations from planarity such as wafer warpage). The objective is to keep the total bond metal thickness to the order of a micron or less.

The APB work to date has used two bonding metallurgies. One is palladium bonding to gallium arsenide, and the other is gold-tin bonding to GaAs. Both have been successfully used to align, bond, and transfer GaAs pillars in DGWs on GaAs IC chips; gold-tin bonding to GaAs is the current method of choice.

A second important process issue is the technique used to remove the substrate of the heterostructure wafer. In the initial APB work the space between the wafers has been back filled with an organic polymer and then the substrate has been etched away, relying on the polymer and one or more etch-stop layers to protect the pillars and the electronics. More recently techniques have been devised, and are under development, which involve etching away "sacrificial" or "etch-release" layers in the heterostructures and that do not require back filling. Most importantly these techniques allow one to reuse the substrate, a feature that will, of course, yield major economic advantages.

APB is currently a topic of active investigation, with the immediate objectives being the APB integration of VCSELs on OPTOCHIP dies, and of 1.55- μm p-i-n detectors on a transimpedance amplifier chip, OEIC-6 [69].

1.7.3 ADVANCED APB: SELF-ASSEMBLY AND PREFORM WAFERS

Both EoE and APB are wasteful of epitaxial material in as much as in both techniques the vast majority of the material deposited during epitaxy is subsequently

etched away, to remove the polycrystalline deposit in the case of EoE and to form the pillars in the case of APB. Selective area epitaxial deposition could eliminate this waste, but this can currently only be done at relatively high temperatures and is not a viable option at present. Another possible solution is to etch the APB heterostructure pillars free from their substrate before they are bonded to the VLSI wafer and reassemble them in the correct pattern on a carrier wafer using self-assembly techniques [70, 71, 72]. The carrier wafer can then take the place of the pillar-etched heterostructure wafer in the current APB process; it should be the same material as the VLSI wafer so TEC mismatch concerns do not arise. Self-assembly techniques might in fact be used to eliminate the carrier wafer altogether by locating the heterostructure pillars directly in the DGWs on the VLSI chip preparatory to bonding them in place. Just as the development of the EoE, SonG, and APB techniques has benefited from work in other areas, work on self-assembly for other applications will undoubtedly provide guidance in reaching this advanced APB goal also.

Self-assembly of heterostructure pillars has several very important advantages in addition to, and probably more significant than, more efficient use of epitaxial material. In particular, self-assembled APB can be done with silicon wafers, and the individual heterostructure pillars can be tested prior to assembly and bonding. With respect to this testing, it can be done at the wafer level prior to etching the pillars free, and defective die can be rejected.

An easy way to appreciate the power of self-assembled APB is to focus on the problem of chip-to-chip optical interconnect and the example of the individual heterostructures being VCSELs. VCSEL pillars that can be self-assembled in any desired pattern on silicon wafers would be the optical equivalent of metal solder bumps now used for electrically interconnecting chips on multichip modules (MCMs). These “optical solder bumps” would be driven by a suitable VCSEL driver circuit, just as conventional solder bumps are driven by a suitable line driver circuit, and the VCSELs’ output light would be directed into and travel through rectangular dielectric waveguides on the MCM substrate to receiver optical solder bumps that could deliver the data to the appropriate gates. In one of these optically interconnected chips, then, many of the solder balls in what would otherwise be a conventional metal ball grid array would be replaced with VCSEL optical solder bumps and others would be detector/receiver sites; the rest would still have conventional metal solder bumps. The companion MCM substrate would contain several patterned optical waveguide “interconnect” layers in addition to the normal layers of patterned metallic electrical interconnects. Viewed in this way, one can see that self-assembled APB promises to make interchip (and even intrachip) optical interconnect as accessible and straightforward to IC designers as the equivalent present electrical interconnects, but with none of the disadvantages of electrical interconnects and with all of the advantages of easier design, higher performance, and general robustness that optical interconnects offer.

1.8 CONCLUSION AND ACKNOWLEDGEMENTS

Three technologies for monolithic optoelectronic integration have been discussed in this review, with the emphasis being placed as much, or more, on the philosophy and perspectives behind their development as on the quantitative technical aspects more traditionally the focus of scientific literature, aspects which have been, and will continue to be, covered in the references and in future articles. The epitaxy-on-electronics technology is the most mature of the three, and is available and accessible to the research community at large. The primary barrier to its use is its reliance on gallium arsenide integrated circuitry. Work on the silicon-on-gallium arsenide technology is much younger, being about two years old, but the technology to produce SonG wafers is already available, and soon techniques to bond SOI CMOS to GaAs will also have been developed. The exciting future for this work lies as much in the area of 3D integration of silicon and compound semiconductor electronics as it does in optoelectronic integration. The final process, aligned pillar bonding, is the “youngest” of the three, but it too generates a great deal of excitement and anticipation. Taken to its extreme with the optical solder bump concept, the possibilities are large.

The development of EoE, SonG, and APB has involved many students and collaborators, and the list continues to expand; the following enumeration is undoubtedly incomplete, but will serve to indicate the scope of the effort. The work began with the thesis work of Dr. K. V. Shenoy, and he quickly developed collaborations (many of which still continue) with Prof. D. Psaltis and his student A. C. Grot at CalTech, with J. Mikkelsen at Vitesse Semiconductor, Dr. B. Ulman and his colleagues at GTE Sylvania, Waltham, MA, Dr. P. Grailovic at Polaroid, Dr. W. Goodhue at MIT Lincoln Lab, Prof. S. Prasad at Northeastern University, and Prof. W. Wang at Columbia University. The early EoE work at MIT was done in part under the DARPA-funded National Center for Integrated Photonics Technology headquartered at the University of Southern California (USC), and benefited from input from the other participants in that center, including Prof. L. Kolodziejski and Dr. G. Petrich at MIT, Prof. A. Sawchuk and Dr. C. Kuznia at USC, and Prof. R. Osgood at Columbia. Dr. J. F. Ahadian continued the direct path of Shenoy and has brought EoE to its present state of development and refinement. Along the way he was aided by other graduate students working on and around the EoE program (at MIT unless otherwise noted), including E. Braun, S. Choy (U-Mass. Lowell), P. T. Vaidyanathan (Northeastern U.), Dr. R. Aggarwal, Dr. Y. Royter, Dr. H. Wang, S. Patterson, and Dr. P. Martin. This work also involved Dr. A. Postigo, K. H. Choy, and T. Knoedl (University of Ulm, Ulm, Germany). Ahadian’s work is being carried on at MIT by E. Barkley, and collaborative efforts on EoE are in place at Nanyang Technical University in Singapore under Prof. S. F. Yoon, and at the University of Madrid under Dr. A. Postigo, work-

ing with Prof. F. Briones. The SonG effort was begun by J. M. London who was aided by very important collaborations with Dr. A. H. Loomis of MIT Lincoln Laboratory, and with important input and assistance from Prof. D. Antoniadis (and his research group), Prof. M. Schmidt, Dr. G. Petrich, and T. Takis at MIT, J. Mikkelsen of Vitesse, and Dr. L. Allen of Ibis Corporation. London and E. Barkley are continuing the SonG effort. The APB effort began two years ago, also, with initial explorations by D. Crankshaw, Dr. H. Wang, and Dr. G. Lullo, and is being continued as a research focus of W. Giziewicz.

Developments in any field depend as much on circumstances and advances in other fields as much as on those in the field itself. We have pointed out repeatedly in this review how the development of the monolithic optoelectronic technologies described—epitaxy-on-electronics (EoE), silicon-on-gallium arsenide (SonG), and aligned pillar bonding (APB)—have built upon lessons from other, often quite different, research efforts and fields. So too will the continued development of these technologies benefit from other work; this will clearly be the case for the last concept described, the extension of aligned pillar bonding to realize “optical solder bumps,” and it will be the case for EoE and SonG as well. The “debt” this work owes to the advances of other researchers in loosely related fields is often not fully acknowledged, or perhaps even recognized specifically, but its importance can not be forgotten.

REFERENCES

- [1] A. B. Glasser and G. E. Subak-Sharpe, *Integrated Circuit Engineering* (Addison-Wesley, Reading, MA, 1977), p. 5.
- [2] David A. B. Miller, “Dense Two-Dimensional Integration of Optoelectronics and Electronics for Interconnects,” in *Heterogeneous Integration: Systems on a Chip: Proceedings of a Conference held January 26–27, 1998 in San Jose, CA* edited by Anis Husain and Mahmoud Fallahi (Critical Reviews of Optical Engineering, Vol. CR70, SPIE Optical Engineering Press, Bellingham, WA, 1998) pp. 85.
- [3] The trend for DRAM density to increase four times every three to four years was recognized by Gordon Moore in the early 1970’s and this trend, which has continued every since, has become to be known as Moore’s Law [R. T. Howe and C. G. Sodini, *Microelectronics—an integrated approach* (Prentice-Hall, Upper Saddle River, NJ, 1997) p. 840.
- [4] For references to the early literature see, for example, R. A. Smith, *Semiconductors* (Cambridge University Press, London, 1959) p. 471.
- [5] P. C. Chen, K. L. Yu, S. Margalit, and A. Yariv, “A new GaInAsP/InP T-laser at 1.2 μm fabricated on semi-insulating substrate,” *Jap. J. Appl. Phys.* **19** (1980) L775–L776.

- [6] R. Davis, N. Forbes, A. C. Carter, and R. C. Goodfellow, "An optoelectronic integrated circuit in the GaAlAs/GaAs system," Inst. Phys. Conf. Ser. No. 63: Chapter 12, Paper presented at the Int. Symp. on GaAs and Related Compounds, Japan, 1981.
- [7] A. V. Krishnamoorthy, L. M. F. Chirovsky, W. S. Hobson, R. E. Leibenguth, S. P. Hui, G. J. Zydzik, K. W. Goosen, J. D. Wynn, B. J. Tseng, J. Lopata, J. A. Walker, J. E. Cunningham, and L. A. D'Asaro, "Vertical-cavity surface-emitting lasers flip-chip bonded to gigabits-per-second CMOS circuits," IEEE Photonics Technology Letters *11* (1999) 128–130; see also the other papers in the present volume.
- [8] *Heterogeneous Integration: Systems on a Chip: Proceedings of a Conference held January 26–27, 1998 in San Jose, CA* edited by Anis Husain and Mahmoud Fallahi (Critical Reviews of Optical Engineering, Vol. CR70, SPIE Optical Engineering Press, Bellingham, WA, 1998).
- [9] H. S. Hinton, "Progress in smart pixel technologies," IEEE Selected Topics in Quantum Electronics *2* (1996) 14–23.
- [10] M. Shur, *Physics of Semiconductor Devices* (Prentice-Hall, Englewood Cliffs, NJ, 1990) pp. 623–635.
- [11] P. Kopperschmidt, G. Kastner, D. Hesse, N. D. Zakharov, and U. Gosele, "High bond energy and thermomechanical stress in silicon on sapphire wafer bonding," Appl. Phys. Lett. *70* (1997) 2972–2974.
- [12] J. M. London, "Wafer bonding for monolithic integration of Si CMOS VLSI electronics with III-V optoelectronic devices," S. M. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, June 1999.
- [13] H. C. Casey, Jr. and M. B. Panish, *Heterostructure Lasers, Part B: Materials and Operating Characteristics* (Academic Press, New York, 1978), p. 300.
- [14] P. Petroff and R. L. Hartman, "Defect structure introduced during operation of heterojunction GaAs lasers," Appl. Phys. Lett. *23* (1973) 469–471.
- [15] M. Fukada, *Reliability and Degradation of Semiconductor Lasers and LEDs* (Artech House, Norwood, MA, 1991) p. 185.
- [16] L. N. Nguyen, A. S. Brown, M. A. Thompson, and L. M. Jelloian, "50-nm self-aligned-gate pseudomorphic AlInAs/GaInAs high electron mobility transistors," IEEE Trans. Electron Dev. *39* (1992) 205–212. This article has numerous references to the early work on pseudomorphic HEMTs.
- [17] J. L. Coleman, "Strained layer quantum well heterostructure lasers," in *Quantum Well Lasers* ed. by P. S. Zory, Jr. (Academic Press, San Diego, 1993) pp. 367–414.
- [18] B.-Y. Tsaur and G. M. Metzger, Appl. Phys. Lett. *45* (1984) 535–537; W. I. Wang, Appl. Phys. Lett. *44* (1984) 1149–1151; J. Palmer, G. Burns, C. Fonstad, and C. Thompson, "Effect of As₄ overpressure on initial growth of gallium

- arsenide on silicon by molecular beam epitaxy," *Appl. Phys. Lett.* 55 (1989) 990–992.
- [19] T. H. Windhorn and G. M. Metze, "Room-temperature operation of GaAs/AlGaAs diode lasers fabricated on a monolithic GaAs/Si substrate," *Appl. Phys. Lett.* 47 (1985) 1031; H. Z. Chen, A. Ghaffari, H. Wang, H. Morkoc, and A. Yariv, *Optics Letters* 12 (1987) 812; T. C. Chong and C. G. Fonstad, "Low-threshold operation of AlGaAs/GaAs multiple quantum well lasers grown on Si substrates by molecular beam epitaxy," *Appl. Phys. Lett.* 51 (1987) 221–223.
 - [20] S. F. Fang, K. Adomi, S. Iyer, H. Morkoc, H. Zabel, C. Choi, and N. Otsuka, "Gallium arsenide and other compound semiconductors on silicon," *J. Appl. Phys.* 68 (1990) R31–58.
 - [21] J. E. Palmer, G. Burns, C. G. Fonstad, and C. V. Thompson, "Effect of As₄ overpressure on initial growth of gallium arsenide on silicon by molecular beam epitaxy," *Appl. Phys. Lett.* 55 (1989) 990–992.
 - [22] S. Sakai, N. Wada, and C. L. Shao, "Thermal annealing effects on the defect and stress reduction in undercut GaAs on Si," Paper B20 at the 18th International Symposium on Gallium Arsenide and Related Compounds, Seattle, WA, Sept. 9–12, 1991 (unpublished).
 - [23] J. De Boeck, C. Van Hoof, K. Deneffe, and G. Borghs, "As grown stress free GaAs on Si by regrowth on chemically released GaAs on Si mesas," *Jpn. J. Appl. Phys.* 30 (1991) L423; J. De Boeck, G. Zou, M. van Rossum, and G. Borghs, "Mesa release and deposition used for GaAs-on-Si MESFET fabrication," *Electron. Lett.* 27 (1991) 22.
 - [24] G. F. Burns and Clifton G. Fonstad, "Complete strain relief of heteroepitaxial GaAs on silicon," *Appl. Phys. Lett.* 61 (1992) 2199–2201; G. F. Burns, H. Blanck, and C. G. Fonstad, Jr., "Complete strain relief of GaAlAs heterostructures and devices grown and processed on Si-substrates," Paper B19 at the 18th International Symposium on Gallium Arsenide and Related Compounds, Seattle, WA, Sept. 9–12, 1991 (unpublished).
 - [25] G. F. Burns, "Monolithic fabrication of strain-free (Al,Ga)As semiconductor lasers on silicon substrates using molecular beam epitaxy and epitaxial separation," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, Feb. 1992 (unpublished).
 - [26] G. F. Burns and C. G. Fonstad, "Monolithic Fabrication of Strain-Free (Al,Ga)As Heterostructure Lasers on Silicon Substrates," *IEEE Photon. Tech. Lett.* 4 (1992) 18–20.
 - [27] MOSIS is operated by the University of Southern California Information Services Institute, 4676 Admiralty Way, Marina del Rey, CA 90292-6695.
 - [28] Vitesse Semiconductor Corporation, 741 Calle Plano, Camarillo, CA 93010.

- [29] K. V. Shenoy, C. G. Fonstad, Jr., B. Elman, F. D. Crawford, and J. Mikkelsen, "Laser diodes and refractory metal gate VLSI GaAs MESFETs for smart pixels," IEEE/LEOS Annual Meeting Digest, Boston, MA, Nov. 16–19, 1992.
- [30] K. V. Shenoy, "VLSI GaAs MESFETs and laser diodes for monolithic optoelectronic circuits," S.M. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, June 1992, unpublished.
- [31] A. C. Grot, D. Psaltis, K. V. Shenoy, and C. G. Fonstad, Jr., "Integration of LEDs and GaAs circuits by MBE regrowth," IEEE Photon. Technol. Lett. 6 (1994) 819–821; K. V. Shenoy, C. G. Fonstad, Jr., A. C. Grot, and D. Psaltis, "Monolithic optoelectronic circuit design and fabrication by epitaxial growth on commercial VLSI GaAs MESFETs," IEEE Photonics Tech. Lett. 7 (1995) 1–3.
- [32] J. E. Cunningham, K. W. Goosen, W. Y. Jan, J. A. Walker, and R. N. Pathak, "Low-temperature growth of 850 nm quantum well modulators for monolithic integration to very scale integrated electronics on GaAs," J. Vac. Sci. Technol. B, 13 (1995) 653–656.
- [33] J. F. Ahadian and Clifton G. Fonstad, Jr., "Epitaxy-on-Electronics technology for monolithic optoelectronic integration: foundations, development, and status," Opt. Eng. 37 (1998) 3161–3174.
- [34] J. F. Ahadian, P. T. Vaidyanathan, S. G. Patterson, Y. Royter, D. Mull, G. S. Petrich, W. D. Goodhue, S. Prasad, L. A. Kolodziejski, and C. G. Fonstad, Jr., "Practical OEIC's based on monolithic integration of GaAs-InGaP LED's with commercial GaAs VLSI electronics," IEEE J. Quantum Elect. 34 (1998) 1117–1123.
- [35] J. F. Ahadian, "Development of a monolithic very large scale optoelectronic integrated circuit technology," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, Feb. 2000, unpublished.
- [36] K V. Shenoy, "Monolithic optoelectronic VLSI circuit design and fabrication for optical interconnects," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, Sept. 1995, unpublished.
- [37] Y. Royter, "Monolithic integration of etched facet lasers with GaAs VLIS circuits," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, January 1998, unpublished.
- [38] P. P. Ruden, A. I. Akinwande, D. Narum, D. E. Grider, and J. Nohava, "High performance complementary logic based on GaAs/InGaAs/AlGaAs HIGFETs," *International Electron Devices Meeting 1989, Technical Digest* (IEEE, New York, NY, 1989, Cat. No.89CH2637-7), pp. 117–120.
- [39] M. Hibbs-Brenner, Honeywell, Inc., private communication.

- [40] E. K. Braun, K. V. Shenoy, C. G. Fonstad, Jr., and J. M. Mikkelsen, "Elevated Temperature Stability of GaAs Digital Integrated Circuits," *IEEE Electron Dev. Lett.* **17** (1996) 37–39.
- [41] *Recent results using the EPI atomic hydrogen source*, Application Note 3/98, EPI MBE Products Group, 4900 Constellation Dr., St. Paul, MN 55127, September 1998, unpublished.
- [42] J. Diaz, H. J. Yi, M. Razeghi, and G. T. Burnham, "Long-term reliability of Al-free InGaAsP/GaAs (808 nm) lasers at high-power high-temperature operation," *Appl. Phys. Lett.* **71** (1997) 3042–3044.
- [43] L. M. F. Chirovsky, W. S. Hobson, R. E. Leibenguth, S. P. Hui, J. Lopata, G. J. Zydzik, G. Giaretta, K. W. Goossen, J. D. Wynn, A. V. Krishnamoorthy, B. J. Tseng, J. M. Vandenberg, and L. A. D'Asaro, "Implant-Apertured and Index-Guided Vertical-Cavity Surface-Emitting Lasers (I2-VCSEL's)," *IEEE Photonics Technology Letters* **11** (1999) 500–502.
- [44] A. C. Grot, "Analog GaAs Optoelectronic Integrated Circuits for Large Scale Arrays," Ph.D. Thesis, Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, 1994, unpublished.
- [45] J. Mikkelsen, Vitesse Semiconductor, private communication.
- [46] P. T. Vaidyanathan, "Fabrication and characterization of GaAs/InGaP double heterostructure light emitting diodes for monolithic optoelectronic integrated circuits," S. M. Thesis, Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, Sept. 1997, unpublished.
- [47] J. P. Donnelly, W. D. Goodhue, C. A. Wang, R. J. Bailey, G. A. Lincoln, G. D. Johnson, L. J. Missaggia, and J. N. Walpole, "CW operation of monolithic arrays of surface-emitting AlGaAs diode lasers with dry-etched vertical facets and parabolic deflecting mirrors," *IEEE Photonics Technology Letters* **5** (1993) 1146–1148.
- [48] H. H. Hsieh and C. G. Fonstad, "Analysis of grating coupled emission in lead-tin telluride diode lasers," *IEEE J. Quant. Electron.* **QE-13** (1977) 681–686; F. S. Choa, M. H. Shih, J. Y. Fan, G. J. Simonis, P.-L. Liu, T. Tanbun-Ek, R. A. Logan, W. T. Tsang, and A. M. Sargent, "Very low threshold 1.55 μm grating coupled surface-emitting lasers for optical signal processing and interconnect," *Appl. Phys. Lett.* **67** (1995) 2777–2779; N. Eriksson, J. Bengtsson, M. Li, P. Modh, and A. Larsson, "Surface-emitting unstable-resonator lasers with integrated diffractive beam-forming elements," *IEEE Photonics Tech. Lett.* **9** (1997) 1570–1572.
- [49] K. D. Choquette and H. Q. Hou, "Vertical-cavity surface emitting lasers: moving from research to manufacturing," *Proc. IEEE* **85** (1997) 1730–1739.
- [50] K. V. Shenoy, "VLSI GaAs MESFETs and laser diodes for monolithic optoelectronic circuits," S. M. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA,

- 1992.
- [51] J. M. Rossler, Y. Royter, D. E. Mull, W. D. Goodhue, and C. G. Fonstad, Jr., "Bromine ion-beam-assisted etching of InP and GaAs," *J. Vac. Sci. Technol. B* 16 (1998) 1012–1017.
 - [52] W. Goodhue and S. Choi, MIT Lincoln Laboratory, private communication.
 - [53] R. J. Aggarwal, "Design of resonant-tunneling diodes for a GaAs integrated SRAM," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Feb. 1997, unpublished; R. J. Aggarwal and C. G. Fonstad, Jr., "High peak-to-valley current ratio InGaAs/AlAs RTDs on GaAs substrates using relaxed InGaAs buffers," *IEEE Electronics Letters*, 31 (1995) 75–77; R. J. Aggarwal, K. V. Shenoy, and C. G. Fonstad, Jr., "A technology for monolithic integration of high-indium-fraction resonant tunneling diodes with commercial MESFET VLSI electronics," *Proceedings of the Conference on InP and Related Compounds*, Sapporo, Japan, May 1995, pp. 85–88.
 - [54] H. Wang, J. Luo, K. V. Shenoy, Y. Royter, C. G. Fonstad, Jr., and D. Psaltis, "Monolithic integration of SEEDs and VLSI GaAs circuits by epitaxy on electronics," *IEEE Photonics Technology Letters* 9 (1997) 607–609.
 - [55] D. A. Hall, B. L. Shoop, J. R. Loy, G. B. Tait, E. K. Ressler, J. F. Ahadian, and C. G. Fonstad, Jr., "Experimental demonstration of a 3x3 monolithically integrated smart pixel neural array based on epitaxy-on-electronics technology," presented at the OSA Annual Meeting, October 12–27, Long Beach, CA, postdeadline paper (unpublished).
 - [56] C.-H. Chen, C. B. Kuznia, B. Hoanca, J.-M. Wu, and A. A. Sawchuck, "Design and demonstration of smart pixel array network interface for 2D parallel data packet networks," presented at the OSA Annual Meeting, October 12–27, Long Beach, CA, (unpublished).
 - [57] I. Y. Yang, C. Vierri, A. Chandrakasan, and D. A. Antoniadis, "Back-gated CMOS on SOIAs for dynamic threshold control," *IEEE Trans. Electron Dev.*, 44 (1997) 822–831.
 - [58] B. E. Burke, J. A. Gregory, R. W. Mountain, B. B. Kosici, E. D. Savoye, P. J. Daniels, V. S. Dolat, T. I. Lind, A. H. Loomis, D. J. Young, G. A. Luppino, and J. L. Tonry, "Large-area back-illuminated CCD imager development," *Experimental Astronomy* 8 (1998) 31–40.
 - [59] J. M. London, A. H. Loomis, J. F. Ahadian, and C. G. Fonstad, Jr., "Preparation of silicon-on-gallium arsenide wafers for monolithic optoelectronic integration," *IEEE Photonics Tech. Lett.* 11 (1999) 958–960.
 - [60] J. M. London, "Wafer bonding for monolithic integration of Si CMOS VLSI electronics with III-V optoelectronic devices," S.M. Thesis, Department of Electrical Engineering and Computer Science, MIT, June 1999, unpublished.
 - [61] Q-Y. Tong and R. W. Bower, "Beyond "Smart Cut": recent advances in layer

- transfer for material integration,” MRS Bulletin (December 1998) 40–44.
- [62] S. N. Farrens, J. R. Dekker, J. K. Smith, and B. E. Roberds, “Chemical free room temperature wafer to wafer direct bonding,” J. Electrochem. Soc. *142* (1995) 3949–3955.
 - [63] Silicon Genesis Corporation, 590 Division St., Campbell, CA 95008.
 - [64] J. M. London, P. A. Postigo, and C. G. Fonstad, Jr., “Quantum well heterostructures grown by molecular beam epitaxy on silicon-on-gallium arsenide substrates,” Appl. Phys. Lett. *75* (1999).
 - [65] T. J. Drabik, “A new bonding technique for microwave devices,” IEEE Trans. on Components, Packaging, and Manufacturing Technology—Part B *19* (1996) 57–63.
 - [66] R. Pu, C. Duan, and C. W. Wilmsen, “Hybrid integration of VCSELs to CMOS integrated circuits,” IEEE Journal of Selected Topics in Quantum Electronics *5* (1999) 201–208.
 - [67] Z. L. Liao and D. E. Mull, “Wafer fusion: a novel technique for optoelectronic device fabrication and monolithic integration,” Appl. Phys. Lett. *56* (1990) 737–739.
 - [68] I-H. Tan, C. Reaves, A. L. Holmes, Jr., J. E. Bowers, and S. DenBaars, “Low temperature Pb bonding of III-V semiconductors,” Electronics Lett. *31* (19) 588–589.
 - [69] H. Wang, “Monolithic integration of 1.55 micron photodetectors with GaAs electronics for high speed optical communications,” Ph.D. Thesis, Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, Sept. 1998, unpublished.
 - [70] Sadik C. Esener, Daniel Hartmann, Michael J. Heller, and Jeffery M. Cable, “DNA Assisted Micro-Assembly: A Heterogeneous Integration Technology for Optoelectronics,” in *Heterogeneous Integration: Systems on a Chip: Proceedings of a Conference held January 26–27, 1998 in San Jose, CA* edited by Anis Husain and Mahmoud Fallahi (Critical Reviews of Optical Engineering, Vol. CR70, SPIE Optical Engineering Press, Bellingham, WA, 1998) pp. 113–140.
 - [71] Joseph J. Talghader, Jay K. Tu, and J Stephen Smith, “Integration of Fluidically Self-Assembled Optoelectronic Devices Using a Silicon-Based Process,” IEEE Photonics Technology Letters *7* (1995) 1321–1323.
 - [72] Ashish K. Verma, Mark A. Hadley, Hsi-Jen J. Yeh, and J. S. Smith, “Fluidic Self-Assembly of Silicon Microstructures,” *Proceedings of the 45th. Electronic Components and Technology Conference* (IEEE Press, New York, NY, 1995) pp. 1263–1268.

Vertical-Cavity Surface-Emitting Lasers Specifically Designed for Integration with Electronic Circuits

Leo M. F. Chirovsky, Ashok V. Krishnamoorthy, William S. Hobson,
John Lopata, Keith W. Goossen, and L. Arthur D'Asaro
Bell Laboratories

2.1 MOTIVATION FOR THE WORK

Several of the authors of this chapter have worked, at various times, on developing what is called “optoelectronic-VLSI” (OE-VLSI) technology, the intimate integration of photonic devices with VLSI electronics. The general purpose is to give the electronics the option of communicating through optical interconnects. Most of that work, thus far, has focused on using the *multiple quantum well electro-absorption* (MQWEA) diode [1, 2] as the photonic device, primarily because the MQWEA diode can serve both as a detector and as a modulator, for data reception and for data transmission, respectively. In a monolithic integration approach [3], the MQWEA diodes were cofabricated with GaAs field-effect transistors (FETs) to produce optoelectronic circuits. This technology produced several system demonstrators and was the basis of a multiproject foundry shuttle of optoelectronic chips to research groups around the world. It was obvious from the start that this monolithic approach would have needed a huge development effort to become a VLSI technology, on the scale of the one that produced the silicon CMOS VLSI technology. Funds for such a large effort could not be justified. However, since the system demonstrators produced encouraging results, attention turned to a hybrid approach, whereby arrays of MQWEA diodes are flip-chip solder-bump bonded to CMOS chips [4, 5, 6], arguably the most advanced electronic VLSI technology. Soon thereafter, this hybrid technology produced many more system demonstrators and was the basis of several multiproject foundry shuttles of optoelectronic chips to research groups around the world [7]. A true OE-VLSI technology was emerging, just as strong arguments developed for the advantages of OE-VLSI technology in general; these were summarized in [8].

In the meantime, many other groups throughout the world were developing alternative approaches to OE-VLSI. Many of these are described in the other chapters of this book by the developers themselves, and most of the others are summa-

rized in reference [8]. In most of these alternative approaches, VCSELs are the photonic devices of choice for data output. Prudence and curiosity suggested that the OE-VLSI technology should include VCSELs in its repertoire of photonic devices, though by no means to the exclusion of MQWEA modulators.

Prudence also suggested that, from the start, the VCSELs should be designed to be amenable to integration with electronic circuits, specifically to our flip-chip solder-bump bonding technique [5]. First of all, that meant VCSELs whose contacts are as nearly coplanar as possible. Dual intracavity contacts fulfill that role best, since the vertical separation between the two contacts to each VCSEL can then be less than half a micrometer. Such a contacting approach is still not conventional, but its feasibility was already proven [9] when a variation of it was developed. With an intracavity contacting scheme comes the option of optically optimized mirrors with no doping or grading.

We chose that option and took it one step further, where the top mirror is a deposited dielectric Distributed Bragg Reflector (dd-DBR). The risk in using a dd-DBR is balanced by several advantages and at least one research group has been using a variation successfully for several years [10]. The alternating layers of a dd-DBR can have a much larger ratio of indices of refraction than any epitaxial DBR. As a result, the spectral bandwidth of a dd-DBR can be much larger than that of epitaxial mirrors, as is shown in Fig. 2.1, and the optical penetration depth into a dd-DBR is much shorter. What is more, a desired reflectivity is achieved by a dd-DBR with far fewer layer pairs. These three facts translate into the following benefits—much greater tolerance to variations in layer thicknesses, a shorter and so more efficient resonant cavity, and a lower profile top mirror that will interfere less with flip-chip bonding. The risk in using a dd-DBR was twofold—its long-term

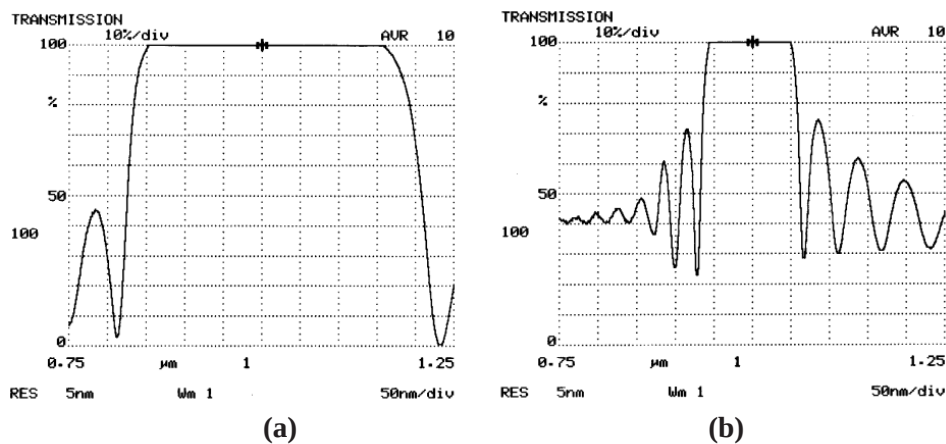


Figure 2.1. Plots of reflectivity vs. wavelength for a dd-DBR (a) and for a binary (AlAs/GaAs) epitaxial DBR (b). The number of pairs for each mirror was chosen to give a peak reflectivity of ~99.5%.

stability as part of a VCSEL had not been tested and there was no established way of patterning it in order to access the contacts.

A somewhat subtler and perhaps even more important advantage of using a dd-DBR, as the top mirror of a VCSEL, became apparent in due order. This mirror can be deposited and patterned as the last steps of the VCSEL-fabrication sequence. As a result, processing options for the other device features become available which are impossible with a thick epitaxial mirror in the way or which are greatly hindered by such a mirror. The paramount option is that in all of the processing steps used, the device features (including the current aperture) can be photolithographically defined. That is an important characteristic of VLSI technologies, since it offers the best chance for reproducible yield and performance. Another important option is that right on the wafer surface, and so ultimately under the top mirror, a separate feature can be formed that determines the diameter of the beam waist throughout the VCSEL structure. Thus, this feature guides the beam as efficiently as possible in the active region, in the current aperture, in the contact openings, and in both mirrors. An opportunity to explore just these options is sufficient motivation for a research project.

Nevertheless, in light of all the above considerations, we formulated the P²I²-VCSEL design strategy to achieve the broad objective of a fabrication platform which would reproducibly yield high-efficiency devices that could easily be integrated and operated with electronic circuitry. The design name is an acronym highlighting the primary design motifs—VCSELs processed with *photolithographic precision*, containing an *intracavity* contacted current guide and an *independent* optical beam guide. The use of a dd-DBR is implied in the above since without it the design strategy could not be realized. Figure 2.2 is a cross-sectional diagram of a generic P²I²-VCSEL, which is described in more detail in Section 2.3. The *intracavity* contacted current guide makes the device amenable to flip-chip bonding for intimate integration with electronic circuitry. Processing with *photolithographic precision* offers the promise of the best possible reproducibility in fabrication. The *independent* optical beam guide in combination with the well-defined current guide is included to produce high efficiency.

The first implementations of the P²I²-VCSEL design strategy were called I²-VCSELs [11]. The shorter acronym stood for *Implant-Apertured and Index-Guided VCSELs*. This shorter name emphasized that these VCSELs had an unusual combination of device features that vastly improved the performance of implant-apertured devices. However, the I²-VCSEL design was only possible because of the broader P²I²-VCSEL design strategy, and so demonstrates its efficacy. Many other implementations in line with that strategy are possible. In this paper we wish to draw more attention to the overall strategy and its potential, so that readers recognize that many other specific implementations are possible. The next section of the chapter gives some background information that contributed to

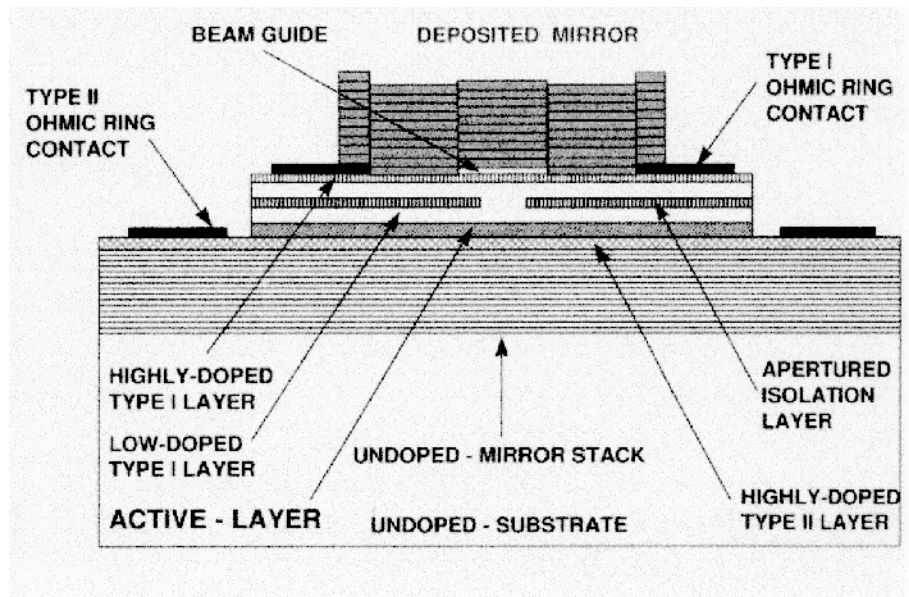


Figure 2.2. A schematic diagram of a cross section of a generic P^2I^2 -VCSEL.

the reasoning that led to the choice of the I^2 -VCSEL design as the first manifestation of the P^2I^2 -VCSEL design strategy.

2.2 BACKGROUND INFORMATION

One of the principal advantages of vertical-cavity surface-emitting lasers is their high-power conversion efficiency at low power levels, in the 1-mW regime. Previously, only oxide-confined current aperture VCSELs [12, 13], using the oxidation of high-Al-concentration AlGaAs, have achieved threshold currents far below 1 mA, while still maintaining high quantum efficiency. This is so because oxide apertures have precise boundaries near the active region and perform beam guiding through an index-guide effect that confines optical beam waists to sizes smaller than all other device features, which minimizes optical loss. The cross section of a generic oxide-confined current aperture VCSEL is illustrated in Fig. 2.3. A problem with those apertures, however, is that it is difficult to control their sizes. The best control of sizes is obtained with photolithographically defined processing steps. Oxide aperture fabrication typically begins with a photolithographic step, but the subsequent, deep etch (through a thick epitaxial top mirror) and the lateral oxidation procedure can easily introduce significant variations in final lateral dimensions.

There are many possible sources of variation [14]. The errors in the diameter of the aperture prior to oxidation are carried unchanged to the significantly smaller

oxide aperture, where they constitute a much larger percentage change in dimensions. The oxidation rate depends rapidly on the Al concentration in the AlGaAs and on the thickness of the AlGaAs layer, which are likely to vary across the wafer and from wafer to wafer. The oxidation rate depends on the composition of the surrounding layers, which may also vary. The rate depends on furnace temperature, the water vapor content of the ambient gas, and the gas flow dynamics in the oxidation furnace. Lateral oxidation typically proceeds from the boundary of a mesa, which is 10–15 μm away from the aperture center. It must be controlled to within a few percent, in order to control the areas of small apertures (3–6 μm diameters) to within $\pm 10\%$, which will result in a threshold uniformity of $\pm 10\%$. Obtaining even that much control may be extremely difficult if not impossible. Whatever the case, the variations in the lateral dimensions of an oxide aperture will certainly be worse than for a strictly photolithographically defined feature, until an “oxidation stop” is forthcoming.

In principle, these problems can be avoided with photolithographically defined ion-implanted apertures [15]. However, the typical implanted aperture VCSELs, made prior to this work, have had even more daunting problems. A schematic cross section of one is shown in Fig. 2.4. The difficulties of previous implantation schemes stem principally from the fact that the implantation must penetrate deeply through the top epitaxial mirror, losing their precision and ability to scale to ever-smaller dimensions. That is because the resulting boundaries of

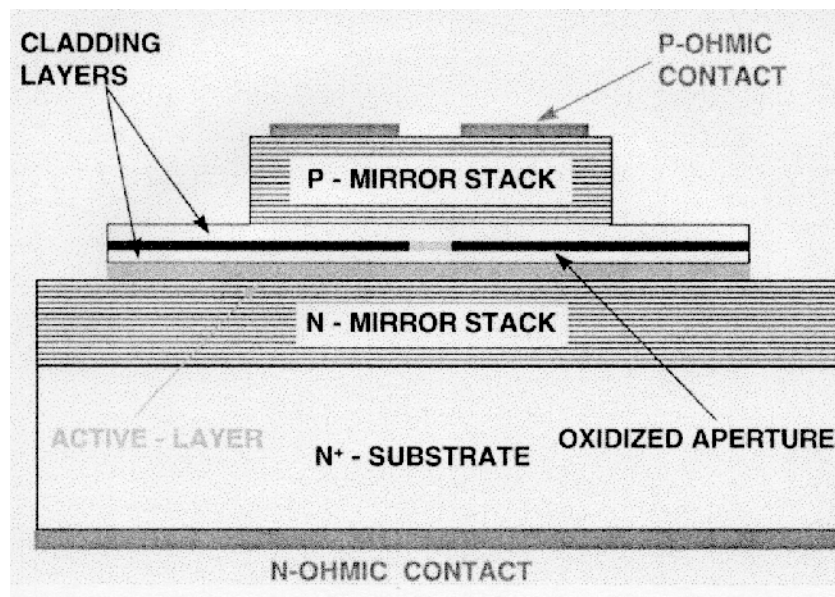


Figure 2.3. A schematic diagram of a cross section of a generic oxide-confined VCSEL.

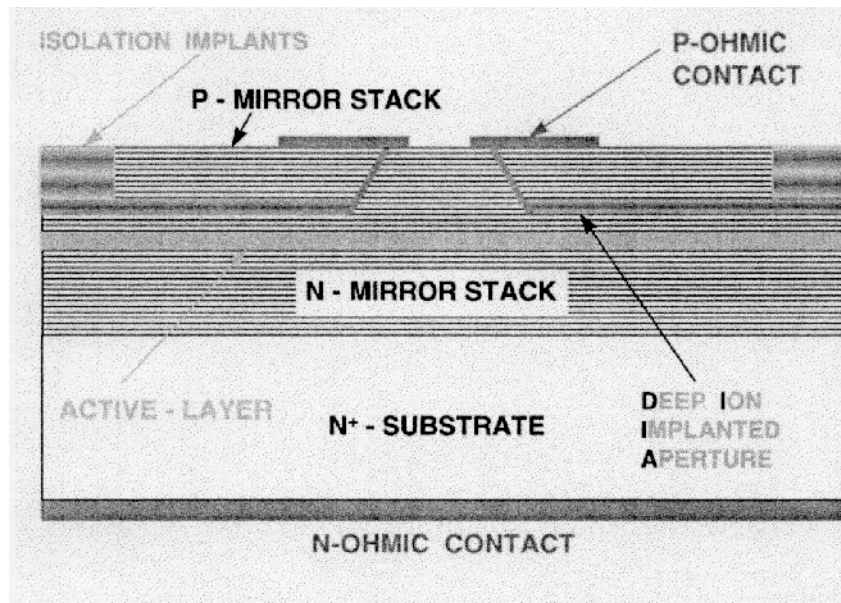


Figure 2.4. A schematic diagram of a cross section of a typical implant-apertured VCSEL, made prior to this work.

the implanted isolation layer become fuzzy due to straggle, and rise to the surface of the wafer in an inwardly sloping funnel shape. As a result, the apertures become narrower near the surface than near the active region far below the surface. In order to overcome high series resistance, the opening in the top ohmic metal contact must be several micrometers smaller than the current aperture near the active region. This necessity and the lack of beam confinement by any index guiding (other than gain guiding) cause more carrier and optical losses in those implant-apertured VCSELs than in oxide-apertured devices. The losses become proportionately greater as apertures are made smaller.

The P^2I^2 -VCSEL design approach provides an alternate way to obtain the advantages exhibited by oxide apertures, without sacrificing any of the dimensional accuracy obtainable with photolithography.

2.3 DETAILED DESCRIPTION OF AN I^2 -VCSEL, THE FIRST IMPLEMENTATION OF THE P^2I^2 -VCSEL DESIGN STRATEGY

In the two approaches discussed in the previous section, difficulties arise for the same reason. In both cases, critical device features (the guides) are formed deep below the surface, and as a result, these features suffer a loss of precision and/or accuracy in their lateral dimensions, leading directly to the aforementioned

problems. (Several micrometers below the surface is deep, especially when the intended lateral dimensions are on the order of 10 micrometers or less.) In the P^2I^2 -VCSEL design approach, the critical features are formed at or near the surface so that photolithographic precision can be accurately transferred to them and then are subsequently “buried” under a deposited top mirror. Similar mirrors have been effectively used before [10, 12], but those were not patterned. As is shown in Fig. 2.2, the generic P^2I^2 -VCSEL comprises an undoped substrate, an undoped (optically optimized) epitaxial bottom mirror, the laser cavity, which contains the entire intracavity contacted p-i-n structure, and a deposited and patterned dielectric top mirror. The cavity region is a few half-wavelengths thick, i.e., no more than a micrometer thick. Five major layers, arranged from the bottom mirror to the surface, make up the cavity region—a layer of one conductivity type (p or n), an intrinsic active layer, a low doped layer of the other conductivity type, a highly doped layer of that other conductivity type, and a cover layer that can be undoped. The beam guide is formed in the cover layer, right at the surface. The current guide is formed just below the surface, by contacting the highly doped layers and by creating an apertured isolation region in the low doped layer just above the active layer. The highly doped layers serve as dual lateral current injectors, sending their respective charge carriers towards the aperture, which then forces the carriers to flow vertically and to recombine in an area of the active layer, approximately equal to the area of the aperture, which is photolithographically defined. Only after all of the above parts are made is a top dielectric mirror, a dd-DBR, deposited on the surface and patterned to expose the contacts. Care is taken to make sure that the lateral dimensions of the top mirror are large enough to avoid diffractive losses. This description of the generic P^2I^2 -VCSEL was intentionally vague about how the beam guide and the apertured isolation layer are to be made, because the design strategy can allow different options.

For the first implementation, the I^2 -VCSEL, the choices were the simplest ones available that satisfied the criteria of the design strategy. The beam guide is a shallow mesa, a small fraction of a wavelength thick and concentric with the current aperture. This mesa causes an abrupt change in the thickness of the cavity at the mesa edge. Light can resonate in such a cavity only if the beam waist becomes small enough for the tails of the Gaussian beam to satisfy the radial boundary conditions for resonance. The situation is analogous to having a small index change, at the same radial position as the edge of the mesa, along the entire length of the cavity, with no mesa. We therefore call this mesa an index guide. Although its effects may not be completely equivalent to that of a typical waveguide, they certainly can dominate over the gain-guiding effects of the current aperture, just as well as the index-guiding effects of an oxidized aperture do. The apertured isolation layer of the I^2 -VCSEL is made with a shallow ion-implantation of a heavy ion. After a proper anneal [16], this implant forms precise current apertures in the

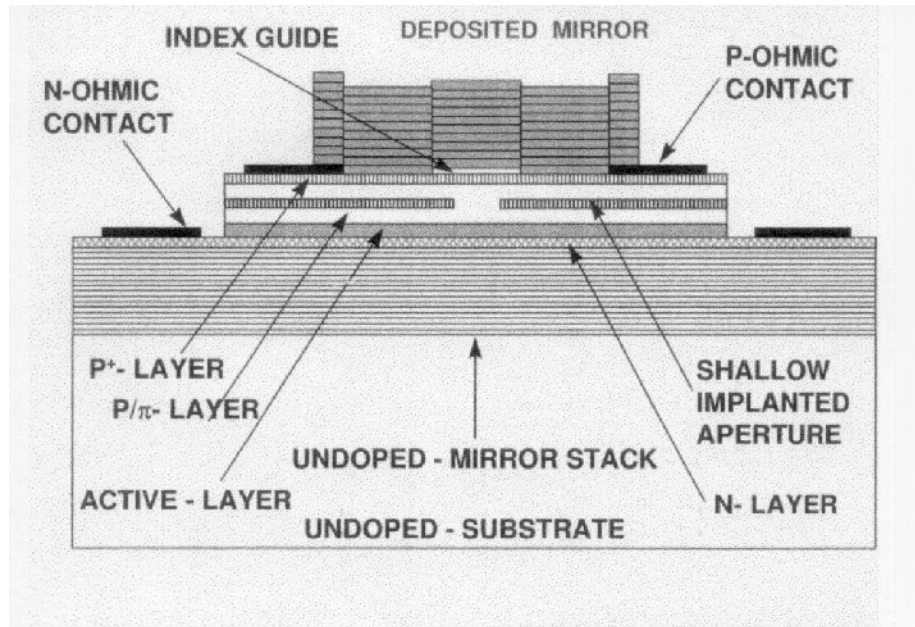


Figure 2.5. A schematic diagram of a cross section of an I²-VCSEL amenable to flip-chip bonding. The bottom mirror is an AlAs/GaAs quarter wave stack, the n-layer is ~1000 Å thick and Si-doped to 1E18 cm⁻³, the active layer comprises 3 quantum wells and suitable barriers for 980 nm, the p-layer is ~3000 Å thick and Be-doped from 2E17 cm⁻³ to 1E18 cm⁻³, the p⁺ layer is 300 Å thick and C-doped to 1E20 cm⁻³, the index guiding mesa is a GaAs layer (a fraction of a wavelength thick) atop a thin undoped InGaP stop-etch layer, the contacts are metals commonly used for alloyed contacts, and the deposited mirror is a dielectric quarter-wave stack.

low p-doped layer under the thin highly conducting lateral injection layer, and avoids the problems encountered with deep implants. *This is the best presently available method for producing a precise photolithographically defined current aperture.* For the technique to work well, it is important to have a doping ratio greater than two orders of magnitude between the highly doped and the low-doped layers above the active layer in Fig. 2.2. Much higher p-doping levels than n-doping levels are possible in GaAs. For these reasons, in an I²-VCSEL, the top layers are p-doped and the current injection layer beneath the active is n-doped. Most importantly, the p⁺-doped layer is kept thin and is positioned at a node of the standing wave pattern of the cavity, to prevent optical losses due to free-carrier absorption. Once these conditions are met, the aperture waist can be made smaller than the opening in the top contact. None of the problems that were suffered previously by implant-apertured VCSELs, described in Section 2.2, persist in I²-VCSELs, and performance levels similar to those of oxide-apertured devices are possible. The two photolithographically defined features of index guide and implanted aperture

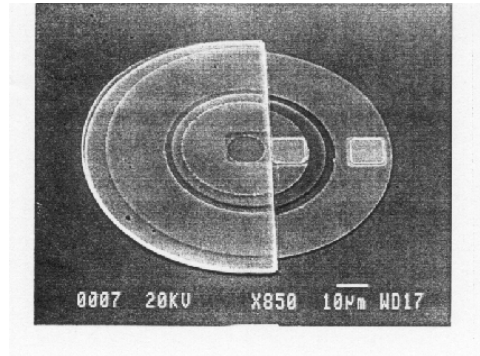


Figure 2.6. An SEM photograph of a completed I²-VCSEL, ready for flip-chip bonding.

duplicate the effects of a laterally oxidized aperture, but in contrast to oxidized apertures, their effects can be independently optimized. A schematic diagram of the cross section of an I²-VCSEL is shown in Fig. 2.5. By comparing Figs. 2.2 and 2.5, one can see how the I²-VCSEL follows the broad design strategy of the generic P²I²-VCSEL. For the purpose of flip-chip solder-bump bonding of I²-VCSEL arrays to CMOS circuits [8] we also often formed gold posts (just slightly taller than the top mirror) on the ohmic contacts. Figure 2.6 is an actual SEM photograph of an I²-VCSEL ready for flip-chip bonding.

This is a good point at which to mention some of the unresolved potential risks in pursuing the P²I²-VCSEL design strategy. We have learned to pattern dd-DBRs, but their long-term stability as parts of VCSELs is untested. In fact, the semiconductor–dd-DBR interface is also a risk factor. The shallow ion implantation technique has not been used with VCSELs before, and even the long-term stability of current guides with lateral current injection is an unknown. We mention these cautions because much more work is needed on this design approach, despite the very encouraging results that will be reported in the subsequent parts of this paper.

2.4 DATA ON THE FIRST, TOP-EMITTING I²-VCSELS

From the start, I²-VCSELs were fabricated amenable to flip-chip bonding; however, the very first ones were top-emitters and so were tested only at wafer level for basic performance properties. They were made to answer the following questions: (1) Will the current guide give a low series resistance, even when the top contact opening is several micrometers larger in diameter than the current aperture? (2) Can devices actually lase with patterned dd-DBRs? (3) Is the performance of I²-VCSELs comparable to that of oxide-confined VCSELs with similarly sized apertures, as predicted? (4) Does the beam guide work as expected? The answers to all four questions were affirmative, as the data in next four figures demonstrates.

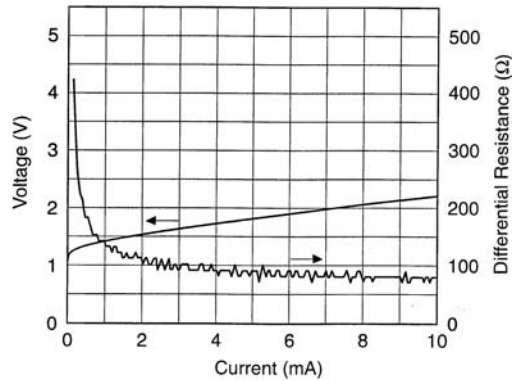


Figure 2.7. Electrical characteristics of a typical 10- μm -aperture I^2 -VCSEL, consisting of plots of the voltage versus the current and the differential series resistance versus the current.

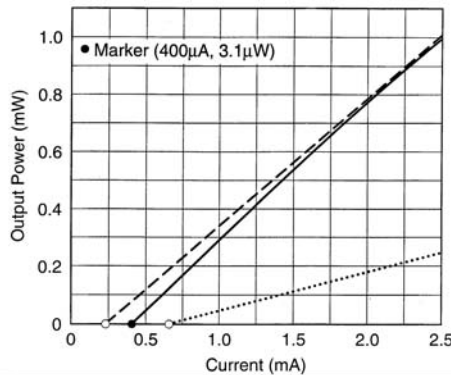


Figure 2.8. Plots of the L-I curves for a 6- μm -aperture I^2 -VCSEL (solid line), the best reported implanted 6- μm -aperture VCSEL prior to this work (dotted line), and a typical oxide-confined 6- μm -aperture VCSEL made by a potential manufacturer (dashed line).

Figure 2.7 presents typical voltage and differential series resistance versus current plots of devices with a 10- μm -diameter aperture, proving the excellent properties of the lateral current injection layer. Figure 2.8 shows that the L-I curve of a 6- μm -aperture I^2 -VCSEL (solid line) is vastly superior to the L-I curve for the previously best similarly-sized implanted VCSEL [15] (dotted line), and rivals the L-I characteristics of a typical similarly-sized oxide-confined VCSEL (dashed line—data quoted by a potential manufacturer of oxide-confined VCSELs). Figure 2.9 shows the richly multiple transverse mode output of a 10- μm device, indicative of a strong index-guided beam confinement which helps to improve L-I characteristics. The one surprise was that despite the strong index-guiding, devices with 6- μm -diameter apertures already had a single mode output, with a side-mode suppression ratio of over 30 dB, even when operating at currents ten times the threshold current, far beyond the expected operating current level. This is illustrated in Fig. 2.10.

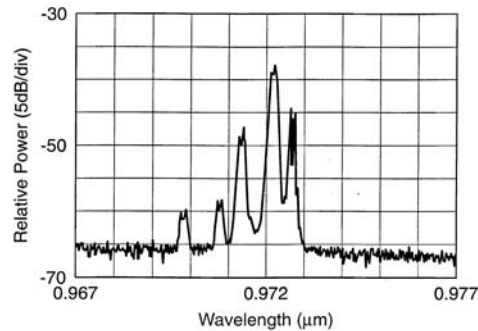


Figure 2.9. The multiple-transverse-mode optical spectrum, at 0.1 nm resolution, of the output of a 10- μ m-aperture I²-VCSEL at 5 mA input, indicating strong index-guided beam confinement.

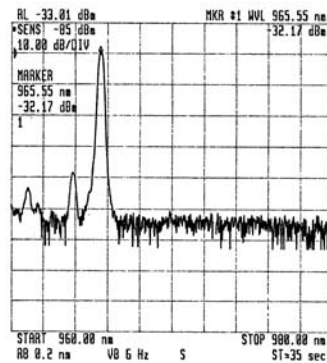


Figure 2.10. Single mode spectrum of a 6- μ m-aperture I²-VCSEL.

2.5 DATA ON BOTTOM-EMITTING AND ON HYBRID INTEGRATED I²-VCSELS

Since the top-emitting devices were already fabricated as if they were to be flip-chip bonded, the fabrication sequence for the bottom-emitters was not much different than for the top-emitters. The one notable change in the epigrowth was a reduction in the number of layer pairs in the bottom reflector stack so that it could be the output mirror. On the other hand, the top dd-DBR mirror became a back mirror by receiving two additional pairs or a metal coating with a phase-matching layer of the low index fluoride. Besides the two above differences between top and bottom emitters in their mirror designs, there was another. Since all our wafers were rough and nearly opaque on the substrate side, that substrate surface was polished on all bottom-emitting samples. That created a second, although weak, resonant cavity between every device's bottom mirror and the now (~30%) reflecting substrate surface. To avoid coupled-cavity effects, that surface was also anti-

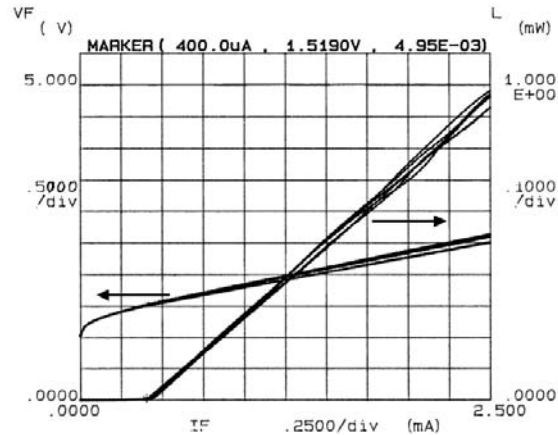


Figure 2.11. The L-I and the V-I curves of six randomly chosen ($6\text{-}\mu\text{m}$ -diameter aperture) bottom-emitting $\text{I}^2\text{-VCSELs}$ on a sample $\sim 3\text{ cm}^2$ in area. They show that good uniformity was obtained over the entire sample with performance comparable to that of top-emitting $\text{I}^2\text{-VCSELs}$.

reflection (AR) coated. However, to prevent damage to the AR coatings during handling, samples intended for flip-chip bonding were not actually AR-coated until after the integration procedure was completed. Therefore, any devices tested prior to flip-chip bonding showed coupled-cavity effects in the form of kinks or undulations in their L-I curves. Many samples needed thinning in addition to polishing before they could be diced into small chips, with one array each, for the flip-chip bonding. The L-I curves of devices on such samples, until they were AR-coated, had much more pronounced kinks and undulations because their second cavity was shorter. Very mild coupled-cavity effects appeared for devices with imperfect AR coatings. These are the reasons that the data in several of the following figures show kinks and undulations to various degrees.

We polished and then immediately AR-coated some samples to characterize bottom-emitting $\text{I}^2\text{-VCSELs}$, as fabricated, in order to compare their performance to that of the top-emitters. Figure 2.11 shows very important results from a sample, with many arrays of bottom-emitting devices, where good uniformity over its $\sim 3\text{ cm}^2$ area was obtained with characteristics similar to those of the top-emitters. On another sample, the metallization, but not the phase-matching layer, was removed from the top mirrors of some of the VCSELs. For those devices, both mirrors had the same output-level reflectivities and the devices lased bidirectionally. Figure 2.12 has the L-I and the V-I curves of a bidirectionally emitting $\text{I}^2\text{-VCSEL}$ and an adjacent bottom-only emitting one (top mirror still metallized). Despite an $\sim 1\%$ output per round trip pass by the light, the bidirectional device has a relatively modest increase in threshold and maintains high slope efficiency, thereby testifying to the high quality of the device material in general. These results were certainly sufficient to attempt the integration of arrays of $\text{I}^2\text{-VCSELs}$ to silicon chips.

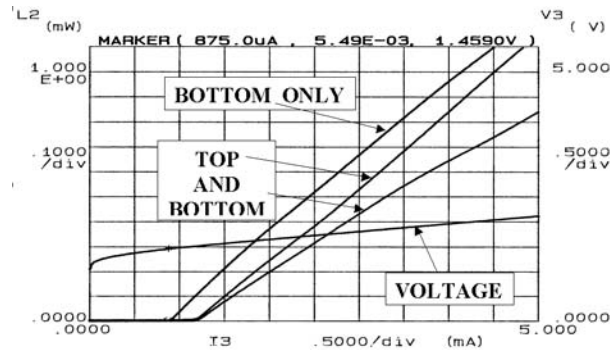


Figure 2.12. The L-I and the V-I curves of a bottom-only and an adjacent bidirectionally emitting I^2 -VCSEL (both 10- μ m-diameter aperture).

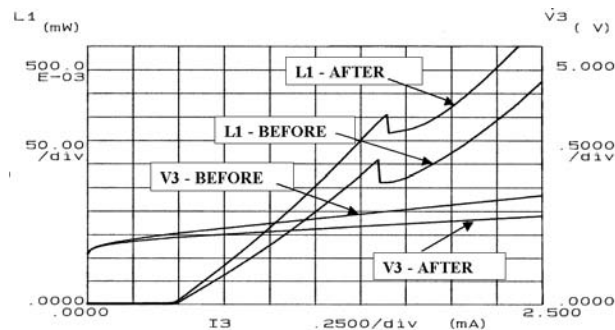


Figure 2.13. The characteristics of an I^2 -VCSEL, obtained just before and then just after flip-chip bonding.

The first flip-chip bonding experiments were with (2×10) I^2 -VCSEL arrays and bare-die silicon chips (solder bumps patterned on the silicon chips). In the beginning, a few VCSELs per array were probed and characterized just prior to the flip-chip bonding, and then were retested just afterwards, through the appropriate contact pads on the silicon chips. Figure 2.13 contains the characteristics, taken before and after flip-chip bonding, of one such typical I^2 -VCSEL. Both the L-I and the V-I responses actually improved as a result of the bonding. These improvements can be attributed to better electrical and thermal conductivity through the solder-bump contacts. Finally, Fig. 2.14 shows that, once AR-coated, a flip-chip bonded, bottom-emitting I^2 -VCSEL could indeed have a smooth, kink-free and undulation-free L-I curve. With such results in hand, we proceeded to the integration of I^2 -VCSEL arrays onto silicon CMOS chips.

We were at the point of testing the central premise of the P^2I^2 -VCSEL design strategy—that it would yield devices easily integrated and operated with electronic circuitry. For our experiments [17], we continued to use (2×10) I^2 -VCSEL arrays,

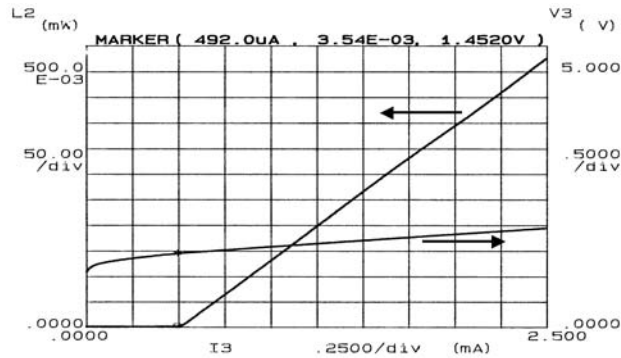


Figure 2.14. The L-I and the V-I curves of a flip-chip bonded I²-VCSEL, with a proper AR coating.

operating at a wavelength of ~ 970 nm. This enabled light transmission through the GaAs substrate, eliminating the absolute necessity for substrate removal, at least for the time being. A schematic cross section of the assembled structure is depicted in Fig. 2.15. The I²-VCSELs are bonded directly to the top level metal on the CMOS chip and do not interfere with the operation or the layout of the underlying circuits [18]. A microphotograph of part of a 2×10 bottom-emitting array prior to bonding is seen in Fig. 2.16. A photograph of the assembled structure is seen in Fig. 2.17. The CMOS drive circuits allowed each laser to be independently modulated and tested by external control. Figure 2.18 is a close-up photograph of the GaAs substrate face of a CMOS/VCSEL chip, with the entire array forward-biased, subthreshold, and with a single device biased over threshold and lasing. All the VCSELs were operational on some chips, but typically, one to two devices per chip did not lase due to observable processing errors. The operational VCSELs continued to lase through several weeks

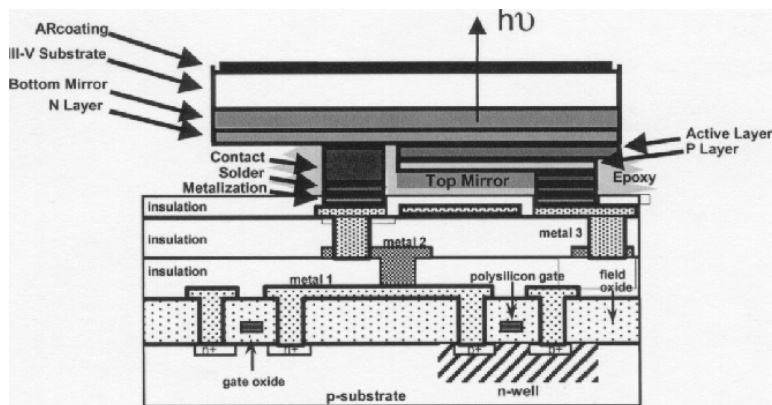


Figure 2.15. Schematic of a flip-chip I²-VCSEL bonded to CMOS. The I²-VCSEL is bonded over active gates.

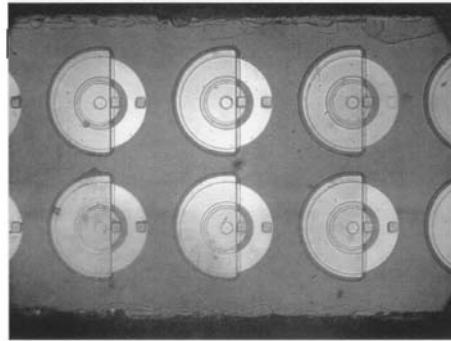


Figure 2.16. Microphotograph of a portion of a 2×10 I²-VCSEL array prior to flip-chip bonding.

of testing, making infant mortality an unlikely issue. Most of the integrated I²-VCSEL arrays comprised 10- μ m-diameter devices, whose threshold voltages and currents, after bonding, remained at ~ 1.5 V and ~ 1 mA, respectively. Both of these values are well within the capabilities of commercial CMOS technologies. The optical outputs of these devices, with 10 μ m diameter apertures, were multimode, similar to those of the top-emitters, described in Section 2.4 and depicted in Fig. 2.9. At room temperature, their outputs typically exceeded 4 mW before rollover occurred at input currents between 16 mA and 18 mA. This data confirmed that good electrical and thermal contact to the lasers was made via the flip-chip bonds.

To investigate the temperature effects on a CMOS/VCSEL chip and its reliability, one such chip was placed directly on a TE-cooler, and L-I curves were recorded for various temperatures. These appear in Fig. 2.19. Note that the 22°C L-I curve corroborates the above general statement about room temperature behavior. Note also that at 80°C, the output is still ~ 1 mW. No appreciable change in performance developed after temperature cycling. We also measured the wave-

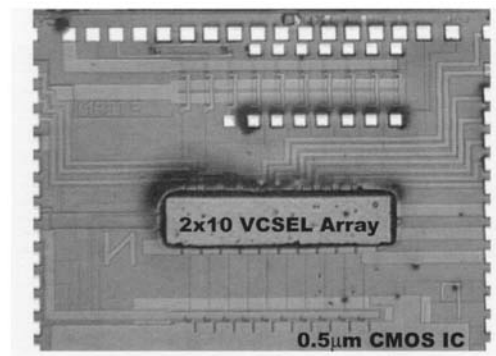


Figure 2.17. I²-VCSEL array after bonding to CMOS chip.

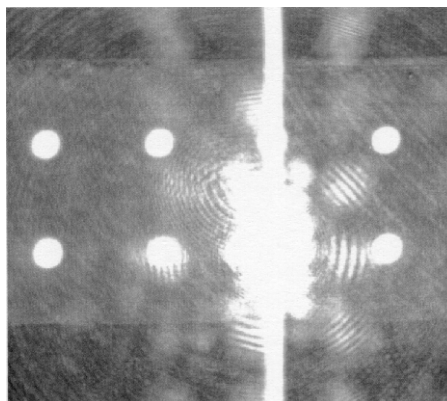


Figure 2.18. A microphotograph, of a portion of the top face of a CMOS/VCSEL chip, taken with one I²-VCSEL biased above threshold and lasing and the rest of the laser array forward-biased below threshold and just luminescing.

length shift in the output of an I²-VCSEL, flip-chip bonded to a CMOS chip, versus temperature and versus electrical input power. The resulting slopes for the linear fits to the two sets of data were 15°C/nm and 0.07 nm/mW. The product of these two values, ~1°C/mW, is the thermal resistance of a hybrid integrated I²-VCSEL.

The VCSEL driver on the CMOS/VCSEL chips was a simple two-transistor version, based on a current-shunting principle, which provides a low-area, tunable-power circuit. A schematic illustration appears in Fig. 2.20. The measured small-signal bandwidth of the driver was on the order of 2 GHz in 0.5-μm CMOS. A PMOS transistor sends an adjustable amount of current through the VCSEL and an NMOS transistor can then shunt the current into and out of the laser. Our first

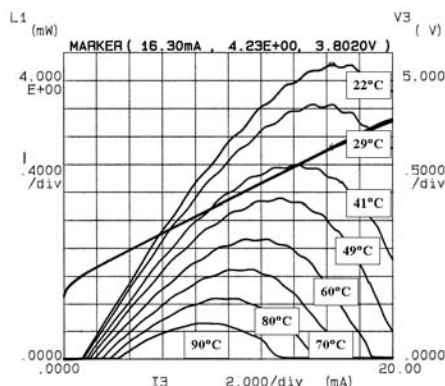


Figure 2.19. A set of L-I and V-I curves showing the temperature dependence of a flip-chip bonded CMOS/VCSEL chip.

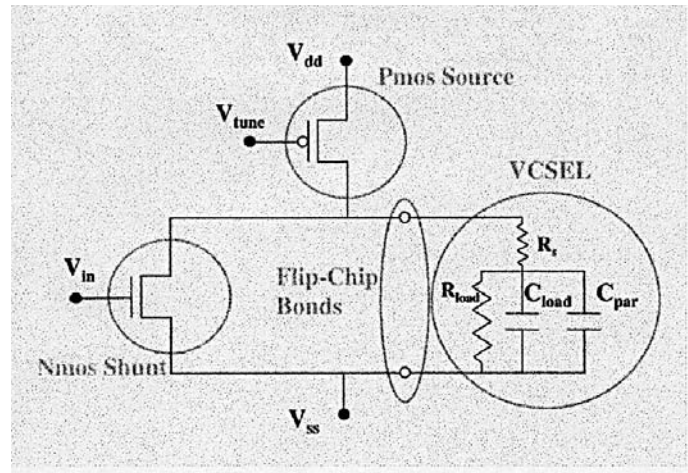


Figure 2.20. A circuit diagram of the CMOS/VCSEL transmitter. The equivalent circuit for the VCSEL includes a series resistance (R_s) and a parallel parasitic capacitor (C_{par}); both conspire to limit the modulation speed of the VCSEL.

drivers were able to provide their lasers up to 6 mA of current. For bit-rate measurements, a commercial 2-GHz detector/receiver monitored the output of each individual VCSEL output. The coupling between a laser and a receiver was through a five-meter-long, 62.5- μm -core multimode fiber patchcord. Figure 2.21 shows an eye diagram, taken at 1.25 Gbit/s, for a CMOS/VCSEL transmitter driven by $(2^{23} - 1)$ -bit-long, pseudorandom bitstream (PRBS). The bit-error rate was $< 10^{-11}$ when the I²-VCSEL was biased above threshold. Without the prebias, adjusted by a change in V_{tune} (Fig. 2.20), the transmitter would still operate at 1.25 Gbit/s, but the eye pattern would then indicate a turn-on delay of ~ 180 ps. The total power dissipated per optical output by the laser, the driver, and all the on-

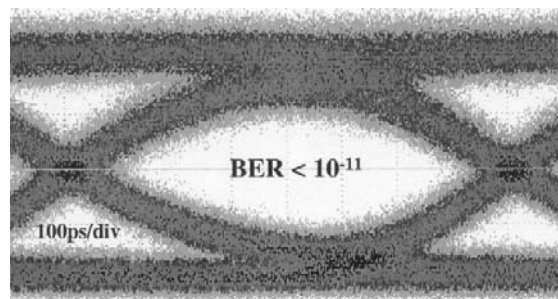


Figure 2.21. An eye diagram for a CMOS/VCSEL transmitter, operating at 1.25 Gbit/s, with the I²-VCSEL biased above threshold and driven by a $(2^{23} - 1)$ -bit-long PRBS.

chip buffers was 17.5 mW from a 3.3 V supply for a -6.9-dBm optical power output at 1.25 Gbit/s.

A much larger CMOS/VCSEL chip with an array of 16×16 I²-VCSELs of the same design was also demonstrated and reported [19].

2.6 DESIGN MODIFICATIONS FOR HIGH-SPEED P²I²-VCSELS

One possible problem with intracavity contacted lateral current injection schemes is their tendency to increase the parasitic capacitance of the VCSELs (note the equivalent circuit for the VCSEL in Fig. 2.20). This increases the RC time constant of the devices and therefore increases their rise and fall times when they are turned on and off. That sets a limit on their modulation speed. Another potential problem with intracavity contacted lateral current injection schemes is that they can make it harder to obtain a uniform current density distribution in the vertical aperture. This is especially true when the lateral current injection is radially symmetric. Holes travel essentially horizontally in their p-doped layer from the p-metal contact towards the aperture, and electrons also travel essentially horizontally, along parallel radial paths, in their n-doped layer from the n-metal contact towards the aperture. The holes and electrons only flow in counter directions vertically inside the aperture region. In such situations the current density can be greater at the rim of the aperture and smaller at the center, whereas for VCSELs the current density should be uniform or even greater at the center of the aperture. Both of the above problems are ameliorated in our new P²I²-VCSEL design, where the conductivity of the bottom doped layer (usually the n-layer) is locally suppressed and where the holes and the electrons are restricted to counterflowing lateral paths (Fig. 2.22). The following paragraphs explain how the redesign attacked the two above problems.

In a VCSEL, the p-doped layer and its contact constitute one plate, and the n-doped layer and its contact is the other plate of the built-in capacitor of the device. Options for reducing the device capacitance are limited. Since the lengthening of a VCSEL cavity is detrimental to the device performance, intracavity plates cannot be placed farther apart. The material between the plates is essentially fixed, so the dielectric constant between the plates is fixed. The area of each plate must be large enough to cover the active area and then must allow room for the metal contacts as well as additional area for fabrication tolerances. Therefore, the plate areas are several times as large as the active areas and in radially symmetric designs can be more than ten times as large, and cannot be reduced without making the fabrication much more difficult. Nevertheless, the excess area contributes to excess parasitic capacitance unless the overlap of the two plates is reduced. In VCSEL designs where the top mirror is deposited and not epitaxially grown, the reduction of the overlap area can be accomplished by a relatively shallow ion implantation that is patterned to locally suppress the conductivity of the bottom-doped layer

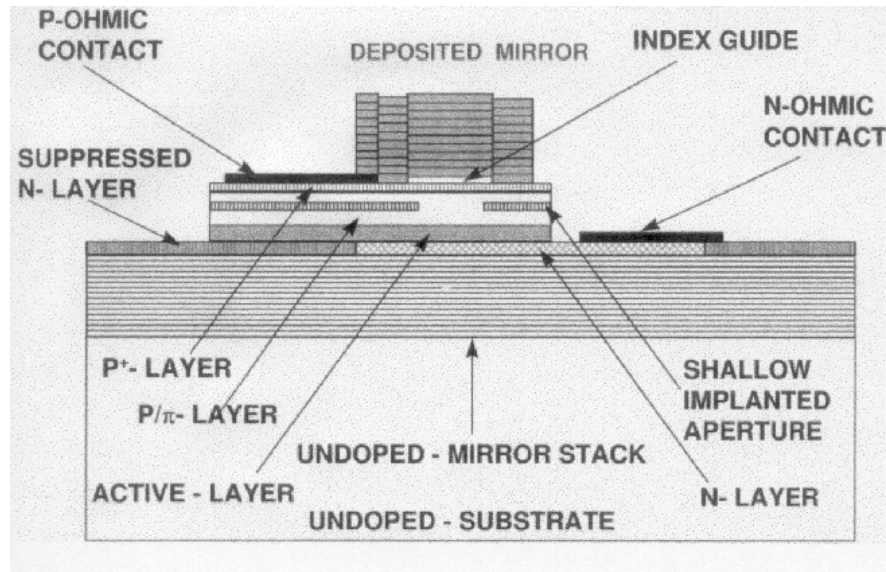


Figure 2.22. A schematic diagram of a cross section of the new high-speed P^2I^2 -VCSEL design. The conductivity of the n-layer is suppressed under the p-contact, so as to reduce parasitic capacitance and the lateral current injection layers and their contacts are so configured as to make all current paths through the aperture essentially equivalent.

(usually n-doped) under most of the top-doped layer (usually p-doped). The overlap area can be made only slightly larger than the active area. The reduction in capacitance can be as large as tenfold. Care must be taken, however, to maintain a low series resistance, R_s (Fig. 2.20), or the net decrease in the RC time constant will not be sufficient.

When the above implantation technique is employed, the VCSELs can have their p-contact to one side of the active area (the left side, for instance) with the n-layer under it suppressed. The n-contact can then be placed on the other side (the right side, for instance) with the p-layer etched away. Under forward bias, holes will then travel from their contact towards the aperture mostly from left to right, and the electrons will travel from their contact predominantly from right to left. If the conductivity of both of the doped layers is about the same, then all current paths through the aperture are equivalent and the current density can be much more uniform than in the radially symmetric case. In the high-speed P^2I^2 -VCSEL design, the current is now carefully guided from one contact to the other.

2.7 DATA ON HIGH-SPEED P^2I^2 -VCSELS AND ITS ANALYSIS

The above processing step has been added as an option to the I^2 -VCSEL fabrication sequence and devices have been made on the same wafer with a

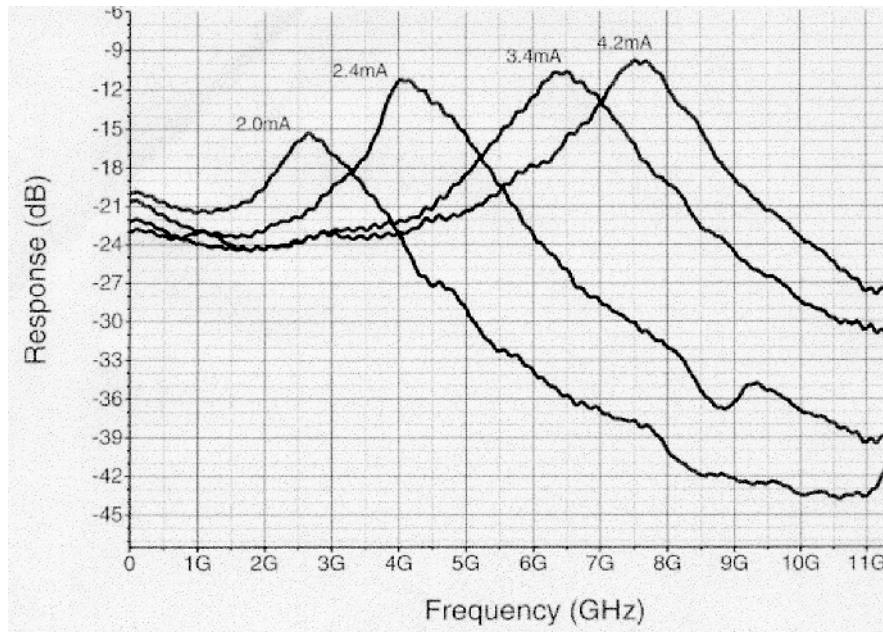


Figure 2.23. The small signal modulation response of a high-speed P²I²-VCSEL at various bias currents, showing a 3-dB bandwidth of ~4 GHz at 2 mA, ~6 GHz at 2.4 mA, ~9 GHz at 3.4 mA, and ~10.5 GHz at 4.2 mA.

mask set containing both the old I²-VCSELs and the new high-speed P²I²-VCSELs. The new designs showed better CW characteristics and were much faster. Figure 2.23 shows the small signal bandwidth reaching 10.5 GHz, at a bias current of ~4.2 mA, for a new 6- μ m-aperture device, lasing single mode, at ~964 nm. During the experiment, the device, still part of a wafer, was just resting on a probe-station chuck, with no special heat sinking or cooling. Heating effects limited the useful range of the device to bias currents < 4.5 mA. Eventually, better alignment of the gain peak with the FP cavity resonance wavelength and flip-chip bonding for better heat sinking should ameliorate the heating effects and thus extend the useful range of devices, at least out to 6 mA, the present limit of the CMOS drivers. The 3-dB bandwidths of the same device, for various currents throughout the useful range, are plotted versus the square root of ($I_{\text{bias}} - I_{\text{threshold}}$) in Fig. 2.24. The data fits a straight line over the entire present range of the device. This fact strongly implies that increases in the bandwidth will follow the reductions in the heating effects. The slope of the line in Fig. 2.24 is the modulation conversion efficiency factor (MCEF) for the VCSEL [20] and is approximately equal to 6.5 GHz/ $\sqrt{\text{mA}}$. The MCEF is a useful parameter for predicting the extent of possible improvements.

An HP lightwave signal analyzer with a built-in tracking generator, which allowed calibrated modulation response measurements from 0–18 GHz, collected the data shown in the above two figures. The signal-analyzer stimulated the VCSEL with a 40-GHz microwave probe through a bias-T (> 12 GHz), while a fiber probe collected the output light, sending it to a high-speed detector and amplifier (25 GHz), which then fed the signal to the analyzer. The response of the bias-T, the photodetector and the amplifier was recorded separately and was then subtracted from the data for the VCSEL.

The expected small signal modulation characteristics of a laser can be approximated as [21]:

$$P(f) = \frac{F_R^4}{(F_R^2 - f^2)^2 + \left(\frac{\Gamma}{2\pi}\right)^2 f^2} \cdot \frac{1}{\left(1 + \frac{f^2}{f_0^2}\right)} \quad (2.1)$$

where F_R is the resonance (relaxation oscillation) frequency of the laser, Γ is the damping frequency, and f_0 is the parasitic (3-dB) cutoff frequency, which includes not only the RC delay but also possible carrier transport limitations. We performed a three-parameter (F_R , Γ , and f_0) fit to each of the data curves in Fig. 2.23, and extracted values for those parameters. The best fit for f_0 , for all of the curves, was 20 GHz. This value is in general agreement with our estimate of the 3RC time constant for that VCSEL, which is ~60 ps, and supports the assertion that the P²I²-VCSEL redesign for high speed (as described in Section 2.6) was effective.

The extracted values for the resonance frequencies and the damping frequencies are plotted in Fig. 2.25, as Γ versus F_R^2 . The data points fit a line and the slope of that line is defined as the K-factor of the laser, which is approximately equal to 95

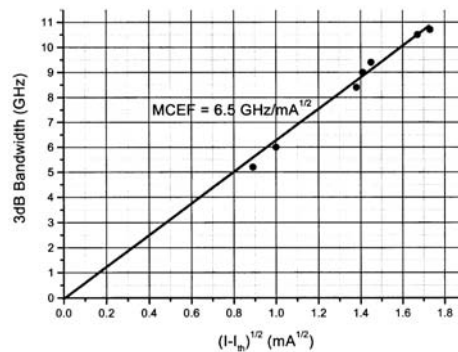


Figure 2.24. The modulation conversion efficiency factor for a 6-μm-aperture, single-mode, high-speed P²I²-VCSEL (defined as the slope of the linear fit to the plot of 3-dB bandwidth versus $(I - I_{th})^{1/2}$).

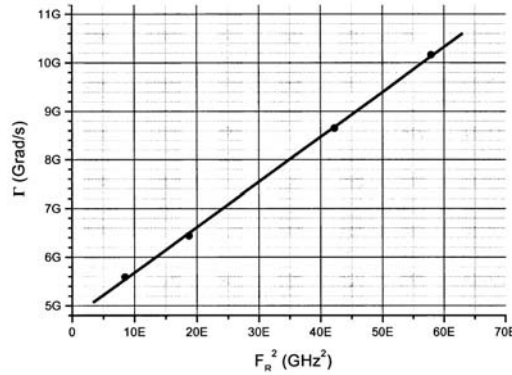


Figure 2.25. A plot of the damping rates versus the squares of the resonance frequencies, for a 6- μm -aperture, single-mode, high-speed P^2I^2 -VCSEL. These values come from the three-parameter fits to the small signal modulation responses of the device. The slope of the linear fit to the data is the K-factor of the laser, used to determine the maximum intrinsic 3-dB bandwidth of the laser, f_{max} .

psec. The importance of the K-factor stems from the fact that it can be used to determine the maximum intrinsic 3-dB bandwidth of the laser, as follows [21]:

$$f_{\text{max}} = \frac{2\sqrt{2}\pi}{K} \quad (2.2)$$

For the 6- μm -aperture, high-speed P^2I^2 -VCSEL under discussion, $f_{\text{max}} = 95$ GHz. This very large value should not immediately raise great expectations. The derivation of Eq. 2.2 assumed strict adherence to Eq. 2.1 with an infinite f_0 . There are many effects that can limit laser bandwidths at values much lower than f_{max} , including the already obvious heating effects and RC parasitics. Carrier transport limits, especially in quantum well lasers, can also be important [22] and are usually indistinguishable from RC delays. The large value of f_{max} does mean that significant bandwidth improvements may be possible and that once the heating effects are ameliorated then f_0 (at 20 GHz) will again become the limitation, either due to RC parasitics or carrier transport problems.

2.8 CONCLUSION

This chapter reviews and describes the major highlights, achieved to date, in an ongoing project. The prime objective is to develop a fabrication platform to reproducibly make high-efficiency VCSELs that easily integrate and operate with electronic circuitry. To that end, the P^2I^2 -VCSEL design strategy was conceived, whereby VCSELs are processed with *photolithographic precision* (for reproducibility) and contain an *intracavity* contacted current guide (for easier integra-

tion) and an *independent* optical beam guide (for increased efficiency). The use of a patterned, deposited dielectric top mirror makes the design strategy feasible, by allowing the formation of crucial device features to occur near the surface with high precision and accuracy, before the top mirror is emplaced. In the first implementation of the P²I²-VCSEL design strategy, *index guiding* was added to an *implanted current aperture* to achieve the desired objective. The resulting devices received the name I²-VCSELs, which emphasizes the novel combination of those two features, which were the most straightforward means for implementing the design strategy. The shorter acronym does fail to point out that the I²-VCSELs are amenable to flip-chip bonding and thus to hybrid integration with electronic circuitry.

The first I²-VCSELs were top-emitting, at ~970 nm, and proved the efficacy of the design with electrical and optical performance characteristics comparable to those of oxide-confined VCSELs, with comparable aperture sizes, and vastly superior to those of implant apertured VCSELs, made prior to this work, with comparable aperture sizes. An added benefit was the fact that, although 10- μ m-aperture devices were multimode, 6- μ m-aperture devices were already single mode. Subsequently, bottom-emitting I²-VCSELs were made and arrays (of 2×10 devices and then of 16×16 devices) were hybrid integrated, via flip-chip bonding, to CMOS chips with specially designed VCSEL drivers, in the 0.5- μ m CMOS technology with a 3.3 V supply voltage. These CMOS/VCSEL chips had optical outputs per channel, at ~970 nm, of up to 4 mW at room temperature and up to 1 mW at 80°C, if driven externally, about half as much if driven by the CMOS transmitter circuits which could deliver a maximum of 6 mA to the VCSELs. The transmitter drivers were able to modulate the VCSELs at data rates of 1.25 Gbit/sec/channel with a BER < 10^{-11} . In one instance, 80 channels operated simultaneously at over 1 Gbit/s each.

The first versions of the I²-VCSELs had a modulation speed limitation characteristic of intracavity contacted VCSELs, attributable to high capacitance and insufficiently uniform current density distributions. The design was modified in subsequent versions to reduce capacitance by suppression of the n-layer conductivity under the p-ohmic contact region, and by placing restricted area contacts on opposite sides of the active region, which improved the current distribution by making all current paths essentially equivalent. The initial results for the new version was a small-signal 3-dB bandwidth of 10.5 GHz, a parasitic cutoff frequency raised to 20 GHz, and other characteristics strongly suggesting the potential for even greater improvements (such as an extracted maximum intrinsic 3-dB bandwidth, $f_{\max}$, of 95 GHz).

The above results demonstrate that the P²I²-VCSEL design strategy is a viable and attractive approach for developing VCSELs for an "Optoelectronic VLSI" technology, where photonic devices are intimately integrated with VLSI electron-

ics. The applications will be systems requiring high bandwidth “data-firehoses” between electronic circuits (aggregate data throughputs of at least a few hundred gigabits-per-second and beyond) [23]. Much more work is needed, however, to guarantee that a P²I²-VCSEL fabrication platform can indeed supply large volumes of reproducibly high yield and reliable VCSEL arrays that would become the backbone of an “Optoelectronic VLSI” technology.

ACKNOWLEDGEMENTS

Numerous colleagues at Bell Laboratories have contributed to this work. We especially appreciate the contributions of J. E. Cunningham, G. Giaretta, S. P. Hui, W. Y. Jan, R. E. Leibenguth, B. J. Tseng, J. M. Vandenberg, J. D. Wynn, and G. J. Zydzik, and the guidance of W. T. Tsang and W. H. Knox.

REFERENCES

- [1] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood, and C. A. Burrus, “Novel hybrid optically bistable switch: The quantum well self-electro-optic effect device,” *Appl. Phys. Lett.*, vol. 45, pp. 13–15 (1984).
- [2] G. D. Boyd, D. A. B. Miller, D. S. Chemla, S. L. McCall, A. C. Gossard, and J. H. English, “Multiple quantum well reflection modulator,” *Appl. Phys. Lett.*, vol. 50, pp. 1119–1121 (1987).
- [3] L. A. D’Asaro, L. M. F. Chirovsky, E. J. Laskowski, S. S. Pei, T. K. Woodward, A. L. Lentine, R. E. Leibenguth, M. W. Focht, J. M. Freund, G. G. Guth, and L. E. Smith, “Batch fabrication and operation of GaAs-Al_xGa_{1-x}As field-effect transistor self-electro-optic effect device (FET-SEED) smart pixel arrays,” *IEEE J. Quantum Electron.*, vol. 29, pp. 670–677 (1993).
- [4] K. W. Goossen, J. E. Cunningham, and W. Y. Jan, “GaAs 850 nm modulators solder-bonded to silicon,” *IEEE Photon. Technol. Lett.*, vol. 5, pp. 776–778 (1993).
- [5] K. W. Goossen, J. A. Walker, L. A. D’Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. D. Bacon, D. Dahringer, L. M. F. Chirovsky, A. L. Lentine, and D. A. B. Miller, “GaAs MQW modulators integrated with silicon CMOS,” *IEEE Photon. Technol. Lett.*, vol. 7, pp. 360–362 (1995).
- [6] J. M. Kuo, Y. C. Wang, K. W. Goossen, L. M. F. Chirovsky, S. P. Hui, B. T. Tseng, J. Walker, A. L. Lentine, R. E. Leibenguth, G. Livescu, W. Y. Jan, J. E. Cunningham, L. A. D’Asaro, A. Ron, D. Dahringer, D. Kossives, D. D. Bacon, R. L. Morrison, R. A. Novotny, D. B. Buchholz, and W. E. Mayo, “Large array of GaAs modulators and detectors flip-chip solder bonded to silicon CMOS using InGaP as the selective etch stop for

- GaAs substrate removal,” *J. Crystal Growth*, vol. 175/176, pp. 971–976 (1997).
- [7] A. V. Krishnamoorthy and K. W. Goossen, “Progress in optoelectronic—VLSI smart pixel technology based on GaAs/AlGaAs MQW modulators,” *Int. J. Optoelectron.*, vol. 11, pp. 181–198 (1997).
 - [8] A. V. Krishnamoorthy and K. W. Goossen, “Optoelectronic—VLSI: Photonics integrated with VLSI circuits,” *IEEE J. of Sel. Topics in Quantum Electron.*, vol. 4, pp. 899–912 (1998).
 - [9] M. H. MacDougall, J. Geske, C.-K. Lin, A. E. Bond, and P. D. Dapkus, “Low resistance intracavity-contacted oxide-aperture VCSELs,” *IEEE Photon. Technol. Lett.*, vol. 10, pp. 9–11 (Jan. 1998).
 - [10] C. Lei, T. J. Rogers, D. G. Deppe, and B. G. Streetman, “ZnSe/CaF₂ quarter-wave Bragg reflector for the vertical-cavity surface-emitting laser,” *J. Appl. Phys.*, vol. 69, pp. 7430–7434 (1991).
 - [11] L. M. F. Chirovsky, W. S. Hobson, R. E. Leibenguth, S. P. Hui, J. Lopata, G. J. Zydzik, G. Giaretta, K. W. Goossen, J. D. Wynn, A. V. Krishnamoorthy, B. J. Tseng, J. M. Vandenberg, and L. A. D’Asaro, “Implant-Apertured and Index-Guided Vertical-Cavity Surface-Emitting Lasers (I²-VCSELs),” *IEEE Photon. Technol. Lett.*, vol. 11, #5, pp. 500–502 (1999).
 - [12] D. L. Huffaker, D. G. Deppe, K. Kumar, and T. J. Rogers, “Native-oxide defined ring contact for low threshold vertical-cavity lasers,” *Appl. Phys. Lett.*, vol. 65, pp. 97–99 (1994).
 - [13] K. D. Choquette and H. Q. Hou, “Vertical-cavity surface emitting lasers: Moving from research to manufacturing,” *Proc. IEEE*, vol. 85, pp. 1730–1739 (Nov. 1997).
 - [14] K. D. Choquette, K. M. Geib, C. I. H. Ashby, R. D. Twesten, O. Blum, H. Q. Hou, D. M. Follstaedt, B. E. Hammons, D. Mathes, and R. Hull, “Advances in selective wet oxidation of AlGaAs alloys,” *IEEE J. Sel. Top. Quantum Electron.* vol. 3, pp. 916–925 (June 1997).
 - [15] R. A. Morgan, J. A. Lehman, and M. K. Hibbs-Brenner, “Vertical-cavity surface-emitting lasers come of age,” in *Proc. SPIE*, M. Fallahi and S. C. Wang, Chairs/Eds., San Jose, CA, vol. 2683, pp. 18–29 (Jan./Feb. 1996).
 - [16] S. J. Pearton, “Ion Implantation in III-V Semiconductor Technology,” *Int. J. Mod. Phys. B*, vol. 7, pp. 4687–4761 (1993).
 - [17] A. V. Krishnamoorthy, L. M. F. Chirovsky, W. S. Hobson, R. E. Leibenguth, S. P. Hui, G. J. Zydzik, K. W. Goossen, J. D. Wynn, B. J. Tseng, J. Lopata, J. A. Walker, J. E. Cunningham, and L. A. D’Asaro, “Vertical-Cavity Surface-Emitting Lasers Flip-Chip Bonded to Gigabits-per-Second CMOS Circuits,” *IEEE Photon. Technol. Lett.*, vol. 11, #1, pp. 128–130 (1999).
 - [18] A. V. Krishnamoorthy, A. L. Lentine, K. W. Goossen, K. W. Walker, J. A. Woodward, J. E. Ford, G. F. Aplin, L. A. D’Asaro, S. P. Hui, B. Tseng,

- R. Leibenguth, D. Kossives, D. Dahringer, L. M. Chirovsky, and D. A. B. Miller, "3-D integration of MQW modulators over active sub-micron transmitter circuit," *IEEE Photon. Technol. Lett.*, vol. 7, pp. 1288–1290 (1995).
- [19] A. V. Krishnamoorthy, K. W. Goossen, L. M. F. Chirovsky, R. G. Rozier, P. Chandramani, W. S. Hobson, S. P. Hui, J. Lopata, J. A. Walker, and L. A. D'Asaro, "16 × 16 VCSEL array flip-chip bonded to CMOS VLSI circuit," OSA Topical Meeting on Optics in Computing, postdeadline paper PD3-1, Snowmass Village, CO (April 14, 1999).
- [20] T. R. Chen, B. Zhao, L. Eng, Y. H. Zhuang, J. O'Brien, and A. Yariv, "Very high modulation efficiency of ultralow threshold current single quantum well lasers," *Electronics Lett.*, vol. 29, pp. 1525–1526 (1993).
- [21] R. Olshansky, P. Hill, V. Lanzisera, and W. Powazinik, "Frequency response of 1.3 μm InGaAsP high-speed semiconductor lasers," *IEEE J. Quantum Electron.*, vol. 23, pp. 1410–1418 (Sep. 1987).
- [22] R. Nagarajan, M. Ishikawa, T. Fukushima, R. S. Geels, and J. E. Bowers, "High-speed quantum-well lasers and carrier transport effects," *IEEE J. Quantum Electron.*, vol. 28, pp. 1990–2008 (Oct. 1992).
- [23] A. V. Krishnamoorthy and D. A. B. Miller, "Firehose architectures for free-space optically interconnected VLSI circuits," *J. Parallel and Distributed Computing*, vol. 41, pp. 109–114 (1997).

Heterogeneous Integration of Vertical-Cavity Surface-Emitting Laser Arrays to CMOS Integrated Circuits

Carl W. Wilmsen and Rui Pu
Colorado State University

3.1 INTRODUCTION

Utilizing photons rather than electrons as information carriers can result in systems with a faster data rate and lower cross talk. However, optical devices, where photons control photons, are presently very difficult to realize [1]. On the other hand, the performance of CMOS integrated circuits (ICs) and VCSEL arrays has increased dramatically in recent years, and provides a strong incentive to devise and study ways of merging the electronic and photonic functions. These integrated circuits are called “smart pixels” and combine electronic processing with optical interconnections. The cost/performance benefit justifies the use of optical technology only when the system clock frequency approaches 1 GHz or when there are a large number of complex I/O interconnects [2]. Therefore, in order to justify replacing the electronic I/O pins with optical interconnections, a smart-pixel array should be large scale and operate at high speed.

The merging of large arrays of VCSELs with CMOS or GaAs MESFET integrated circuits requires a reliable, low-cost method of heterogeneous integration. This is not an easy task because of the material and structural incompatibility between photonic and electronic devices. Several research groups have successfully integrated VCSEL arrays as large as 16×16 to CMOS using a variety of techniques. This is a very positive indication that a practical integration technique will be developed in the foreseeable future. Each of the different integration techniques has strengths and weaknesses, but it is too early to declare the “winner” since it is not clear which of the weaknesses will turn out to be fatal or excessively limiting.

Most researchers agree that the VCSELs must be embedded in the IC chip and not reside as a separate chip adjacent to the CMOS chip. There is, however, disagreement on whether the VCSEL-chip substrate should be removed, and if so, whether the VCSELs should be physically separated from each other. Questions

related to the best bonding metallurgy and the best method of electrical confinement are also important, but these will probably not be the defining ones. The issue of midprocess testability must also be part of the evaluation equation, but testability may be overshadowed by reliability and yield considerations. This chapter first provides a general discussion of these issues and then describes the details of various integration techniques. This is followed by a description of a proposed VCSEL-based communication routing switch.

3.2 GENERAL DISCUSSION OF POSSIBLE METHODS OF BONDING VCSELS TO CMOS

The heterogeneous integration of VCSELS with CMOS or GaAs MESFET integrated circuits requires the resolution of three main questions:

- Should the VCSEL substrate be removed?
- How should the VCSELS be attached to the ICs?
- How should the VCSELS be fabricated?

This section discusses each of these questions in general terms in order to point out strengths and weaknesses. The true answers to the questions, however, will only be found from experimentation in the laboratory and fabrication in a commercial environment.

3.2.1 SUBSTRATE REMOVAL

The bonding of large arrays of VCSELS to electronic chips is best accomplished with the VCSELS still attached to the substrate since this structure is robust and the substrate can serve as a common electrode. Thus, only one electrode per VCSEL needs to be flip-chip bonded to a metal pad on the electronic chip. However, this technique requires the VCSELS to operate at a wavelength transparent to the substrate since the flip-chip-bonded VCSELS must be back-emitting. This usually requires a VCSEL-lasing wavelength ≥ 980 nm since the GaAs substrate is opaque to shorter wavelengths. Using a 980-nm wavelength requires long-wavelength photodetectors (PD). This limits the choice of PD materials; for example, Si is only weakly absorbing at 980 nm, which rules out monolithic PDs on the CMOS chip and requires heterogeneous integration of the PD. Even though monolithic integration is highly desirable, significant progress has been made on the heterogeneous integration of PDs and sensitive receiver circuits. Thus, 980-nm VCSELS may not be a major drawback. The VCSEL array can also be grown on an AlGaAs substrate instead of GaAs allowing the VCSELS to emit

at wavelengths ≤ 850 nm [3, 4]. However, AlGaAs substrates are much more expensive than GaAs and are not commonly used.

The most critical issue related to substrate removal is the thermal-expansion mismatch between the VCSEL array and the CMOS circuitry, which may break the solder attaching the VCSELs to the CMOS chip or reduce the lifetime of the solder bonds. Bonding the array to a transparent structure before removing the substrate, but not separating the VCSELs into individual lasers, could alleviate some of the mismatch strain, but would leave a large, thin structure, which could make the overall structure less robust. The process of etching vias through the VCSEL layers and filling these vias with metal may also reduce yield and reliability. Separating the VCSELs into individual lasers offers a reasonable compromise to these problems. Thermal stress and aging tests should be performed in order to determine the severity of the thermal stress problem and serve as a guide to future development.

The ability to test the VCSEL array before attaching it to the CMOS chip is important in order not to incur the cost of bonding nonfunctioning arrays to expensive CMOS chips. Pretesting can be performed on almost all the reported integration techniques. Those techniques that do not remove the substrate or bond to a transparent structure can probe test each of the VCSELs. Test structures, such as those discussed in Section 3.4.1 can be used to test back-emitting VCSEL arrays. While this does not provide verification that every VCSEL meets specification, it does determine whether the wafer and the basic fabrication process are satisfactory.

3.2.2 METHODS OF ATTACHMENT

Hybrid integration requires attaching the VCSEL to the electronic chip array and making electrical connections between the VCSELs and the electronic circuits. Flip-chip solder bonding is the most common approach since this technique is used for electronic components and provides good electrical and thermal contact. Flip-chip bonding requires postprocessing of the CMOS bonding pads in order to achieve good bonding. To avoid this, the VCSELs could be attached with conducting epoxy, which is screen printed on the CMOS chip or a glass plate. This type of process is used to attach electronic components [5], but so far has not been used for VCSELs. However, it is hard to stencil a fine pattern ($< 60^2 \mu\text{m}^2$) with solder paste because very high paste viscosity is required in order to prevent the paste from flowing off the small bonding pads ($\leq 30 \times 30 \mu\text{m}$) [5]. Also, solder bonding normally requires reflow. This may reduce reliability since an intermetallic reaction may occur between the solder and the metal pad during reflow [7]. Therefore, a solder-bonding technique, which uses thermal compression without reflow, may be better.

Probably the simplest method of VCSEL attachment “glues” the entire VCSEL wafer, epitaxial-layer-side down, to the CMOS chip. The substrate is then removed

and the layers processed into individual VCSELs. The processing is similar to processing of the VCSELs not intended for heterogeneous integration. This was first accomplished using polyimide [8, 9] and later with epoxy [10] as the glue. Metal solder has also been used [11]. The glue must be able to withstand the temperature and environment required to process the VCSEL. The use of oxide confinement is probably the harshest treatment and therefore it may be best to use ion implantation for this integration technique. Obviously, the VCSELs can not be tested before attachment to the CMOS chip since they are fabricated after attachment. Also, polyimide and epoxy have low thermal conductance and therefore must be very thin. This has been shown not to be a significant problem [8, 11].

Electrical connection must also be made between the CMOS chip pads and VCSELs. If the VCSEL chip is flip-chip bonded to the CMOS and the substrate is not removed, the solder forms one of the contacts and the VCSEL substrate provides the other. The other techniques require two contacts, flip-chip solder or thin film, neither of which should cause significant fabrication problems. Forming a low-resistance contact is the main goal. To date, the resistance of the heterogeneously integrated VCSELs is significantly higher than that of standard VCSELs, thus more research is required to ensure high-performance devices.

From the above discussion it appears that the best commercial method for attachment and electrical connection of VCSELs to CMOS chips is not obvious, and thus, several techniques must be investigated. In this way, the strengths and weaknesses of the techniques can best be evaluated.

3.2.3 VCSEL FABRICATION

The main issues related to VCSEL fabrication are:

- The compatibility of the fabrication process with the integration-processing sequence,
- The efficiency of the resulting VCSELs, and
- The uniformity and repeatability of the process.

There are at least five different ways of fabricating VCSELs: air post, ion implantation, oxide confinement, buried heterostructure, dielectric aperture, and combinations of these [13]. The main considerations for the choice of fabrication technique are high VCSEL efficiency and compatibility with the other processes. Oxide confinement holds the record for highest efficiency and lowest threshold current and therefore is very appealing for large arrays [14]. However, the process results in a nonplanar structure that effects both the photolithography and bonding. Standard ion implantation results in a planar structure if the substrate is not

removed since this allows using the substrate as a contact. The scraggle from the deep-proton implantation can introduce defects in the active area, thereby increasing the threshold current and reducing the efficiency. If the VCSEL is back-emitting, the implant must be into the thicker mirror. This increases the scraggle. On the other hand, implantation after bonding is not feasible if the substrate is not removed. If the substrate is removed, then the surface topography will probably hinder processing the thick photoresist mask required for ion plantation. Also, postbonding handling is not desirable.

It probably will be best to develop new VCSEL-fabrication processes which are specifically designed for bonding. This would involve combining some of the present processes to establish an optimum design for a given bonding technique. Chirovsky et al. [15] report using a combination of shallow implantation of fluorine or oxygen and dielectric mirrors to form an efficient VCSEL structure that bonds well to CMOS. Other such structures are possible.

3.3 REVIEW OF HETEROGENEOUS INTEGRATION OF VCSELS AND ELECTRONIC CIRCUITS

The hybrid integration of large arrays of VCSELs to integrated circuits is a much more challenging process than the attachment of electronic components to ceramic substrates since VCSELs are small, have a complex structure, and must be integrated with a much greater density. VCSELs require relatively small bonding pads which adds to the processing difficulties. Furthermore, VCSELs are not completely fabricated before bonding and both they and the ICs require postfabrication processing.

A number of research laboratories are investigating VCSEL bonding and much of their success is reported in other chapters of this book and therefore will not be discussed here. Pu et al. [16] were the first to report bonding an array of VCSELs directly to a foundry-fabricated integrated circuit. They bonded an 8×8 -VCSEL array to GaAs MESFET IC [8]. After that, Krishnamoorthy and his coworkers at Lucent Technology reported bonding a 2×10 array of 970-nm VCSELs to a CMOS chip [17]. Despite several differences in details, both of these works are based on coplanar flip-chip-bonding techniques, which had been used to make SEED arrays [7, 18]. Trezza and coworkers at Sanders [19] and Liu and coworkers at Honeywell [20] have also reported coplanar bonding processes.

Previous to these array-bonding reports, Mathine et al. [21] at Arizona State University developed the appliqué technique to transfer each individual VCSEL to its corresponding metal pad on the CMOS chip using a mechanical probe tip. The electrical contacts between the VCSEL and the CMOS chip were then formed. The appliqué technique is similar to that developed by Jokerst and coworkers at Georgia Tech for photodetectors [22].

Yeh and Smith at UC Berkeley [11] used an In/Au alloy to bond a VCSEL wafer and a Si substrate. After the substrate of the VCSEL wafer was removed, the VCSEL wafer was processed into individual devices. Although the conductive wafer-bonding material provides the VCSELs' electrical contacts, the VCSELs are not individually addressable. Matsuo et al. at NTT Optoelectronic Laboratories used polyimide to replace the In/Au [8, 9]. Polyimide is IR-light transparent; therefore, one can view through the VCSEL layers and the electronic circuits with an IR light source. Unfortunately, polyimide has low thermal conductivity, which may affect the performance of VCSELs, although this has been shown not to be a serious problem [12]. Lee et al. [23] at UC San Diego etched large holes in CMOS chips and then placed the VCSEL array, still attached to the substrate, through the holes [24]. They reasoned that the grafting of silicon devices is less critical and easier than the grafting of GaAs devices. However, drilling holes on a CMOS chip occupies a large area of the electronic chips and may affect the yield of the CMOS chip. The etching of these large holes is not allowed in standard CMOS design; therefore, this technique is not compatible with CMOS foundry fabrication.

These early bonding investigations proved that the VCSEL structure can withstand the rigors of postfabrication bonding and substrate removal. Their success, along with the development of the SEED bonding process, has shown the way for the bonding of VCSEL arrays to foundry-fabricated integrated circuits.

3.4 OVERVIEW AND DETAILS OF BONDING TECHNIQUES

Three types of heterogeneous-integration techniques using both conductive and nonconductive attachment were investigated. All of these technologies require substrate removal. The three technologies are (1) coplanar flip-chip, (2) conductive-attachment top-bottom contacts, and (3) nonconductive attachment top contacts. In this section, these three fabrication techniques are described in detail.

VCSEL wafer design for these bonding techniques is similar to that of a standard 850-nm oxide-confined VCSEL wafer, except the n-mirror has fewer layers than the p-mirror. This allows stronger emission from the bottom side of the VCSEL stack after substrate removal. In order to facilitate substrate removal, the VCSEL wafer should also have an AlAs stop-etch layer between the mirror stack and the substrate. In addition, the VCSEL wafer should have intracavity contact layers for both p- and n-contacts.

A 250- μm pitch was used for all of the VCSEL arrays; this is equal to the density of 1600 VCSEL/ cm^2 . The bonding pads can be made as small as $10 \times 10 \mu\text{m}^2$ [17]; however, considering that a reasonable mesa size for oxide-confined VCSELs is $40 \times 40 \mu\text{m}^2$ and small bonding-pad size increases the difficulty of heat dissipation, the bonding pads used were $30 \times 30 \mu\text{m}^2$. Arrays of 4×4 or 8×8 were used in the research.

Aiming at hybrid integration of foundry-fabricated electronic chips, patterned GaAs substrates were first used as a test structure in order to develop the bonding techniques. Later, VCSEL arrays were bonded to MESFET and CMOS chips. There appears to be no significant difference between bonding VCSELs to a GaAs dummy chip or a foundry-fabricated microelectronic chip. Furthermore, it is unlikely that the integration process will affect the foundry-fabricated chips as long as high temperature ($> 500^{\circ}\text{C}$) or a large force are not applied to the chip.

3.4.1 COPLANAR FLIP CHIP BONDING

The coplanar-bonding process is illustrated in Fig. 3.1. The oxide-confined VCSELs were fabricated by dry etching the device mesa down to the n-mirror and wet oxidizing the confinement layers [Fig. 3.1(a)]. The mesas were oxidized at 430°C for 25 minutes in water vapor resulting from N_2 bubbled through 75°C water. The oxidation resulted in approximately $15 \times 15\text{-}\mu\text{m}^2$ apertures when the mesa was $40 \times 40\text{ }\mu\text{m}^2$. The mesa etching results in a wafer surface with $4\text{--}5\text{-}\mu\text{m}$ steps. While this decreases the resolution of subsequent photolithography, VCSEL processing only requires a few-micrometer resolution, which is readily obtainable. Furthermore, the coplanar flip-chip structure does not require a metal aperture (this is the smallest feature of the VCSEL). Thus the resolution requirement is even less critical.

The p- and n-contacts were then deposited and annealed [Fig. 3.1(b)]. Then, $500\text{-}\text{\AA}$ Ti/ $1500\text{-}\text{\AA}$ Au layers were deposited for the p-contacts, while the n-contacts were composed of $250\text{-}\text{\AA}$ Ge/ $520\text{-}\text{\AA}$ Au/ $150\text{-}\text{\AA}$ Ni/ $1000\text{-}\text{\AA}$ Au. These contacts were annealed at 375°C for 40 seconds. The metal contacts were deposited after oxidation to avoid the corrosive atmosphere and high temperature of the oxidation step, both of which can degrade the contact resistance. When better alignment and fine structure are required, the metal contacts need to be deposited before the mesa etch, therefore before the oxidation. Then, special design of the metal contacts is necessary in order to protect the metal from the corrosive atmosphere and prevent the Au from diffusing into the VCSEL deeply during high temperature.

After annealing the contacts and depositing a seed layer, a gold post was electroplated on the n-contact to a height approximately level with the p-contact. This was followed by electroplating gold on both p- and n-contacts in order to enhance the thermal-compression bonding [Fig. 3.1(c)]. The accuracy of the height of the two electroplated coplanar gold posts must be controlled to within $0.3\text{ }\mu\text{m}$, otherwise the bonding process has a low yield or the bonded VCSEL will have nonuniform characteristics. Even if electrical contact is achieved, a barely-touching contact results in high thermal resistance and low mechanical strength. Therefore, the quantum efficiency and the maximum optical-output power and other characteristics of the VCSELs can vary over a large range if the bonding is not done properly. In order to achieve uniformity, the rate of electroplating must be accurately controlled. The electroplating rate depends on the size of the sample, the pattern on the sample, elec-

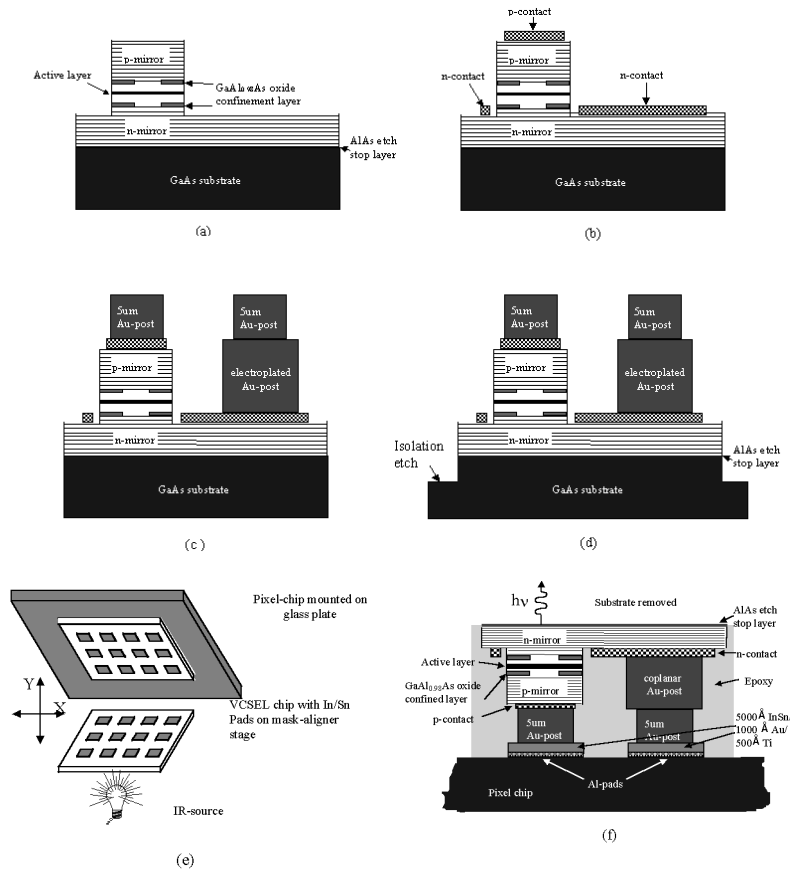


Figure 3.1. Process of coplanar flip-chip bonding [25].

troplating current, and temperature. While the latter two are easy to control, the other two cause the electroplating rate to vary, and repeatability is difficult. With similar samples, the rate can be estimated, but it is seldom exactly the same.

After electroplating, a mesa surrounding the entire VCSEL structure was formed by wet etching down to the substrate [Fig. 3.1(d)]. This step is required to separate the VCSELs when the substrate is removed. Although wet etch results in an undercut, a smooth surface is obtained. The smooth surface is very helpful for viewing through the VCSELs for the alignment of the subsequent bonding to the electronic chip. The VCSEL processed through step D is shown in the scanning-electron micrograph of Fig. 3.2.

Since the bonding pads of foundry-fabricated ICs are aluminum, 500-Å Ti/1500-Å Au/6000-Å InSn was deposited onto these pads. Ti is the adhesive layer and the Au serves as a solder-wettable layer for the InSn solder. The flip-chip bonding of the VCSELs to the IC chip was accomplished by mounting the

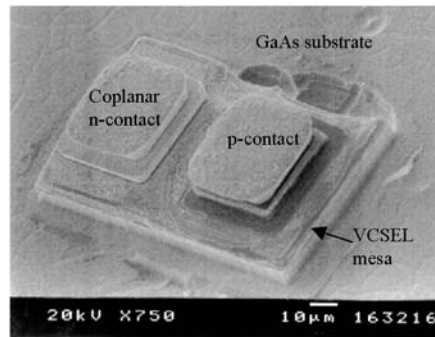


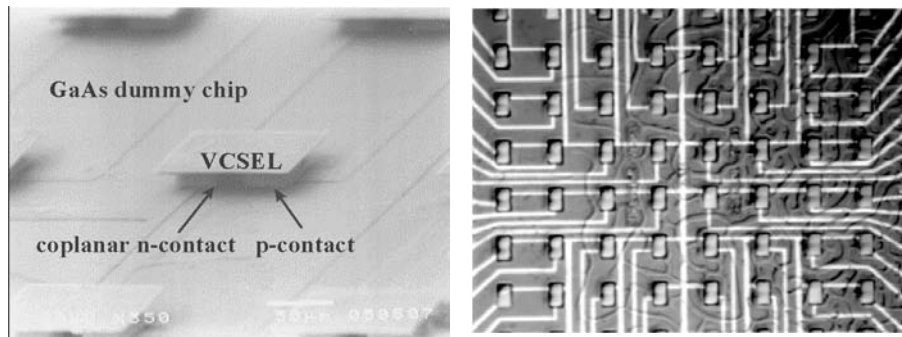
Figure 3.2. SEM photomicrograph of a coplanar VCSEL.

VCSEL chip onto a glass plate with a crystal bond, aligning it with the pixel chip in a mask aligner using IR to view the two structures, and then bonding the two chips together by applying pressure [Fig. 3.1(e)]. If this part of the process is done properly, then the unit can be easily transferred to a hotplate and heated to 180°C for 12 minutes. This melts the InSn and forms a eutectic with the Au pads. Since the present volume of the InSn is small, it is believed that alloying of the InSn with Au occurs, and not reflow. The yield of the bonding is limited by the tilt between the VCSEL arrays and the pixel chip. To improve the yield, the mask aligner can be calibrated so that the two glass plates are parallel. However, the crystal bond can also introduce a tilt. In order to reduce this problem, a very thin crystal bonding layer is formed by melting it at a slightly higher temperature (200°C). However, experience indicates that the tilt is mainly from the polishing of the VCSEL substrate. VCSELs are normally grown wafers with an unpolished backside. However, in order to view through the chips, the VCSEL wafer was polished. The quality of polishing was limited by the available lapping facility. Replacing the mask aligner with a bonding machine will eliminate this problem.

After the bonding, epoxy is wicked in between the chips to enhance the robustness of the structure and protect the pixel chip from attack by the subsequent polishing and the selective etch used to remove the GaAs substrate [Fig.1(f)]. To remove the substrate, it is first polished to ~100 µm, then selectively etched away. This process leaves free-standing VCSEL mesas bonded to the electronic chip. It has been found that a PA solution (H₂O₂:NH₄OH 250 ml:8 ml) spray etch provides extremely good selectivity at the AlAs etch-stop layer and consequently a very smooth top-emitting surface. The epoxy can be removed by plasma etch if necessary.

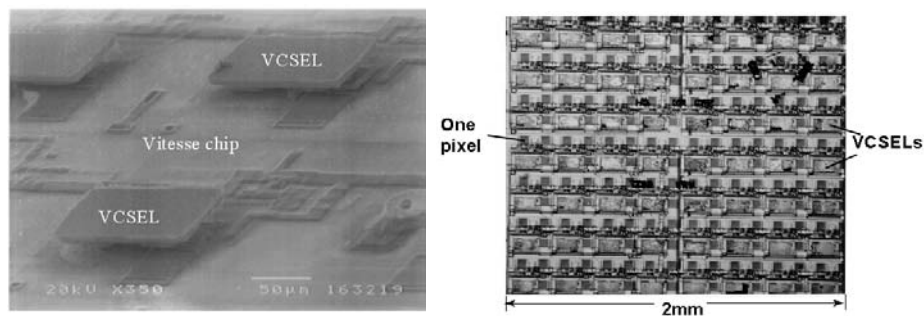
The coplanar flip-chip technique was developed by bonding the 8 × 8-VCSEL array onto a GaAs dummy chip, as shown in Fig. 3.3.

Using this technique, an 8 × 8-VCSEL array was bonded onto a MESFET chip as illustrated in Fig. 3.4. The MESFET chip was designed with the Vitesse



(a) SEM photomicrography

(b) Optical photomicrograph

Figure 3.3. 8×8 VCSEL array coplanar flip-chip bonded to a GaAs dummy chip.

(a) SEM photomicrograph

(b) Optical photomicrograph

Figure 3.4. 8×8 -VCSEL array coplanar flip-chip bonded to a MESFET chip [16].

Semiconductor $1.0\text{-}\mu\text{m}$ E/D MESFET process and fabricated by Vitesse through the MOSIS Foundry Service. Each smart-pixel circuit contained 32 MESFETs and three MSM photodetectors, which perform the selection function of a data filter [16, 25] and also provides the current drive for the VCSEL. Figure 3.5 shows a fabricated pixel with and without a coplanar-bonded VCSEL.

Using this flip-chip process, a 4×4 -VCSEL array was also bonded to a CMOS chip, as shown in Fig. 3.6. The CMOS chips were designed with $0.5\text{-}\mu\text{m}$ design rules and fabricated by Hewlett-Packard through the MOSIS Foundry Service. Each pixel contained a four-stage CMOS amplifier and an NMOS single-stage current driver with a parallel prebiasing transistor.

3.4.2 TOP-BOTTOM CONTACT BONDING

The top-bottom contact-bonding process is illustrated in Fig. 3.7. Depositing the p-contacts is the first step in the process. This is followed by electroplating a

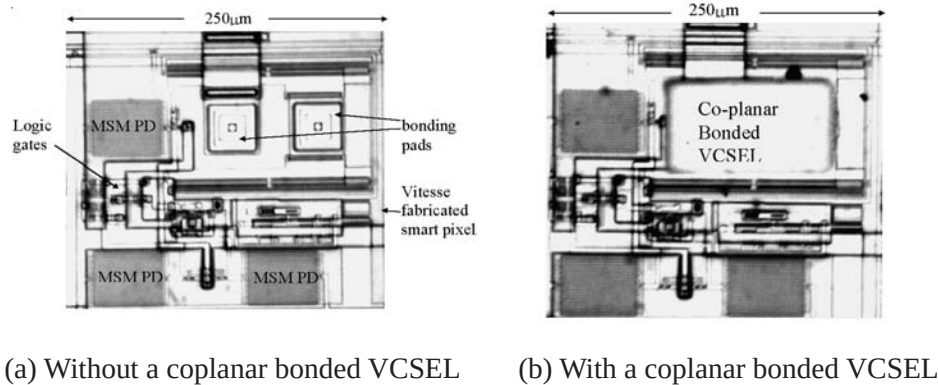


Figure 3.5. Optical photograph of a Vitesse fabricated smart pixel [16].

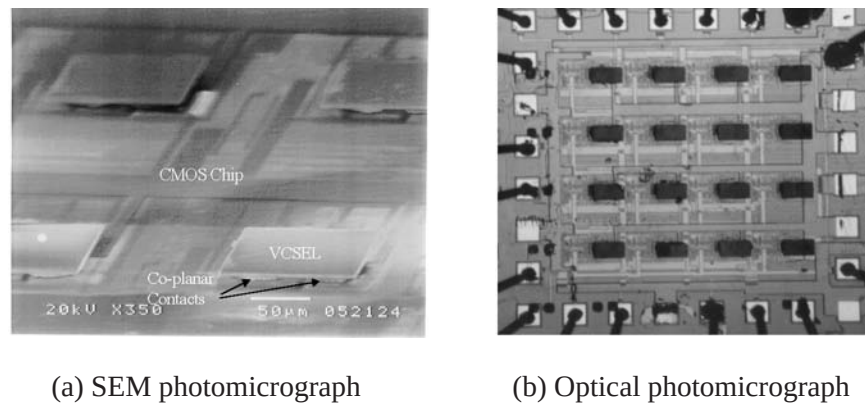


Figure 3.6. 4×4 -VCSEL array coplanar flip-chip bonded to a CMOS chip [16].

gold post [Fig. 3.7(a)]. The VCSEL mesa was then isolated by dry etching into the substrate, and the confinement layers oxidized by heating in steam at 430°C [Fig. 3.7(b)]. The gold post serves as the etch mask. The VCSEL mesa is the same size as that of coplanar flip-chip bonding discussed previously and is oxidized under the same condition. The metalizations were carried out before the etching, and therefore before the oxidation, because these VCSEL mesas are about $9\text{ }\mu\text{m}$ high, which reduces the resolution of the photolithography. Furthermore, using the gold posts as the mask for mesa etching also saves one photolithography step. The p-contact was formed with $400\text{-}\text{\AA}$ Ti/ $200\text{-}\text{\AA}$ Pt/ $1500\text{-}\text{\AA}$ Au metalization instead of Ti/Au. The Pt was needed as a diffusion barrier to prevent Au from diffusing deeply into the device during the 430°C oxidation. The electroplated gold post also protects the ohmic contact from the corrosive environment.

Ti/Au/InSn was deposited onto the Al p-contact pads of the pixel chip. An additional gold post was electroplated on the n-contact Al pad of the pixel chip to

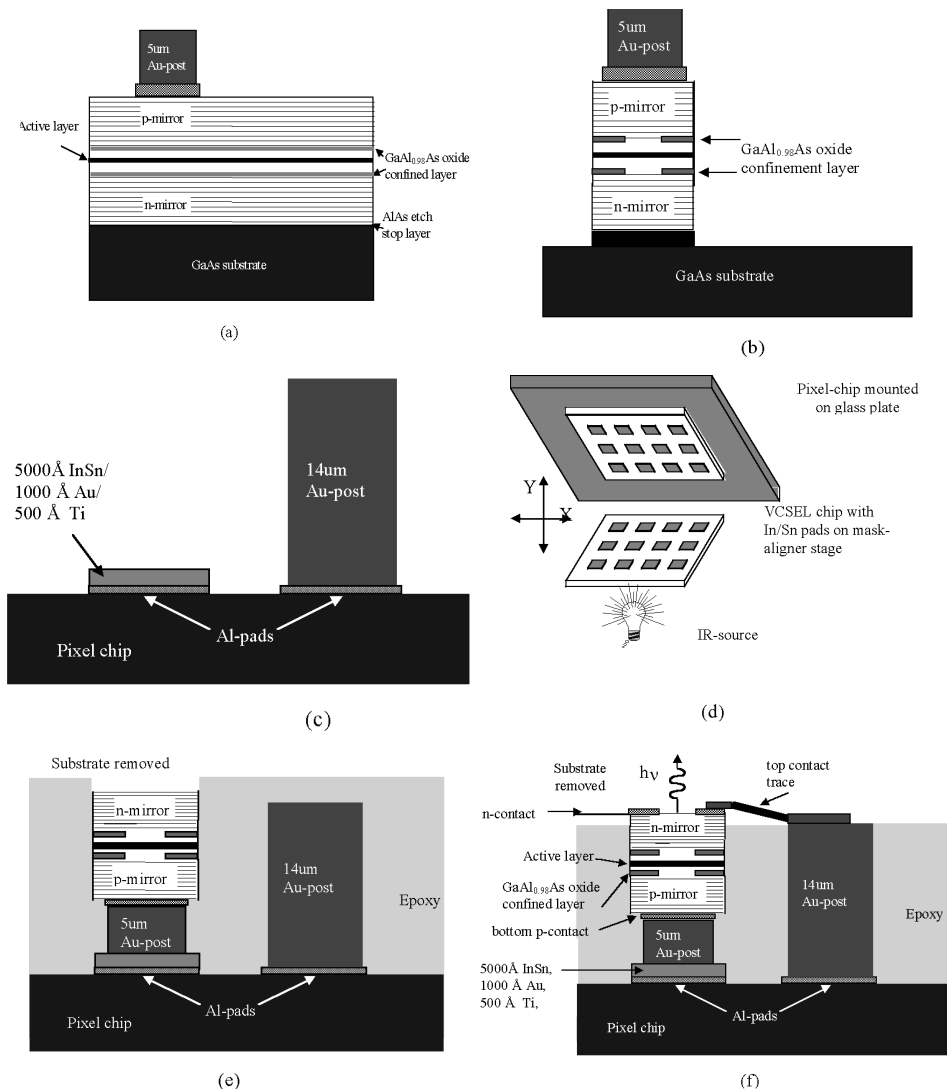


Figure 3.7. Process of top-bottom contact bonding.

the approximate height of the VCSEL in order to form the top contact [Fig. 3.7(c)]. The height tolerance of this post is greater than that with the coplanar technique, because this post is not solder-bonded.

The bonding [Fig. 3.7(d)] and substrate-removal process [Fig. 3.7(e)] is similar to that used in the coplanar flip-chip, except the bonding alignment is not as critical because there is only one solder-bonding pad for each VCSEL. Furthermore, there is no coplanar height accuracy problem. Therefore, higher bonding yield compared with coplanar flip-chip bonding is expected.

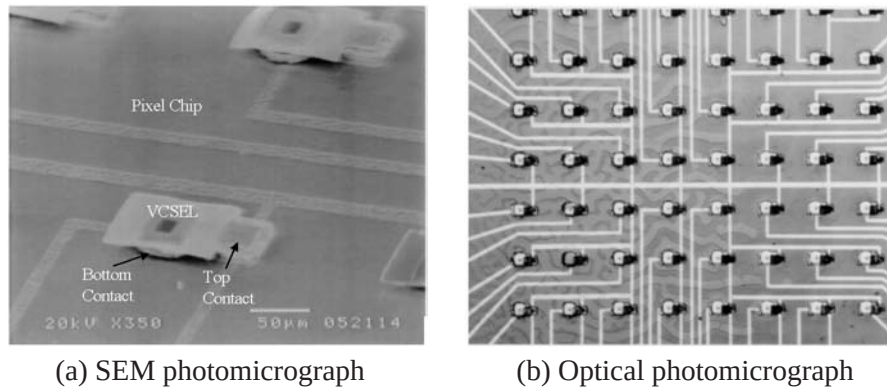


Figure 3.8. 8×8 -VCSEL array top-bottom bonded to a GaAs dummy chip [10].

If the Au posts on the pixel chip were lower than the VCSEL mesas then epoxy will cover the top of the Au posts on the pixel chip. This can be removed with an oxygen plasma etch [Fig. 3.7(f)]. Following this step, the n-contacts and the contact traces between the VCSEL n-contacts and the Au posts were deposited. These were then thickened by electroplating in order improve the step coverage and make the electrical connection more reliable. After these steps, the chip was annealed.

The aperture on the n-contact is the smallest pattern required by this technique. Therefore, it is the most critical photolithography step. Since this photolithography step was performed after bonding the VCSEL, a non-planar surface is unavoidable. According to experience, the radii of the patterned apertures are normally $2 \mu\text{m}$ smaller than the mask. Designing a bigger-aperture-size mask can solve this problem. Step coverage with this process is not a serious problem since the epoxy provides a bridge between the VCSEL. In addition, increasing the thickness of the metal trace by electroplating also enhances the step coverage.

The SEM photomicrograph of Fig. 3.8(a) illustrates part of an 8×8 array of VCSELs bonded by this technique to a GaAs dummy chip. Figure 3.8(b) is an optical photomicrograph of this 8×8 bonded array.

3.4.3 TOP-CONTACT BONDING

Another approach to attaching individual VCSELs is the technique of “glueing” the entire VCSEL to the electronic chip, removing the VCSEL substrate and processing the remaining layers into individual VCSELs. Figure 3.9 illustrates the process of top-contact bonding with this kind of attachment.

The VCSEL wafer was first glued face-down onto the electronic wafer, using a nonconducting material such as epoxy or polyimide [8–10] as shown in Fig. 3.9(a). In order to prevent the very thin epitaxial layers from being distorted, the nonconductive material should have low shrinkage after curing. Also, since nonconductive materials normally are not good thermal conductors, a thin layer is pre-

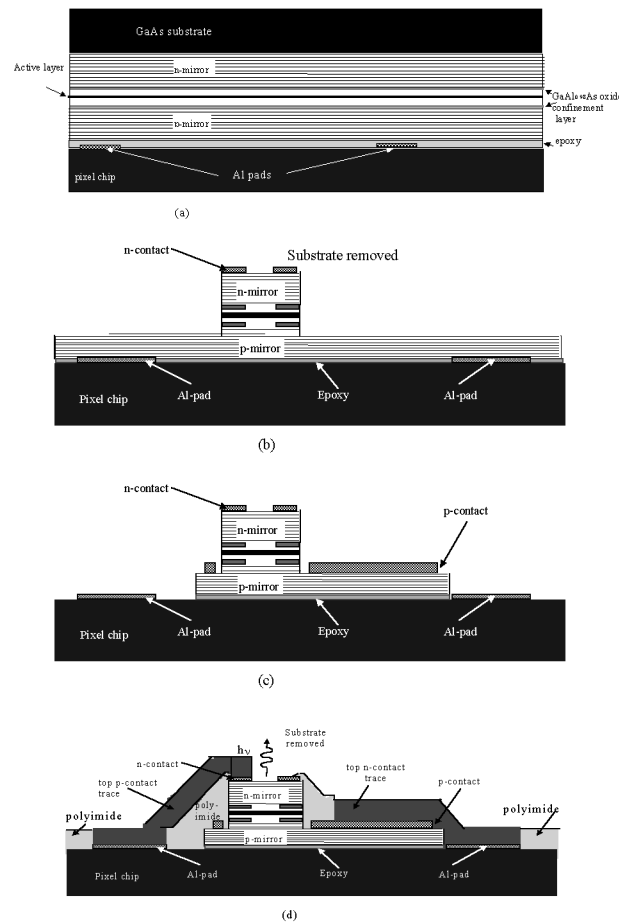


Figure 3.9. Process of top-contact bonding.

ferred in order to reduce the thermal resistance. Low viscosity is required in order to obtain a thin uniform layer of bonding material. Epoxy was chosen as the bonding material for this research since it has low viscosity. Experiments also verified that better bonding is accomplished with epoxy than with polyimide: even though epoxy data sheets do not recommend heating above 150°C, the VCSELs were firmly attached after the 430°C oxidation.

After the substrate was removed, the processing steps for this technique were similar to the first two techniques. First, n-contacts/apertures were deposited on the exposed n-mirror and the device mesas were dry etched down to the p-mirror as shown in Fig. 3.9(b). Since the metal aperture is the smallest pattern in this technique, the required photolithography for this step was performed first.

The p-contacts were then deposited and isolation mesas were wet etched with the epoxy serving as the wet-etch stop layer [Fig. 3.9(c)]. The oxidation was car-

ried out after these etching steps, leaving individual VCSEL mesas on the surface of the pixel chip. Oxidation before the isolation etch could cause the layers to break due to the difference in thermal expansion.

The very thin layer of epoxy or polyimide used in this bonding ($< 1 \mu\text{m}$) cannot protect the Al bonding pads on the pixel chip from the corrosive environment of the oxidation. Therefore, the Al bonding pads need to be protected before oxidation by other methods, such as depositing Au.

The epoxy not under the VCSEL mesas was removed with an oxygen plasma etch [Fig. 3.9(c)].

Polyimide was spun on in order to cover the VCSEL mesas and was then patterned [Fig. 3.9(d)] to provide an insulating layer and better hold the VCSEL mesas in place. Gold traces were then electroplated to connect the p- and n-contacts to their corresponding bonding pads on the pixel chip. The Au trace step coverage is reliable, since the polyimide planarizes the surface and the polyimide sidewalls in the contact holes are 45° .

The SEM photomicrograph of Fig. 3.10 shows individual VCSELs of an 8×8 array bonded to a specially prepared GaAs dummy substrate using the top-contact technique. The optical photomicrograph illustrates this 8×8 array.

3.5 ELECTRICAL AND OPTICAL CHARACTERISTICS

An ideal bonding technique should not affect the performance of either the VCSEL and the CMOS chips. The best way to evaluate the bonding technique is to compare the VCSELs before and after bonding. However, since the bonded VCSELs are bottom-emitting, the characteristics of the bonded VCSEL, e.g. threshold current and output optical power, could be quite different from the top-

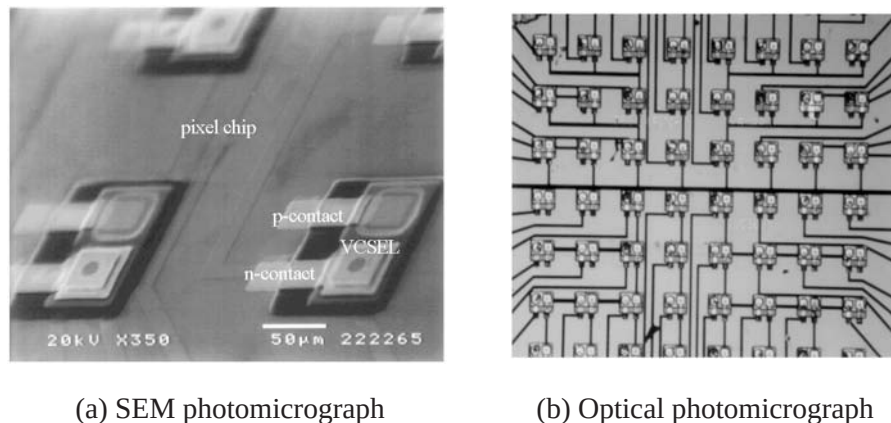


Figure 3.10. 8×8 -VCSEL array top-contact bonded to a GaAs dummy substrate [10].

emitting VCSELs before bonding (on-wafer test). The comparison can only be made by considering the differences between the top-emitting and the bottom-emitting VCSELs. Details are given below.

In order to make a better comparison of these three bonding techniques, the bonded VCSEL arrays fabricated from the same VCSEL wafer and with the same aperture size were measured.

3.5.1 PREDICTING THE CHARACTERISTICS OF BONDED VCSELs FROM ON-WAFER TESTING

The characteristics of the bonded VCSELs depend on the initial quality of the VCSEL wafer, the VCSEL processing, and the bonding processing. Since the VCSEL wafer used emitted at 850 nm and was grown on a GaAs substrate, which absorbs 850-nm-wavelength light, the on-wafer test could only be top-emitting. However, the bonded VCSELs are bottom-emitting after the substrate was removed.

The difference between the threshold current for the top- and back-emitting VCSELs is due to the changes in the reflectivity of the mirrors as a result of integration. In coplanar flip-chip bonding and top-bottom contact bonding, the p-contact metal served as a reflective layer, thereby increasing the reflectivity of the p-mirror. Substrate removal also changed the reflectivity of the n-mirror. Considering the relationship between threshold and VCSEL mirror reflectivity, the characteristics of bonded VCSELs can be predicted from the on-wafer test.

The reflectivity R of the quarter-wavelength DBR mirror at the resonant wavelength is given by [26]:

$$R = \left[\frac{1 - \frac{n_s}{n_0} \cdot \left(\frac{n_1}{n_2} \right)^{2m}}{1 + \frac{n_s}{n_0} \cdot \left(\frac{n_1}{n_2} \right)^{2m}} \right]^2, \quad (3.1)$$

where m is the number of the nonabsorbing DBR pairs, n_1 and n_2 are the reflective indices of the low- and high-reflection layer of the DBR mirror, and n_s and n_0 are the reflective indices of the materials on each side of the DBR mirror.

If the cavity and mirror losses are ignored, the relation between the mirror reflectivity R_1 , R_2 and the threshold gain g_{th} is given by [26]:

$$g_{th} = \frac{1}{2 \cdot l} \cdot \ln \left(\frac{1}{R_1 \cdot R_2} \right), \quad (3.2)$$

where l is the length of the active region.

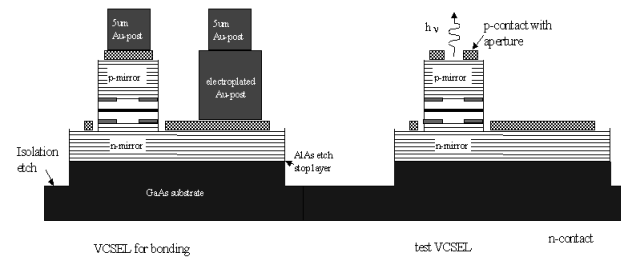


Figure 3.11. Coplanar flip-chip bonding VCSEL and test VCSEL.

For the coplanar flip-chip bonding technique, the mask set provided for the simultaneous fabrication of a top-emitting test VCSEL as shown in Fig. 3.11. Therefore, the test VCSEL and the bonding VCSELs are fabricated simultaneously and have the same oxide-defined aperture and electrical contacts.

Figure 3.12(a) shows the I-V-L characteristics of the top-emitting test VCSEL with a threshold current of 10 mA for the 14- μm aperture size. After the VCSEL arrays were coplanar flip-chip bonded to a GaAs dummy chip, and the substrate of the VCSEL array removed, the back-emitting VCSEL was tested. The results are shown in Fig. 3.12(b). The threshold current is reduced to ~ 4 mA. The dramatic reduction of the threshold current fits very well with the calculation results, as shown in Table 3.1. According to the calculation, the threshold gain of the coplanar flip-chip bonded VCSEL is about one third of the top-emitting VCSEL, which is the approximate change in the ratio of the threshold gains, $470/1978 = 0.32$. Therefore, except for changing the reflectivity of the mirrors, the process itself will not impose any significant change on the characteristics of the VCSELs.

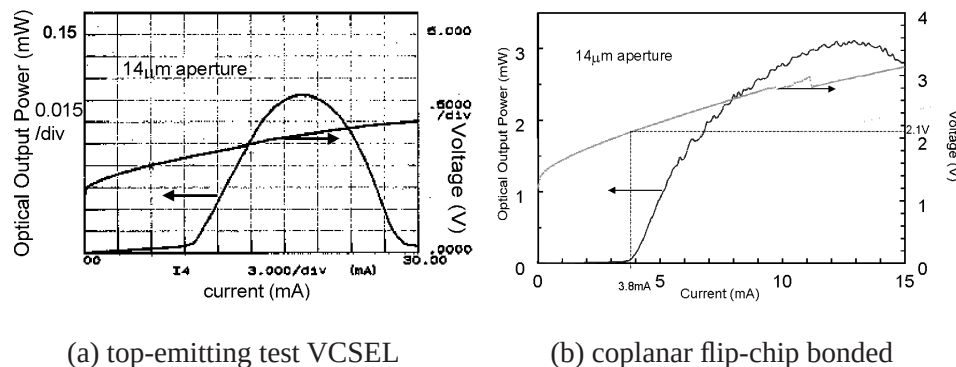


Figure 3.12. I-V-L of the VCSELs made from wafer EMC1632.

	Calculated			Experimental
On-Wafer Test	$R_{p\text{-mirror}}$	$R_{n\text{-mirror}}$	$g_{th} \text{ (cm}^{-1}\text{)}$	$I_{th} \text{ (mA)}$
Top-emitting test	0.99853	0.98970	1478	10
VCSEL				
Coplanar flip-chip bonded VCSE	~ 1.0	0.99625	470	3.8

Table 3.1. Summary of the calculated and experimental results of the on-wafer test VCSEL and flip-chip bonded VCSEL.

3.5.2 CHARACTERISTICS OF THE BONDED-VCSEL ARRAYS

In order to compare the characteristics of VCSELs bonded by the three different techniques, VCSEL arrays from the same wafer and oxidized at the same time were fabricated and bonded to patterned GaAs wafer. The VCSEL wafer was not ideal since there were no intercavity contact layers and the wafer was designed for top emission. Thus, the output power was low and the contact resistance higher.

Figure 3.13 shows the I-V and I-L curves for 12 of these VCSELs coplanar flip-chip bonded to the GaAs substrate. The threshold current is in the range of 1.7–2.1 mA, and the threshold voltage between 2.7 and 2.9 V. The average series resistance is 100 Ω . The nonuniformity of the series resistance is due to the metals being deposited on the n-mirror layers and not on an interacting layer. Since the mirror layers are only a few hundred angstroms thick and the nonuniformity of the dry etch is 2000 Å, the device mesa etch can not be controlled to stop at the GaAs layer rather than the AlGaAs layer. Furthermore, the AlGaAs layer was also exposed to the corrosive environment for the oxidation and the Al_2O_3 is difficult to remove.

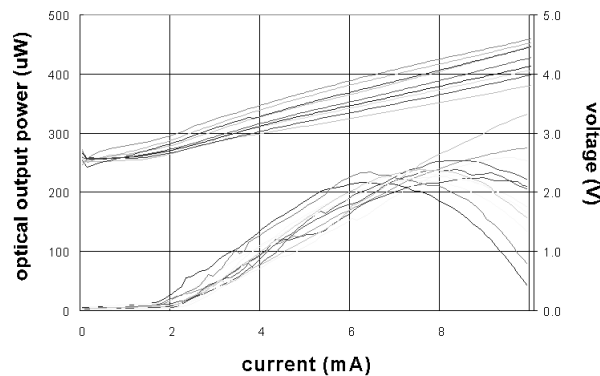


Figure 3.13. I-V-L curves for the VCSEL array coplanar flip-chip bonded to the GaAs substrate.

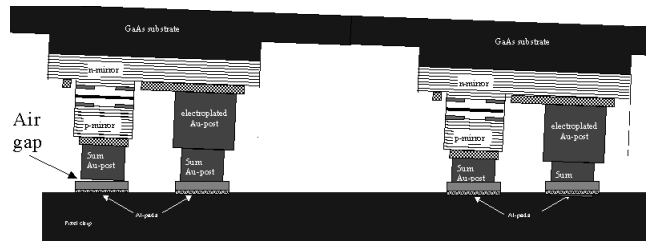


Figure 3.14. The contact which provides electronic connection but poor thermal conduct.

As expected, the maximum output optical power is low, approximately 200 μ W, because of the high reflectivity of the n-mirror. This is due to the fact that the VCSEL was designed for top emission. The nonuniformity of the maximum optical-output power also shows that the thermal properties of the bonded VCSEL are not the same as shown in the experimental results in the next section. The reason may be that the bonding process did not result in good contacts on all of the bonding pads due to the tilt, as show in Fig. 3.14. Since the volume of InSn used is very small and no flux has been used during the heating, the InSn probably did not fill the air gap during heating. An air gap at the band increases the thermal resistance, which changes the threshold current. The nonuniform thermal resistance across the array contributes to the threshold-current variation.

Figure 3.15 shows the I-V and I-L curves of the VCSEL array top-bottom contact bonded to a GaAs substrate. The threshold current is about 2 mA and shows better uniformity than the coplanar flip-chip bonded VCSEL array. The optical-output power also shows good uniformity. This indicates that the bonded VCSELs have uniform thermal properties and suggests that it is easier to make better con-

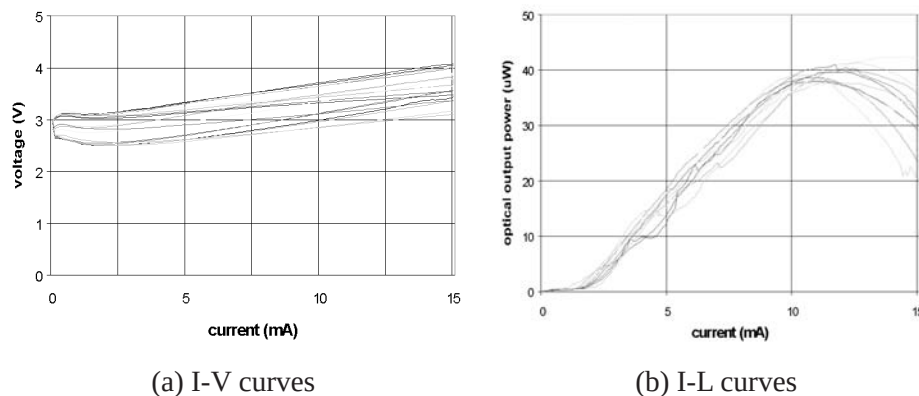


Figure 3.15. Characteristics for the VCSEL array top-bottom contact boded to the GaAs substrate.

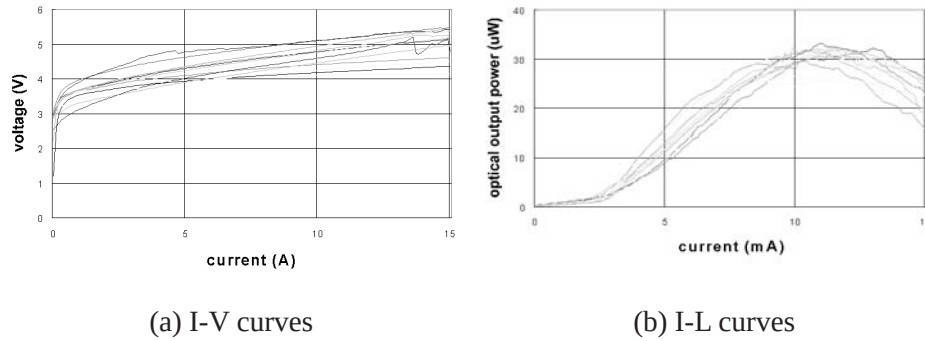


Figure 3.16. Characteristics for the VCSEL array top contact bonded to the GaAs substrate.

tacts with top-bottom contact bonding than with coplanar flip-chip bonding, at least using this equipment.

The optical output power of top-bottom-bonded VCSELs is smaller than coplanar flip-chip bonded VCSELs, because the top metal contacts may block some of the output light. The metal aperture size is only $16\text{ }\mu\text{m}$, which is similar to the oxide-confined aperture, thus if there is any misalignment the top metal contacts may block some of the output light.

The series resistance is approximately $70\text{ }\Omega$, which is smaller than coplanar flip-chip bonding, because top-bottom-contact bonding does not require intercavity contacts.

Figure 3.16 shows the I-V and I-L curves of the VCSEL array top contact bonded to a GaAs substrate. The threshold current is about 2.5 mA. The threshold voltages vary from 3.5 to 4.5 volts. The relatively high voltage is due to the lack of intercavity contact layers for both p- and n-contact. The series resistance is approximately $120\text{ }\Omega$, which is also larger than coplanar flip-chip bonding.

3.5.3 SMART-PIXEL FUNCTIONALITY

The results discussed in the previous two sections were obtained from VCSELs bonded to patterned GaAs substrates to evaluate the quality of the VCSELs after bonding. This section presents results and discusses issues for VCSELs bonded to foundry-fabricated Si CMOS and GaAs MESFET integrated circuit chips.

An 8×8 -VCSEL array was coplanar flip-chip bonded to a foundry fabricated MESFET chip. The chip performs the selection function of a data filter. The chip was fabricated by Vitesse through the MOSIS foundry service. The logic function performed by this smart pixel is shown in the block diagram of Fig. 3.17. Following the truth table (Table 3.2), MSM C was illuminated with a fiber. A and B were off as shown in Fig. 3.18; the bright area above the VCSEL is the reflection from the fiber providing the optical input. As the input optical power increases, the output VCSEL

A	B	C	F
0	0	1	1
1	1	1	1

Table 3.2. Truth table of the smart pixel for a data-filter system.

turned on from below threshold [Fig. 3.18(a)] to above threshold [Fig. 3.18(b)]. The chip was subsequently operated up to 200 MHz. This simple test demonstrates that the Vitesse chip was not significantly affected by the bonding process.

A 4×4 -VCSEL array was also coplanar flip-chip bonded to a CMOS chip to form a transmitter array for an optical ATM switch. The CMOS chip was fabricated by HP through the MOSIS foundry service. Figure 3.19 shows the 4×4 transmitter array with nine of the VCSELs turned partially on. Preliminary high-speed measurement shows that the transmitter has ~ 0.7 ns rise time and ~ 1.5 ns fall time approximately independent of the VCSEL bias voltage, as shown in Fig. 3.20. The longer fall time is probably due to the long discharge time of the photodetector used in this measurement or the reflection due to the impedance mismatch.

3.6 THERMAL CHARACTERISTICS

The bonding pads for the integration not only provide the electrical connection between the electronic chip and the VCSELs, but also provide the heat path for the VCSELs, which might be the most power-consuming devices in the circuits. Thus, the quality of the bonding affects the maximum optical-output power and efficiency of VCSELs as well as the scalability and density of the integrated arrays [14, 27, 28]. Thermal resistance and thermal cross talk are indicators of the thermal properties of the integrated structure. In this section, the results of thermal resistance and cross-talk measurements on bonded VCSELs are reported.

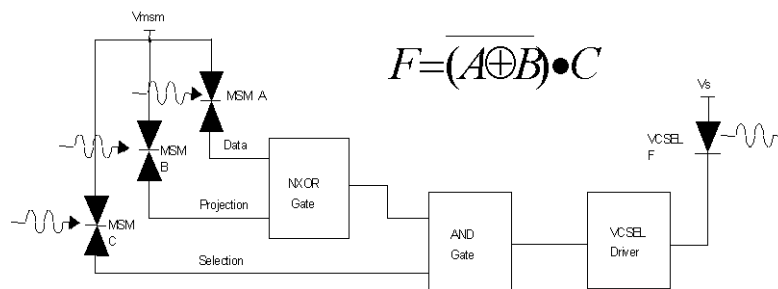
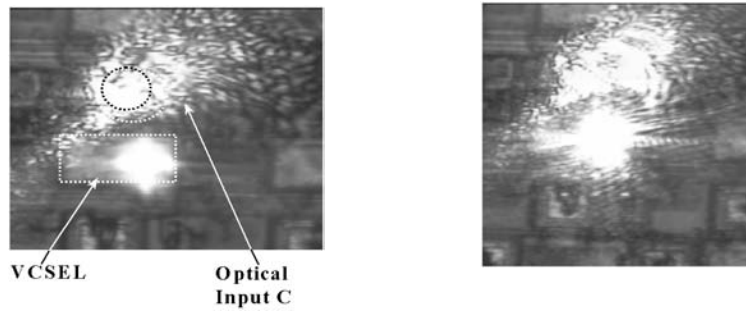


Figure 3.17. Logic function of the smart pixel for a data-filter system.



(a) VCSEL was below threshold

(b) VCSEL was above threshold

Figure 3.18. CCD camera image of the output of a VCSEL-MESFET smart pixel with the pixel photo-activated [16].

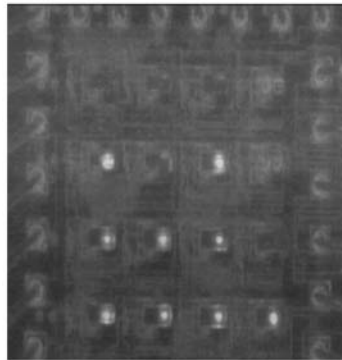


Figure 3.19. CCD camera image of the 4×4 VCSEL-CMOS smart pixel array with some of the VCSELs turned on [41].

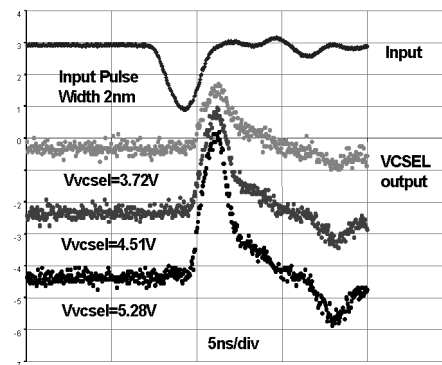


Figure 3.20. High speed measurement of the VCSEL-CMOS pixel.

3.6.1 THERMAL RESISTANCE AND CROSS-TALK MEASUREMENT

The thermal resistance R_{th} of a device is generally defined as the temperature rise ΔT in the device with the change of power dissipation, ΔP :

$$R_{th} = \Delta T / \Delta P \quad (3.3)$$

Since the index of refraction changes with temperature, the longitudinal-mode spectrum of the VCSEL will also change. By first measuring the wavelength change with temperature $\Delta\lambda/\Delta T$, and then measuring the effect that dissipated power has on wavelength, $\Delta\lambda/\Delta P$, the temperature change can be correlated to dissipated electrical input power change [24].

The thermal cross talk R_{12} can be described as the temperature change of device 1, ΔT_1 with the power dissipation change of device 2, ΔP_2 :

$$R_{12} = \Delta T_1 / \Delta P_2 \quad (3.4)$$

The thermal cross talk can be measured by driving VCSEL1 at constant input power and varying the electrical input power of its neighbor [30].

In order to isolate the thermal resistance of the bonding from the thermal resistance of the CMOS or MESFET chip, the thermal resistance of VCSELs bonded to a GaAs dummy substrate have been measured. The results show that if the bonding process was not properly done, the thermal resistance can be increased significantly. The thermal resistance of a VCSEL bonded to Si CMOS or GaAs MESFET chips was then measured. Since the thermal cross talk is very small, the measurement is not very accurate. However, it still gives a general idea about the range of the thermal cross talk.

All of the VCSELs used for bonding were made from the same VCSEL wafer and have approximately 15- μm apertures. The thicknesses of the p-mirror and n-mirror are 3.5 μm and 5.0 μm respectively. Normally, if the VCSELs have similar I-V-L characteristics, the VCSELs have similar thermal properties, because the thermal properties influence the maximum optical-output power and efficiency of VCSELs. Therefore, although the measurements only show one typical device on each bonded array, similar thermal properties can be expected for the devices on the same array.

The results of measurement on a VCSEL coplanar flip-chip bonded to GaAs is shown in Fig. 3.21. The wavelength of the fundamental mode is measured as a function of temperature at 4 mA drive current [Fig. 3.21(a)], which is 2 mA above the threshold. The temperature of the VCSEL was controlled by a thermoelectric cooler. This results in a straight line with a slope (wavelength shift) of 1.918 K/Å. The dissipated power of the VCSEL is taken as the input-electrical power minus the output-optical power. The wavelength shift dependence on power dissipation

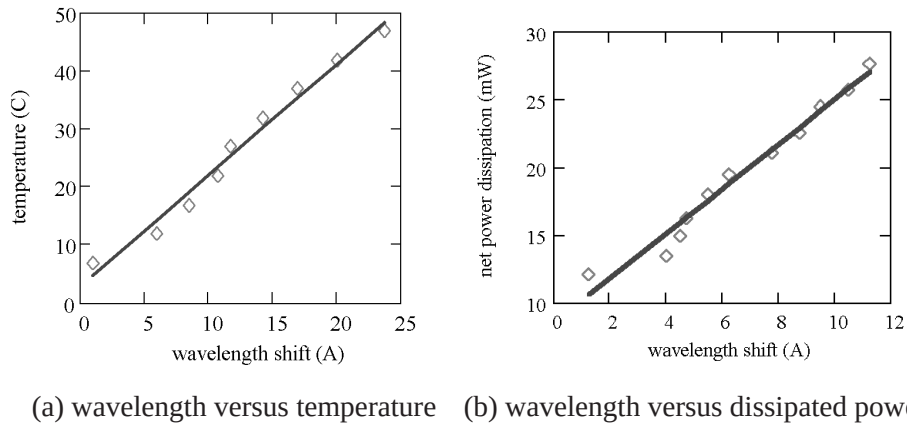


Figure 3.21. Thermal-resistance measurement of the coplanar A VCSEL properly coplanar flip-chip bonded to GaAs substrate.

is $1.650 \text{ mW}/\text{\AA}$ [Fig. 3.21(b)]. By dividing the temperature dependence by the power dependence, we obtain 1100 K/W for the thermal resistance.

Table 3.3 summarizes the measured thermal resistance of the VCSELs bonded to GaAs substrates. They are similar to most of the reported measurements of the VCSEL devices or bonded VCSELs. Matsuo and his coworkers at NTT developed

Author	Aperture (μm)	Bonding technique	thermal resistance (K/W)
Present work	15	coplanar flip-chip	1100
	15	top-bottom	1230
	15	top-contact with epoxy	1240
NTT [9]	15	polyimide (0.3–1.4 μm) bonded	1600–2900
UC Berkley [33]	5	Unbonded oxide confined VCSEL	1200–1500
UCSB [30]	4–12	Unbonded air post VCSEL	4900–2000
USC [29]	7	Unbonded oxide confined VCSEL	2500
Sandia [14]	7	Unbonded proton- implant and oxide confined VCSEL	2500

Table 3.3. Thermal resistance of the bonded and standard VCSELs

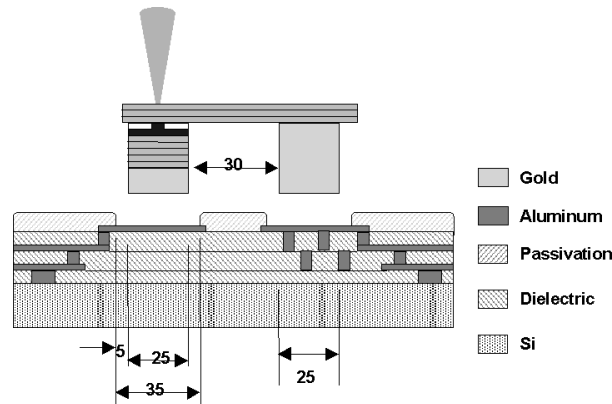


Figure 3.22. Schematic drawing of a VCSEL flip-chip bonded to a CMOS chip.

a bonding technique using polyimide to attach the VCSELs onto the Si substrate [8, 9]. The thermal resistance of the VCSEL bonded with 0.25- μm polyimide is 1600 K/W. Thermal resistance of 1200–1500 K/W for oxide-confined VCSELs was also reported by Lascola and his coworkers at UC Berkeley [31]. The thermal resistance of the unbonded VCSEL mainly depends on the mirror layers underneath the active region. Different VCSEL mirror designs can result in quite a different thermal resistance. Therefore, this comparison only shows a reasonable range of the thermal resistance of a working VCSEL.

In addition to the investigation of VCSELs bonded to dummy GaAs chips discussed previously, VCSELs coplanar flip-chip bonded to the CMOS chip have also been measured and compared to VCSELs bonded to GaAs dummy substrates. In fact, since coplanar flip-chip and the top-bottom bonding have similar thermal transfer paths, one can also predict the thermal properties of VCSELs top-bottom bonded to the electronic chips by these measurements.

The thermal resistance of the VCSELs coplanar flip-chip bonded to the CMOS is as high as 2490 K/W. Compared to 1100-K/W thermal resistance of the VCSEL bonded to the GaAs dummy substrate, the high thermal resistance of the VCSEL bonded to the Si CMOS chip is caused by the dielectric layers underneath the bonding pads, as shown in Fig. 3.22. Standard CMOS pads were used on the chips. There are three dielectric layers between metal layer 3, 2, 1 and the Si substrate. For lower thermal resistance, as many Al vias as the design rules allow can be put directly underneath the bonding pads (as shown underneath the right bonding pad). The number of the vias is also limited by the available room at each metal bonding pad.

The thermal cross talk mainly depends on the pitch of the VCSEL arrays. The VCSEL arrays bonded by all three techniques have a 250- μm pitch. Therefore, the

thermal cross talk should be the same for all three and was only measured between coplanar flip-chip-bonded VCSELs.

With the thermal cross talk measurement, the total wavelength shift with the power dissipation of the nearest neighbor VCSEL is as small as 0.5 Å. The information obtained from this measurement is that an array with 250-μm pitch has less than 100-K/W thermal cross talk. This number corresponds to a 1-K temperature rise of the laser when its nearest neighbor laser operates with 10-mW dissipated power.

3.6.2 SIMULATION RESULTS OF STATIC THERMAL TRANSFER

The static thermal transfer was simulated with finite element analysis software (FEHT). The heat was considered to be generated only in the active region. Since the series resistance of the VCSELs bonded by these three techniques does not show a big difference between 70 and 120 Ω, ignoring the heat generated by the series resistance will not significantly affect the following comparison. Furthermore, after the processing is optimized, the series resistance will be reduced. The series resistance of an unbonded VCSEL can be as low as 30 Ω. The surface radiation and convection are also ignored since they only contribute a very small portion of the heat loss, less than 1% [23].

Although the model is very simple, what is of interest is not the exact value of the simulation, but the trends. In this case, the model is good enough for the analysis.

The thermal conductivities of the materials used in this model are listed in Table 3.4. These values differ from the bulk-material thermal conductivity, since experimental results have shown that the conductivity of the DBR mirror is lower than the thermal conductivity of the same average composition [30]. Because the layers of the DBR mirrors are only a few hundred nanometers thick, interface scattering contributes a large portion of the thermal transfer mechanism. The ther-

Material	Thermal conductivity (W K ⁻¹ m ⁻¹)	Reference
Au	315	[31]
Al _{0.16} Ga _{0.84} As/ Al _{0.92} Ga _{0.08} As mirror	25	[30]
GaAs	45	[31]
polyimide and epoxy	0.2	[32]
Si	141	[31]
SiO ₂	1.02	[31]

Table 4. Thermal conductivity of the materials.

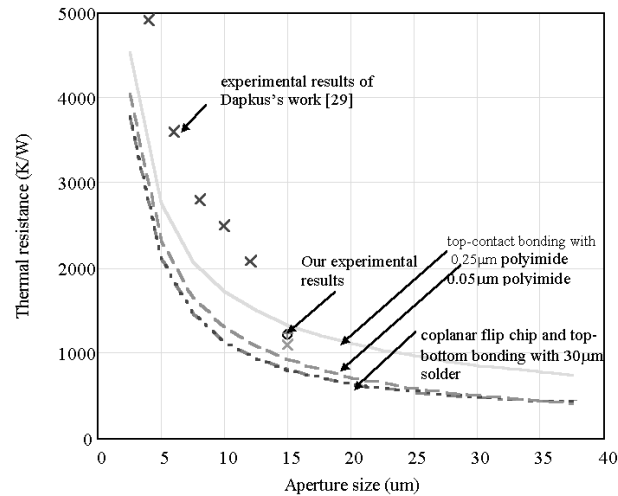


Figure 3.23. Thermal resistance of bonded VCSEL versus aperture size.

mal conductivity of the polyimide and epoxy are the same. Although we only use the polyimide in the following simulation, the results will be approximately the same with the epoxy.

In the measurements presented in the previous sections, the temperature of the VCSEL was controlled by a temperature controller through a thermoelectric cooler, which maintained the bottom of the substrate at a constant temperature. The same condition was used in the simulation. Heat was generated in the active region of the simulation model, and the temperature distribution in the VCSEL and the substrate were calculated. Using these data, the thermal resistance could be calculated using Eq. (3.3).

3.6.2.1 Thermal resistance versus aperture size

The thermal resistance of the VCSELs bonded to GaAs substrate vs. aperture size was calculated, as shown in Fig. 3.23. The calculations were made with a solder size of 30 μm. A polyimide thickness of 0.25 μm was used with the top-contact-bonded VCSEL. The measured thermal resistance of VCSELs bonded onto the GaAs substrate by the three techniques is also shown in Fig. 3.23.

The simulation results indicate that the coplanar flip-chip and top-contact-bonded VCSELs have almost exactly the same thermal properties. This is because the Au post on the n-contact only contributes a small portion to the heat flow. The thermal resistance of the top-contact-bonded VCSEL approaches that of the other two techniques when the thickness of the polyimide is reduced to 0.05 μm. The measured thermal resistance of the top-contact-bonded VCSEL also shows similar values to that of the VCSELs bonded with the other two techniques.

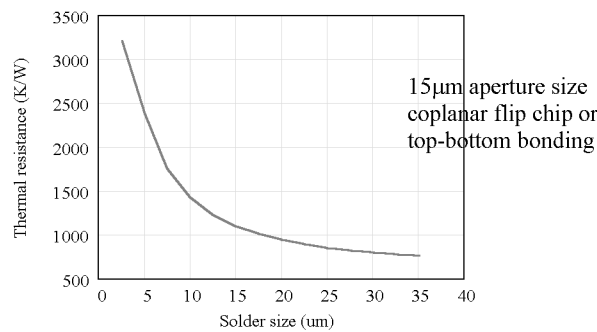


Figure 3.24. Thermal resistance versus the solder size.

The measured thermal resistance is in the same range of the simulation results. However, the measured thermal resistance of the coplanar flip-chip and top-bottom-bonded VCSELs are higher than the simulation results. The reason may be due to the value of the thermal conductivity of the VCSEL mirror used in the simulation, since the VCSEL used in this research does not have exactly the same composition as that in reference [30].

Since the simulation results show little thermal-resistance variation with the aperture size around 15 μm (particularly for larger sizes), variation in the VCSEL fabrication will not result in significant thermal resistance variation and is probably not the cause of the difference in the measured and simulated results.

3.6.2.2 The limit of bonding density

Since the thermal properties of the VCSELs bonded with the three techniques are similar, the simulation in this section is based on coplanar flip-chip-bonded VCSELs. In order to accomplish higher bonding density, smaller solder-bonding-pad sizes are desirable. Figure 3.24 shows the thermal resistance versus the solder size. When solder size is smaller than 10 μm , the thermal resistance of the VCSEL increases very rapidly.

The thermal cross talk versus pitch was also calculated, as shown in Fig. 3.25. Considering the size of the oxide-confined VCSEL, the pitch can not be smaller than 40 μm . The thermal cross talk is already less than 100 K/W when the pitch is larger than 40 μm , according to the simulation. Therefore, the thermal cross talk does not limit the VCSEL array density.

3.6.2.3 Thermal resistance reduction with top-contact-bonded VCSEL

When the VCSEL is bonded to a CMOS chip with the top-contact technique, the VCSEL is attached to the surface of the CMOS chip with polyimide or epoxy. Since there is an approximately 1- μm -thick passivation layer on the surface of the

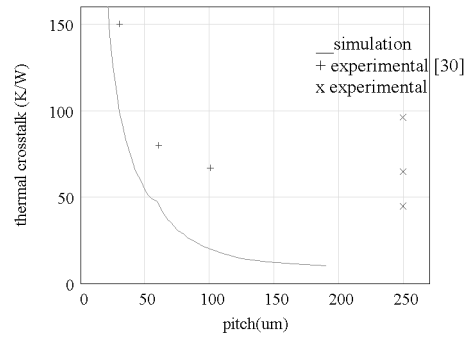


Figure 3.25. Thermal cross talk versus the pitch of VCSEL array.

CMOS chip, the thermal resistance can not be reduced by reducing the thickness of the polyimide/epoxy.

However, the thermal resistance of top-contact-bonded VCSELs can be reduced by increasing the thickness of the top gold trace. Figure 3.26 shows the thermal resistance versus the thickness of the gold trace. It is assumed that:

- the polyimide layer is so thin that it can be ignored,
- the passivation layer is 1 μm , and
- the aperture of the VCSEL is 15 μm .

The simulation indicates that the thermal resistance can be reduced 30% by increasing the thickness of the gold traces from 0.5 to 5 μm . The thermal resistance can be further reduced by increasing the width of the gold traces.

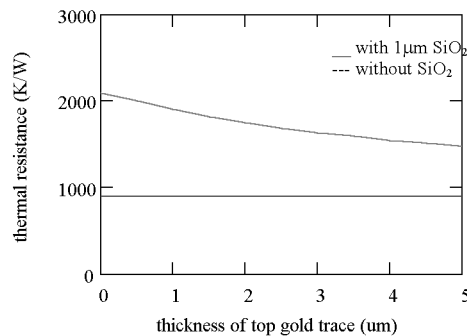


Figure 3.26. Thermal resistance versus the thickness of the top gold trace in top-contact bonded VCSELs.

3.6.2.4 VCSELs on CMOS

The thermal resistance of the VCSEL coplanar flip-chip bonded to a CMOS chip was also calculated with the same parameters as the fabricated device, e.g., 15- μm aperture and 30- μm solder. The dielectric layer underneath the bonding pads (Fig. 3.22) increases the thermal resistance to as high as 2150 K/W. This compares favorably with the measured thermal resistance of 2450 K/W.

In order to reduce the thermal resistance, as many vias can be placed underneath the bonding pads as the design rules allow. In the best case, which means there is enough room on metal layer two and one, the thermal resistance can be reduced by about 1000 K/W to 1165 K/W, which is approximately the resistance of VCSELs bonded to a dummy GaAs substrate.

3.7 COMPARISON OF THESE RESULTS WITH OTHER METHODS AND APPRAISAL OF THESE TECHNIQUES

An ideal hybrid-integration technique should have the ability to be scaled to large size with high yield and low cost. In addition, the integration-induced parasitics should be minimized and the high-quality performance of the discrete devices should be maintained after integration.

Three hybrid-integration techniques have been developed for the fabrication of VCSEL-based smart pixels. Since the VCSELs are attached directly on top of the pixel chip in all three techniques, the distance between the VCSEL contacts and the bonding pads of the pixel chip is very short. Thus, very small parasitics can be expected and measurements have shown that all three techniques maintain the high quality of the VCSELs. However, techniques differ in other respects and this section provides a comparison of the three techniques, based on the processing and the electrical, optical, and thermal characteristics of the bonded VCSELs.

3.7.1 COMPARISON OF THE PROCESSING

Generally, the number of mask levels is taken as a measure of the complexity and yield of a process. In addition, there are critical processing steps which effect bonding yield.

The coplanar flip-chip technique requires the most mask levels (seven mask levels compared with five) and had the lowest yield. However, the use of a wafer-bonding machine should greatly improve the yield. In addition, the coplanar flip-chip-bonding technique fully processes the VCSEL array before bonding and is therefore the most compatible technique with foundry fabrication. This advantage makes it the most promising technique of the three investigated; however, the top-contact technique bonds the VCSEL flat on the CMOS and is therefore the most robust.

Structure	Coplanar flip-chip	Top-bottom contact	Top-contact
Required accuracy of attachment parallelization	Critical	critical	not as critical
Height tolerance of electroplated Au	0.3 μm	1 μm	N/A
Metal step coverage	N/A	not as critical	critical
Critical photolithography	many mask levels	metal aperture	metal aperture
Number of mask levels	7	5	5
Sequence of metal deposition	after oxidation	before oxidation	before oxidation
Sequence of attachment	after oxidation	after oxidation	before oxidation

Table 3.5. Comparison of the important processing steps for the three VCSEL-bonding techniques.

All of the bonding techniques require many processing steps. However, the majority of the processing steps are routine, e.g., rinsing, plasma ashing, dry etching, metal deposition, etc. These are common to all three of the techniques and do not provide a basis for differentiating the difficulty and/or scalability of the processes. The more critical processing steps and processing characteristics are listed in Table 3.5.

The difficult processing steps that may limit the yield of the smart-pixel arrays are different for each of the three techniques. For coplanar flip-chip bonding, besides the low solder-bonding yield due to tilt, the difficulties include:

- (1) Many processing steps carried out on a nonplanar surface topography. Although the other two techniques also require processing steps carried out on a nonplanar surface topography, the number of the mask levels is less than with coplanar flip-chip bonding.
- (2) High resistance metal-semiconductor contacts. The present VCSEL wafer does not have an intracavity contact layer for the metal contact. The metal contact is deposited on the n-mirror layers. The composition of Al in the mirror layers is very easily oxidized and the oxide is hard to remove.

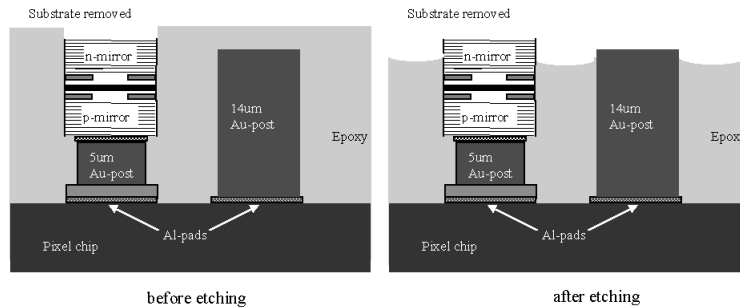


Figure 3.27. Epoxy etch with top-bottom-bonding technique.

- (3) The coplanar flip-chip bonding technique requires electroplating the n-mirror contact to the same level as the p-mirror contact. Although the required accuracy is within $0.3\ \mu\text{m}$, this can be obtained by careful control.

The solder-bonding yield due to tilt is the most significant limit on the total yield with the top-bottom-bonding technique, but this can be overcome by using a solder-bonding machine. The other problems associated with top-bottom-bonding technique are:

- (1) A deposit metal aperture is required as part of the top contact. This is the smallest feature among the bonding structures and since this step occurs after bonding and substrate removal, the surface is a combination of VCSEL mesas, Au posts and epoxy. Tilt also adds to the problem. Furthermore, the dies are small, and thus it is hard to perform edge-photoresist removal. This would not be the case in a commercial process. The nonplanar surface makes the contact between the surface and the mask poor. Thus, the patterned metal aperture is smaller than the mask.
- (2) The epoxy serves as a bed upon which the contact metal sets up a bridge between the Au post and the VCSEL mesa. However, the oxygen plasma etching of the remaining epoxy is not uniform as shown in Fig. 3.27.

Top-top contact bonding is different from the other two bonding techniques since it does not require solder bonding, thus there is no problem with tilt. However, there are still problem steps.

The most critical processing step is patterning the insulation layer for the step coverage of the metal traces for the n-mirror contact. Polyimide was chosen to serve as the insulation layer, since a few-micrometer-thick layer can be obtained with polyimide. The thicker layer not only provides better insulation, but also reduces the parasitic capacitance. In addition, polyimide forms 45°

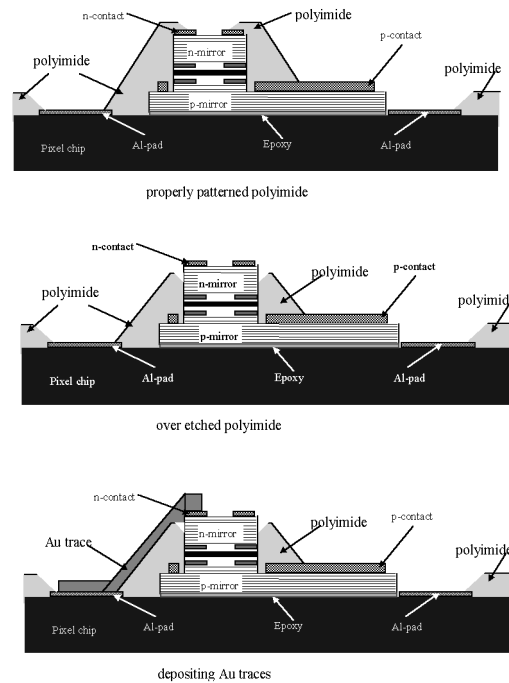


Figure 3.28. Insulate layer for top-top contact-bonding technique.

sidewall after patterning, which makes the following metal step coverage easier. However, patterning the polyimide on a nonplanar surface topography is difficult since the polyimide tends to smooth the nonplanar surface. Therefore, the thickness of the polyimide is not uniform and the etching of the polyimide is also not uniform. The overetched polyimide layer cannot provide a good insulation layer for the following metal step coverage, as shown in Fig. 3.28; however, a more suitable polyimide and better process conditions might be found with more experience. This problem can also be solved by depositing SiO_2 instead of polyimide.

- (3) The top-contact technique requires attaching the VCSEL wafer to the pixel chip with a nonconductive adhesive such as epoxy or polyimide. After removal of the VCSEL wafer substrate and the formation of the VCSEL mesa, the bonded VCSEL and pixel chip must be exposed to steam at 430°C for 25 min. As a result, the bonding pads are also oxidized, since the thin nonconductive adhesive layer does not provide sufficient protection. The oxide on the surface of the bonding pad affects the resistance and yield of the subsequent metal contacts. Therefore, better protection before the oxidation or treatment after the oxidation needs to be designed to overcome this problem.

- (4) Top-contact-bonding technique also requires depositing a metal aperture on a nonplanar surface topology, similar to the top-bottom technique.

Most of the limits on the fabrication yield for all three bonding techniques can be overcome by using suitable equipment or optimized design. Therefore, much better yield can be expected from a commercial fabrication facility. Although fabrication was based on 4×4 and 8×8 arrays, these techniques should be scalable to 32×32 or more.

3.7.2 COMPARISON OF ELECTRICAL AND OPTICAL CHARACTERISTICS

All three bonding techniques resulted in working VCSELs with reasonable electrical and optical characteristics. This indicates that all of these techniques are possible candidates for large-scale hybrid integration of VCSELs. In an attempt to obtain a more meaningful comparison of the bonding technique, the characteristics listed in Table 3.6 were measured from bonded VCSELs fabricated from the same VCSEL wafer, EMC1148, and with similar aperture size.

The most important electrical characteristics are the series resistance and threshold voltage. These depend on:

- the metal-semiconductor contacts of the p- and n-contacts of the VCSEL,
- the series resistance of the p- and n-mirror of the VCSEL,
- the bonding metal contacts,
- the band gap of the contacted semiconductor,
- oxides on the semiconductor.

Since there are no intracavity contact layers in the VCSEL wafer for the semiconductor-metal contacts, the changing series resistance and threshold voltage VCSELs from run to run reflects the quality of the semiconductor-metal contact rather than the bonding contact.

The VCSEL bonded with the top-contact technique has the highest series resistance and threshold voltage among the three bonding structures. This may be caused by the p-contact on the Al-rich AlGaAs layer of the p-mirror. The n-contact could also be on the AlAs etch-stop layer. The coplanar flip-chip-bonded VCSEL has a smaller series and threshold voltage resistance because it avoids the AlAs layer. The top-bottom-bonded VCSEL has the smallest series resistance and threshold voltage among the three bonding structures, since it does not require intracavity contacts for either the n- or p-mirror.

Group	Resistance (Ω)	Threshold current (mA)	Threshold Voltage (V)	Aperture (μm)	Optical output power (mW)
Presented works	1632 coplanar	3.8	2.1	14	3.0
	1584 top-bottom	0.8	1.5	6	0.07
	1148 coplanar	2.0	2.8	~ 15	0.25
	1148 top-bottom	2.0	2.5–3.1	~ 15	0.04
	1148 top-contact	2.5	3.5–4.5	~ 15	0.03
Kishnamoorthy[17] Lucent	150	0.9–1.0	1.4–1.5	N/A	5.0
Matsuo[9] NTT	70	3.5	2.3	15	1.2
Maracas[21] Arizona State	90	5.5–6.4	3.0	40	1.0
McLaren[34] Univerity of Colorado	200	1.5	1.9		> 0.4
Smith[11] UC Berkeley	70	30	17.0	40	0.03
Dapkus [35] USC	60	0.5	1.3	5	6
Morgan[36] Honeywell (proton-implanted)	29	2.7	1.55		8.0
Choquette[37] Sandia (oxide-confined)	32	4.0	2.0	14	5.0
Honeywell product HFE 4080-321	20	3.8	1.5		5.0

Table 3.6. Comparison of the electrical and optical characteristics of the three bonding techniques with those of other researchers

As mentioned previously, the threshold current of the bonded VCSEL depends on the mirror reflectivity and each of the bonding techniques results in different mirror reflectivity. Calculations indicate that the coplanar flip-chip and top-bottom-contact-bonded VCSELs should have lower threshold than the top-contact-bonded VCSEL.

The optical-output power of the VCSELs is low because the VCSEL wafer (wafer EMC1148), was not designed for back emitting, since its n-mirror has many DBR pairs (39), which results in very high reflectivity and less optical output power.

The top-bottom and top-contact-bonded VCSELs have even smaller optical-output power, because the metal aperture blocked part of the optical output. The difficulty of processing the metal aperture was discussed in the last section. Table 3.6 also lists published values for bonded VCSELs from other research groups, data from a Honeywell commercial product, and VCSELs from standard unbonded-VCSEL arrays. The parameters for the presented VCSELs are in the same range as reported for bonded VCSELs by other groups.

It is also observed that the series resistance of all of the bonded VCSELs listed in the table is higher (50–200 Ω) than the standard/commercial VCSELs (20–32 Ω) and the threshold voltage of the bonded VCSELs is also usually higher. There are several possible causes for the higher values for the bonded VCSELs:

- the VCSEL wafers for flip-chip bonding were not properly designed to reduce the contact/mirror resistance,
- the ohmic contacts were not annealed properly,
- an oxide exists on either the bonding metal pads or on the AlGaAs.

This is an area of research of all types for flip-chip bonding of VCSELs.

3.7.3 COMPARISON OF THE THERMAL PROPERTIES OF THE VCSEL-BASED SMART PIXEL

Since the VCSELs are bonded as individual devices, the heat generated by each VCSEL must be removed through the attachment to the substrate. Therefore, the thermal properties of the bonded VCSEL indicate the quality of the bonding.

The thermal transfer model of Section 3.5.2 shows that coplanar flip-chip and top-bottom-bonding techniques have similar thermal properties. With both bonding techniques, the VCSELs are attached to the bonding pads on the pixel chip with a Au post. Therefore, the heat transfer of the bonded VCSEL are limited by:

- (1) the size of the Au post,
- (2) the quality of the solder bonding,

(3) the thermal property of the bonding pads on the pixel chip.

According to the simulation, when the size of the Au post is $< 10 \mu\text{m}$, the thermal resistance of the bonded VCSEL increases rapidly. The experimental results demonstrate that when the bonding is not properly done, the thermal resistance can be as high as 2250 K/W, while the thermal resistance of a properly bonded VCSEL is only 1100 K/W.

The material underneath the bonding pads on the pixel chip also influences the thermal resistance of the bonded VCSEL, e.g., when the VCSEL was bonded to a CMOS chip. In a standard CMOS bonding-pad design, there are three or more dielectric layers underneath the bonding pads; this causes the thermal resistance to increase to 2450 K/W. On the other hand, the thermal resistance of a VCSEL bonded to a GaAs substrate is only 1100 K/W. The thermal resistance of the VCSEL bonded to a CMOS chip can be reduced by putting many Al vias underneath the bonding pads.

With the top-contact-bonding technique, the VCSELs are attached to the surface of the pixel chip with nonconductive adhesive, such as epoxy and polyimide. The electrical connection between the VCSEL and the pixel is provided by the gold traces on the top of the VCSEL. Therefore, the heat-transfer properties of the top-contact-bonded VCSEL are limited by:

- the thickness of the nonconductive adhesive layer,
- the uniformity of the nonconductive adhesive layer,
- the thickness of the passivation layer on the surface of the pixel chip, and
- the thickness of the top gold traces.

In order to obtain low thermal resistance, the thickness of the nonconductive adhesive layer should be as small as possible. Very thin layers can be obtained with epoxy because of its very low viscosity. The measured thermal resistance of the VCSEL bonded with epoxy, which is 1250 K/W, is close to the thermal resistance of the VCSEL bonded with the other two techniques. Thus, the epoxy layer does not significantly increase the thermal resistance. However, nonuniformity of the epoxy or polyimide layer will affect the uniformity of the thermal resistance of the bonded VCSEL, especially when the layer becomes thick.

There is a potential thermal-resistance problem with top-top contact technique applied to a foundry-fabricated chip. When the VCSEL is attached on the surface of the microelectronic chip, either CMOS or MESFET, the surface has a passivation layer which is not a good thermal conductor. This will increase the thermal resistance. If a window is opened in the passivation layer underneath the VCSEL,

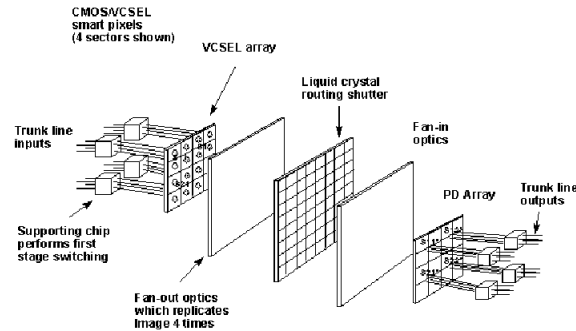


Figure 3.29. Conceptual block diagram [38].

the window will be filled up with the epoxy or polyimide when it is applied. Therefore, this problem cannot be avoided in this method.

When the heat path underneath the VCSEL is blocked by the passivation layer, the gold traces on the top of the VCSEL, which connect the VCSEL to the bonding pads on the microelectronic chip, become important. The thermal resistance can be reduced by increasing the thickness and width of the gold traces.

3.8 APPLICATIONS

The bottlenecks encountered in cross-point communication routing for asynchronous-transfer-mode (ATM) and internet-protocol (IP) systems can be greatly reduced with free-space processors using VCSEL-based smart pixels. Two versions of such a routing switch are outlined here.

3.8.1 VCSEL- AND FLC-SLM-BASED OPTOELECTRONIC ATM SWITCH

A conceptual block diagram of the optoelectronic ATM switch is illustrated in Fig. 3.29. This approach can be mapped identically to a Clos three-stage switching array [38].

To explain how the system operates, a small switch with 16 inputs and 16 outputs can be used as an example [39]. Spatially, the 16 electrical input channels are grouped into four sectors with four channels in each sector, as shown in Fig. 3.30(a). The first stage switching is carried out by four 4×4 switches (represented by four blocks at the input in Fig. 3.29), which rearranges all the channels to reduce or eliminate the collision that may occur in the subsequent stages. The 16 channels, after being rearranged by the first stage, drive a 4×4 -VCSEL array. The outputs from the VCSEL array are therefore in a format of an array of 4×4 light beams. Each light beam is normal to the VCSEL-array plane. The array can be considered as four sectors with four channels in each sector.

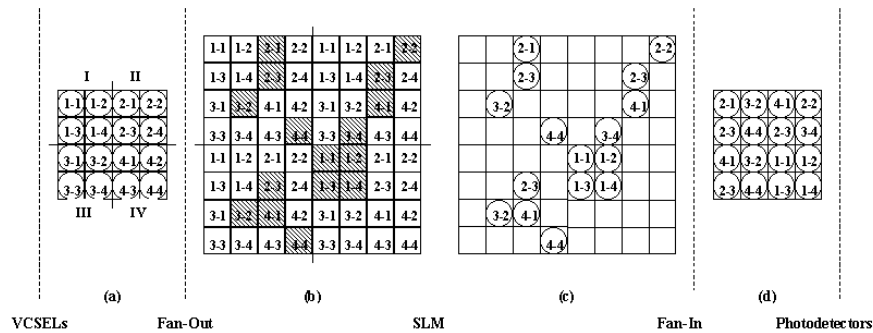


Figure 3.30. Patterns at different stages.

The 4×4 light-beam array from VCSEL passes through the fan-out optics, which replicates the 4×4 beam array four times. The output pattern after fan out can be either an 8×8 or a 4×16 array depending on the choice of fan-in optics. Both can be considered to consist of 16 sectors. Figure 3.30(b) shows an 8×8 array. A liquid-crystal spatial-light modulator (SLM) is placed after the fan-out optics and functions as the cross-point routing shutter. The SLM allows one beam out of each of 16 sectors. Each shaded square in Fig. 3.30(b) indicates that the cross point is closed and the channel in the area is passed to the next stage. The other light beams are blocked by the shutter. Figure 3.30(c) shows the outputs from the FLC shutter.

The fan-in optics then map these 16 beams back into the 4×4 configuration. The output plane of fan in is a 2D array consisting of four sectors. Each sector is a 2×2 array, as illustrated in Fig. 3.30. There are a total of 16 output channels, represented by a 4×4 light beam array. These 16 optical channels are converted into the electrical domain by a 4×4 photodetector array.

3.8.2 VCSEL- AND SPA-BASED OPTOELECTRONIC ATM SWITCH

One of the major speed and scalability limitations of the above SLM-based demo system is the liquid crystal used as the SLM shutter, which prevents the overall system speed from reaching the speed of modern networks even with data stream demultiplexing. The system also requires fan-in optics that limit the performance and the scalability of the system. Moreover, the supporting electronic circuit for the SLM control is complicated and slow.

These difficulties inspired another solution, which is easier to implement and has higher performance [40]. By studying all the problems present in the SLM-based system, it was decided to replace the SLM and the fan-in optics. The block diagram of the new system is shown in Fig. 3.31. It replaces the slow SLM and the difficult fan-in optics with a smart-pixel circuit which is referred to as a smart-photodetector array (SPA).

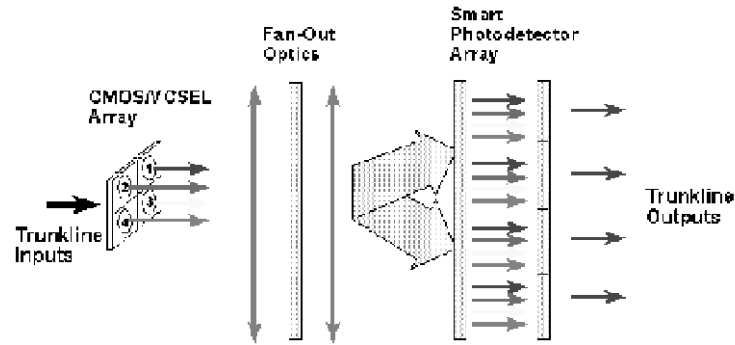


Figure 3.31. Block diagram of the SPA-based system.

It can be seen that the architecture of the new system is identical to the SLM-based system. The most significant change in the new system is that the selection and fan in are performed by an SPA in the electrical domain instead of by SLM and fan-in optics in the optical domain.

Figure 3.32(a) illustrates the SLM selection and fan-in optics in the SLM-based system. The SLM functions as a switch on each optical channel. At any time, only one switch is closed. Following the SLM are the fan-in optics that simply merge multiple inputs into one output. However, a component with identical functionality can be easily implemented in the electrical domain. Figure 3.32(b) shows an electronic digital selector that performs select/fan in. This is a standard electronic digital circuit with a photodetector added in front of each input of the selector to enable the circuit to receive optical data. The output of the selector is converted back into a light signal by driving a laser diode.

The operation speed of the system increases greatly if the selector is realized with fast digital IC technology. Aberration exists in each of the two 4-f beam-relay systems and the degree of aberration accumulates when two 4-f systems are placed back-to-back. Using the SPA, the digital-optical signal is regenerated by

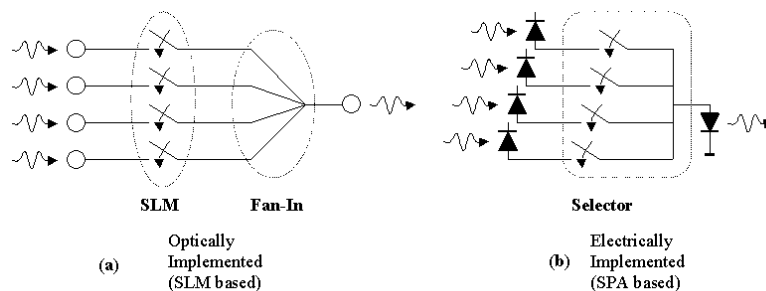


Figure 3.32. Two ways to implement select/fan in.

the SPA; therefore, aberrations occurring in the previous stages can be eliminated and the optical cross-talk performance of the whole system is greatly improved. Also optical power amplification is included in the signal-regeneration circuit, which compensates for the optical power loss in the first-stage optical link, and therefore leaves more room in the power budget. The SPA-based system inherits all the advantages of free-space optical interconnection while taking advantage of well-developed digital electronic IC technology. The SPA has the potential to be implemented in large scale and can operate at high data rate.

3.9 SUMMARY

There are a variety of techniques for bonding VCSELs to integrated circuits and a number of different techniques have been successfully demonstrated. Each of these techniques has advantages and disadvantages. This chapter investigates three techniques, all of which remove the substrate, resulting in individual VCSELs bonded to locations on the integrated circuit. This has the advantage of reducing the strain on the bonds which may be important in determining the lifetime of the bonded optoelectronic chips. Removal of the substrate also provides direct optical access to the integrated circuit and the photodetectors. For these bonding techniques, the following has been shown:

- (1) Bonding VCSELs to CMOS & MESFET circuits is viable. The proper functioning of the smart pixel with the bonded-VCSEL array verifies that the bonding process can maintain the high quality not only of the VCSEL, but also foundry-fabricated microelectronic integrated-circuit chips, such as GaAs MESFET and Si CMOS chips.
- (2) The measured series resistance and threshold voltage of the bonded VCSELs are comparable with previously reported bonding results, but are higher than for unbonded VCSELs. The series resistance should be reduced by adding an intracavity contact layer into the VCSEL wafer for the metal contact.
- (3) The measured thermal resistance of the VCSELs bonded to GaAs substrates is also comparable with unbonded VCSELs.
- (4) The measured thermal cross talk is not as low as expected for the 250- μm pitch, but is the same range as with 100- μm pitch. Therefore, thermal cross talk will not limit the density of the array down to a pitch of 100 μm .
- (5) The thermal transfer simulation results show that the size of the bonding pads may be determined more by the need to heat sink the VCSELs than other factors.

- (6) Partially attached bonding can provide good electrical connection but not thermal conduction; thus, the thermal properties of the bonded VCSELs can reflect the quality of the bonding better than electric characteristics.
- (7) Although bonding characteristics need further investigation in order to transition these techniques into practice, these techniques have the potential to make large optoelectronic arrays and maintain the high quality of the VCSELs.

ACKNOWLEDGEMENTS

This work was funded in part by National Science Foundation grant #ECS-9900493 and Colorado State University through the Program of Research and Scholarly Excellence. The authors also thank the reviewer for many helpful suggestions, and acknowledge the contribution of C. Duan to Section 3.8.

REFERENCES

- [1] Christopher Tocci, and H. John Caulfield, *Optical Interconnection, Foundations and Applications*, Artech House, Boston, London, 1994.
- [2] Y. S. Liu, "Lighting the way in computer design," *Circuits and devices*, pp. 23–31, Jan. 1998.
- [3] Yoshitaka Ohiso, Kouta Tateno, Yoshitaka Kohama, Atushi Wakatsuki, Hideki Tsunetsugu, and Takashi Kurokawa, "Flip-chip bonded 0.85- μ m bottom-emitting vertical-cavity laser array on an AlGaAs substrate," *IEEE Photon. Tech. Lett.*, vol. 8, pp. 1115–1117, 1996.
- [4] K. D. Choquette, B. Roberds, K. M. Geib, H. Q. Hou, R. D. Twستن, K. L. Lear and B. E. Hammons, "Bottom-Emitting 850 nm Selectively Oxidized VCSELs Fabricated Using Wafer Bonding," *LEOS Conf.*, San Francisco, CA, November, 1997.
- [5] *Chip on board technologies for multichip modules*, Edited by John. H. Lau, International Thomson Publishing company, 1994.
- [6] Ralph W. Woodgate, *"The handbook of machine soldering SMT and TH,"* A Wiley-Interscience publication, 1996.
- [7] K. W. Goossen, J. A. Walker, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. D. Bacon, D. Dahringer, L. M. F. Chirovsky, A. L. Lentine, and D. A. B. Miller, "GaAs MQW modulators integrated with silicon CMOS," *IEEE, Photon. Tech. Lett.*, vol. 7, pp. 360–362, 1995.
- [8] S. Matsuo, K. Tateno, T. Nakahara, H. Tsuda and T. Kurokawa, "Use of polyimide bonding for hybrid integration of a vertical cavity surface emitting laser on a silicon substrate," *Electron. Lett.*, vol. 22(12), pp. 1148–1149, 1997.
- [9] S. Matsuo, T. Nakahara, K. Tateno and T. Kurodawa, "Novel technology for

- hybrid integration of photonic and electronic circuits," IEEE, Photon. Tech. Lett., vol. 8, pp. 1507–1509, 1996.
- [10] Rui Pu, Chunjie Duan and C. W. Wilmsen, "Hybrid Integration of VCSELs to CMOS Integrated Circuits," IEEE, Journal of Selected Topics in Quantum Electronics, vol. 5, pp. 201–208, 1999.
 - [11] Hsi-Jen J. Yeh and S. Smith, "Integration of GaAs vertical-cavity surface emitting laser on Si by substrate removal," Appl. Phys. Lett. Vol. 64, pp. 1466–1468, 1994.
 - [12] Rui Pu, "Fabrication and characterization of VCSEL based smart pixel," Ph.D. thesis, Colorado State University, Aug. 1999.
 - [13] *Vertical Cavity Surface Emitting Lasers*, Edited by Carl Wilmsen, Henryk Temkin, and Larry A. Coldren, Cambridge University Press, 1999.
 - [14] K. L. Lear, K. D. Choquette, R. P. Schneider Jr., S. P. Kilcoyne and K. M. Geib, "Selectively oxidized vertical cavity surface emitting lasers with 50% power conversion efficiency," Electronics Letters, vol. 31, pp. 208–209, 1995.
 - [15] L. M. F. Chirovsky, W. S. Hobson, R. E. Leibenguth, S. P. Hui, J. Lopata, G. J. Zydzik, G. Giaretta, K. W. Goossen, J. D. Wynn, A. V. Krishnamoorthy, B. J. Tseng, J. M. Vanderbery and L. A. D'Asaro, "Inplant-apertured and index-guided vertical cavity surface emitting lasers (I²-VCSELs)," vol. 11, pp. 500–502, 1999.
 - [16] Rui Pu, E. M. Hayes, R. Jurrat, C. W. Wilmsen, K. D. Choquette, H. Q. Hou, K. M. Geib, "VCSEL's bonded directly to foundry fabricated GaAs smart pixel arrays," IEEE Photon. Tech. Lett., vol. 9, pp. 1622–1624, 1997.
 - [17] A. V. Kishnamoorthy, L. M. F. Chirovsky, W. H. HObaon, R. E. Leibenguth, S. P. Hui, G. J. Zydzik, K. W. Goossen, J. D. Wynn, B. J. Tseng, J. Lopata, J. A. Walker, J. E. Cunningham, and L. A. D'ASaro, "Vertical cavity surface emitting laser flip-chip bonded to gigabit-per second CMOS circuits," IEEE Photon. Tech. Lett., vol. 11, pp. 128–130, 1999.
 - [18] K. W. Goossen, Cunningham J. E. and Jan W. Y., "GaAs 850 nm Modulators solder-bonded to silicon," IEEE Photon. Tech. Lett., vol. 5, pp. 776–778, 1993.
 - [19] C. Garvin, J. A. Tressa, J. S. Ahern, K. Kang, J. S. Powell and R. Stack, "Overview of high-speed MQW optical modulator devices at Lockheed Martin Sander," Spatial Light Modulators and Integrated Optoelectronic Arrays Conf., Snowmass Village at Aspen, Colorado, April 1999.
 - [20] Yue Liu, "Smart pixel array module based on VCSEL to Si ASIC integration," Spatial Light Modulators and Integrated Optoelectronic Arrays Conf., Snowmass Village at Aspen, Colorado, April 1999.
 - [21] D. L. Mathine, R. Droopad, and G. N. Maracas, "A vertical-cavity surface-emitting laser appliquéd to a 0.8-um NMOS driver," IEEE. Photon. Tech. Lett., vol. 9, pp. 869–871, 1997.

- [22] S. M. Fike, B. Buchanan, N. M. Jokerst, M. A. Brooke, T. G. Morris, and S. P. DeWeerth, "8 × 8 array of thin-film photodetectors vertically electrically interconnected to silicon circuitry," *IEEE. Photon. Tech. Lett.*, vol.7, pp. 1168–1170, 1995.
- [23] Jianglong Zhang, Victor M. Bright and Y. C. Lee, "Thermal Interaction between Laser and Micro-mirrors," *Spatial Light Modulators and Integrated Optoelectronic Array Conf*, p. 111, Aspen, Colorado, April, 1999.
- [24] Michael S. Jin, Volkan Olguz and Sing H. Lee, "Integration of microlaser arrays with thinned and drilled CMOS Silicon Driver Arrays," paper OTuB7, pp. 68–69, *Optical Computing 96*, Sendai, Japan.
- [25] E. M. Hayes, R. Jurrat, R. Pu, R. D. Snyder, S. A. Feld, P. Stanko and C. W. Wilmsen, K. D. Choquette, K. M. Geib, H. Q. Hou, "Foundary fabricated array of smart pixels integrating MESFETs/MSMs and VCSELs," *International Journal of optoelectronics*, vol. 11, pp. 229–237, 1997.
- [26] T. E. Sale, "*Vertical Cavity Surface Emitting Lasers*," Research Studies Press Ltd. and John Wiley & Sons Inc., New York, 1995.
- [27] G. M. Yang, M. H. MacDougal and P. D. Dapkus, "Ultralow threshold current vertical-cavity surface-emitting lasers obtained with selective oxidation," *Electronics Letters*, vol. 31, pp. 886–888, 1995.
- [28] Nobuaki Hatori, Akimasa Mizutani, Nobuhiko Nishiyama, Akihiro Matsutani, Takahiro Sakaguchi, Fumichika Motomura, Fumio Koyama, and Kemichi Iga, "An over 10-Gb/s Transimission Experiment Using a p-Type Delta-doped InGaAs-GaAs Quantum-well Vertical Cavity Surface Emitting Laser," *IEEE Photonics Technology Letters*, vol.10, pp. 194–196, 1998.
- [29] Michael H. MacDougal, Jon Geske, Chao-Kun Lin, Aaron E. Bond, and P. Daniel Dapkus, "Thermal impedance of VCSELs with AlOx-GaAs DBRs," *IEEE. Photon. Tech. Lett.*, vol. 10, pp. 15–17, 1998.
- [30] T. Wipiejewski, D. B. Young, B. J. Thibeault, and L. A. Coldren, "Thermal cross talk in 4 × 4 vertical-cavity surface-emitting laser arrays," *IEEE. Photon. Tech. Lett.*, vol. 8, pp. 980–982, 1996.
- [31] *CRC Handbook of Chemistry and Physics*, Edited by D. R. Lide, 72nd ed. CRC, Boca Raton, 1992.
- [32] *Polymer Handbook*, Edited by J. Brandrup and E. H. Immergut, Wiley, New York, 1989.
- [33] K.M. Lascola, W. Yuen, and C. J. Chang-Hasnain, "Structrual Dependence of the thermal resistance of Vertical Cavity Surface Emitting Lasers," 1997 Digest of the LEOS Summer Topical Meeting, pp. 79–80, Montreal, Quebec, Canada.
- [34] T. S. McLaren, S. Y. Kang, W. Zhang, T. H. Ju and Y. C. Lee, "Thermosonic bonding of optical transeiver based on an 8 × 8 vetical cavity surface emitting laser array," *IEEE, Trans. Compon. Packag. Manuf. Technol.*, vol. 20,

- pp. 152–159, 1997.
- [35] C.-K. Lin, S.-W. Ryu, W.-J. Choi and P. D. Dapkus, “Wafer bonded bottom emitting 850 nm VCSELs on GaP substrates,” *IEEE Photonics Technol. Lett.*, vol. 11, pp. 937–939, 1999.
 - [36] R. A. Morgan, M. K. Hibbs-Brenner, R. A. Walterson, J. A. Lehman, T. M. Marta, S. Bounnak, E. L. Kalweit, T. Akinwande and J. C. Nohava, “Producibile GaAs-base top surface emitting lasers with record performance,” *Electron. Lett.*, vol. 31, pp. 462–463, 1995.
 - [37] Unpublished data.
 - [38] K. Y. Eng, M. J. Karol and Y. S. Yeh, “A growable packet (ATM) switch Architecture: Design principles and applications,” *IEEE Trans. Comm.*, vol. 40, pp. 423–430 1992.
 - [39] C. W. Wilmsen, C. Duan, J. R. Collington, M. P. Dames and W. A. Crossland, “VCSEL-Based Optoelectronic Asynchronous Transfer Mode Switch,” *Optical Engineering*, vol. 38, pp. 1216–1222, 1999.
 - [40] Chunjie Duan, “VCSEL based high performance optoelectronic ATM switches,” M.S. Thesis, Colorado State University, Dec. 1998.
 - [41] Rui Pu, E. M. Hayes, C. W. Wilmsen, K. D. Choquette, H. Q. Hou and K. M. Geib, “Comparison of the techniques for bonding VCSELs directly to ICs,” *J. Opt. A: pure Appl. Opt.*, vol. 1, pp. 324–329, 1999.

Heterogeneous and Monolithic Integration for Optical Interconnects

Duane A. Louderback, Larry A. Coldren, *University of California/Santa Barbara*

Ola Sjölund, *Lucent Technologies*

Eric R. Hegblom, *Hewlett-Packard*

Eva M. Strzelecka, *Honeywell Technology Center*

Shigeru Nakagawa, *University of California/Santa Barbara*

Syn-Yem Hu, *Hewlett-Packard*

Dan D. Lofgreen, Ryan L. Naone, *University of California/Santa Barbara*

Jack Ko, *Picometrix*

The large size and density of VCSEL arrays required for many applications place strict requirements on the VCSELs themselves. To avoid thermal problems and achieve single-mode operation, small, yet efficient, VCSELs are needed. The scaling of device characteristics such as threshold current density, differential efficiency, and wall-plug efficiency to small sizes is examined. Theoretical and experimental results on small, high-performance, tapered-aperture VCSELs are presented. The design, fabrication, and characterization of monolithically integrated refractive microlenses are examined. The microlenses are etched into the substrate side using mask shape transfer and can be used to eliminate or at least relax the tolerances on external optics. The design, fabrication, packaging, and characterization of monolithically integrated arrays of VCSELs and resonant-cavity photodetectors are presented. The devices are monolithically integrated without regrowth and the substrate emitting/detecting devices have lithographically aligned refractive microlenses integrated on the backside of the substrate. Arrays of these integrated devices were flip-chip bonded onto high-speed submounts for modulation testing. High-speed and free-space link results are presented. Finally, monolithically integrated WDM pie-shaped VCSEL arrays for short-distance multimode fiber links are presented. The dense WDM arrays are flip-chip bonded and cross talk results are discussed.

During the preparation of this chapter, all the authors were affiliated with the University of California/Santa Barbara.

4.1 INTRODUCTION

Although computer clock speeds continue to increase, the bandwidths of board-to-board and chip-to-chip interconnects have not kept pace. Unless this interconnect bottleneck is addressed, the current growth of system performance will not continue. Free-space optical interconnects (FSOIs) have received considerable interest for their potential to relieve bottlenecks associated with electrical board-to-board and chip-to-chip interconnects [1–6]. Using arrays of VCSELs or modulators and arrays of photodetectors, several FSOI systems have been demonstrated [7–12], including optical backplanes [13–16] and optoelectronic data filters [17–18]. Besides the potential for achieving large data throughput and high interconnect densities, optical interconnects also have the advantages of alleviating problems caused by electromagnetic cross talk, capacitive loading, and signal distortion. Unlike electrical interconnects, free-space optical interconnects are inherently two-dimensional, and therefore are capable of scaling with chip size better than electrical interconnects. In addition, optical interconnects can provide fan-out without the speed limitations imposed by capacitive loading.

Despite the potential of free-space optical interconnects, many challenges remain, both in the devices themselves as well as their packaging. Many applications of free-space optical interconnects require large, dense arrays of optoelectronic devices, with VCSELs and photodetectors being the devices of choice. In addition to the difficulties involved in fabricating large arrays of devices with good yield, uniformity, and reliability, the operation of large numbers of these devices in a dense array places equally demanding requirements on the device characteristics themselves, particularly the transmitters. To keep the total thermal dissipation at manageable levels, the VCSELs need to be efficient and have low operating currents [7]. This implies simultaneously having both low threshold current and high efficiency. The first approach to reducing the threshold current is simply to make the VCSELs smaller. However, it was learned that the device characteristics don't necessarily scale well, with smaller devices having higher optical, current, and carrier losses [19–20]. Section 4.3 examines these scaling issues and methods of realizing small, efficient VCSELs that are compatible with operation in large, dense arrays.

Optics are an important component of any optical interconnect, regardless of whether the transmission medium is fiber or free-space. Microlenses integrated at the wafer scale, whether they are refractive or diffractive, can be used to eliminate or least relax the tolerances on external optics [21–23]. Section 4.4 discusses the design, fabrication, and characterization of monolithically integrated refractive microlenses.

With packaging comprising a large portion of the cost of optical interconnects, it is imperative that devices and integration techniques are developed to make the packaging process as simple and straightforward as possible. For free-space inter-

connects, it is advantageous to have both the transmitter and receiver on the same chip. This monolithic integration may not only allow denser interconnects, but also might reduce the complexity of packaging further down the line. Monolithic integration of VCSELs and detectors has been demonstrated by many groups using various integration techniques [24–27]. Although monolithically integrating VCSELs and detectors may have some inherent packaging advantages, it doesn't address the challenges of integrating the optoelectronic devices with electronics. Flip-chip bonding is the preferred method of heterogeneously integrating arrays of optoelectronics with electronics such as CMOS. Integration of the necessary optics, such as the microlenses discussed in Section 4.4, is another important facet of packaging devices for free-space optical interconnects. Section 4.5 discusses the monolithic integration of VCSELs and detectors without regrowth. The devices have through-the-substrate emission and detection facilitating the monolithic integration of substrate side refractive microlenses. Flip-chip bonded device and free-space optical interconnect results are presented.

As computer clock speeds and Internet use continue to increase, the demands placed on interconnects become more demanding and challenging to meet. The requirements of short-haul networks are currently being met with multimode optical interconnects [28]. The need for more bandwidth has led to the use of WDM in short distance interconnects [29–30] and future demands will likely necessitate the use of single-mode fibers in these interconnects. Probably the most stringent requirement placed on short-distance fiber interconnects is cost. One approach to reducing the cost of short-haul WDM networks is to monolithically integrate the different wavelength channels into a single pie-shaped structure that can simultaneously couple all the wavelength channels into a multimode fiber [31–32]. These pie-shaped WDM VCSEL arrays were heterogeneously packaged into a test fixture using flip-chip bonding for cross talk measurements. Packaging and cross talk results are presented in Section 4.6.

4.2 SCALING VCSEL PERFORMANCE: SMALL, HIGH-PERFORMANCE VCSELS

Manufacture of relatively large, multimode vertical cavity lasers for data links using multimode fiber is now commonplace. However, in the future, relatively small ($\leq \sim 5 \mu\text{m}$) vertical cavity lasers may also become important for other applications, such as board-to-board or chip-to-chip free-space optical interconnections or high-speed printing [33]. These applications require arrays of lasers with low power consumption and operating at output powers of 100–500 μW . In the case of free-space interconnections (or when multimode fiber is not used), one usually wants single mode devices to reduce optical cross talk and intermodal noise and that generally means using small diameters, $\sim 3 \mu\text{m}$. But smaller lasers are not only

desirable for single-mode operation. If properly scaled (constant J_{TH} , η_{EX}), the peak wall-plug efficiency will be higher in smaller devices despite resistance increasing as $\sim 1/(\text{aperture diameter})$ [34]. This higher wall-plug efficiency will occur, as desired, at lower output powers. Properly scaled devices will also provide higher modulation bandwidths at lower drive power. However, to realize these improvements one must overcome several obstacles to scaling device characteristics, namely: optical scattering losses, current spreading, carrier diffusion and increased electrical and thermal resistance.

Although current spreading or carrier diffusion will raise the threshold in larger ($\sim 10\text{-}\mu\text{m}$ diameter) devices, optical scattering losses present a much greater barrier at smaller sizes because they not only lower the slope efficiency, but they also increase the threshold gain which in turn leads to higher lateral leakage currents. The excess loss generally turns on sharply [roughly as $\sim (1/R_d)^3$, where R_d is the device radius] [19, 35], and will have an exponentially greater effect on threshold. These losses have been dramatically reduced by confining the mode with apertures rather than etched pillars [35–36]. In addition, previous work with thin apertures and tapered apertures has demonstrated further reductions in scattering loss [37–38].

Current spreading between the aperture and the active region and carrier diffusion within the barriers and quantum wells raise the threshold, but by themselves should not effect the slope efficiency due to the clamping of the carrier population in the active region. One might expect the unclamped voltage at the aperture to increase the current spreading above threshold and drop the slope efficiency, but an estimate of this increased spreading above threshold suggests it should not be significant even at diameters of $1\text{ }\mu\text{m}$ [39]. Even so, the contribution of current spreading to threshold is significant at much larger diameters.

Use of dielectric apertures instead of etched mesas for current confinement has also improved the scaling of the electrical resistance so it varies approximately as $1/R_d$ instead of $1/R_d^2$ as for the etched pillar case. Under this geometry, both the thermal and electrical resistance scale in a similar manner. This scaling implies that at a constant current density, the temperature and voltage (above the diode drop) should actually decrease linearly with device radius. Unfortunately, due to the parasitics of current spreading and carrier losses, a constant output power density does not translate into a constant current density versus device size. And even if all these lateral leakage and scattering loss parasitics were removed, the scaling of the thermal and electrical resistance implies an optimum device radius for a given output power. Nevertheless, we can conclude that as the operating power is lowered (provided parasitic optical and current/carrier leakage is removed), smaller devices will be more efficient despite increased electrical and thermal resistance.

Increased resistance at smaller sizes is a limitation when one wants to drive VCSELs directly from 3.3-V CMOS, and it is expected that the device resistance

still could be improved with better control of the doping profile in the first few mirror periods. Nevertheless, single-mode devices with aperture openings of 2–3 μm , 20% wall-plug efficiency at 150 μW , and low (154 μA) threshold still had a drive voltage under 2.5 V at 300 μW output power. One might also expect the increased resistance at smaller sizes to increase the high frequency RC time constant. However, it is the parasitic capacitance of the aperture that will allow the radial current distribution to widen at higher frequencies and lower the dynamic resistance. Thus, at high speed one is limited more by the vertical resistance to reach the oxide (in series with the source resistance) and the parasitic capacitance across the oxide and diode depletion region, rather than the lateral resistance from funneling current into the aperture. This difference in resistance at high and low frequencies is also needed to explain the RC parasitic observed in apertured intracavity contacted devices [40]. Having summarized all the major limitations to device scaling, the aim of the remainder of this section is to quantify the improvements to the device scaling in comparing three VCSEL structures.

Although other studies have also shown the improvement in optical confinement by positioning apertures at the standing wave nulls [41–43], this chapter quantifies the effects for tapered apertures which also depend on the gradient of the taper. In addition, this chapter goes further to quantify the improvement in current confinement (by comparing two different structures both with apertures at the nulls) and how these improvements combine with the scaling of the electrical resistance to impact the wall-plug efficiency at lower powers and at sizes small enough to be single-mode. While high wall-plug efficiencies of 50% [44] and 57% [45] at output powers of 1.3 mW and 3.3 mW, respectively, have been achieved in larger 7- μm and 5- μm diameter devices, the improvements being touted here are for smaller, single-mode devices at lower powers where the scaling issues are much more important.

4.2.1 SUMMARY OF DEVICE STRUCTURES

A general schematic for the VCSELs discussed in this section is shown in Fig. 4.1. All the structures are bottom-emitting VCSELs with 18.5 bottom mirror periods as shown schematically in Fig. 4.1. They all have the same calculated mirror transmission (0.7%) and the same number (three) and composition ($\text{In}_{0.18}\text{Ga}_{0.82}\text{As}$) for the quantum wells. A schematic of the central portion of the three different VCSEL structures examined in this section is shown in Fig. 4.2. The aperture is centered at different points in the standing wave: the first null from the active region, the first peak and the second null. The distance between the aperture tip and the active region varies from ~ 230 nm in the second-null device to ~ 50 nm for the first-null device. All the tapers are created by oxidation of a thin AlAs layer adjacent to $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}$. For the first- and second-null devices, the thin layer of AlAs is 10 nm thick and for the first-peak device the layer is 8.5 nm thick. The

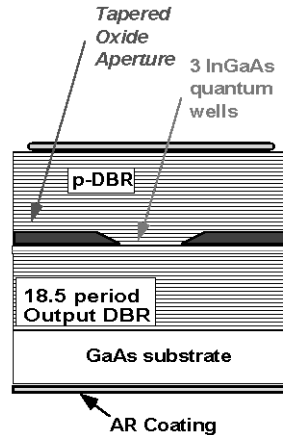


Figure 4.1. General schematic for the three tapered-aperture VCSEL structures discussed in this chapter.

first-peak and second-null devices are not intracavity contacted and this impacts the broad-area slope efficiency because of some substrate absorption (but comparison of scattering losses will not be affected). The first-peak device is carbon doped in the top mirror using a graphite filament in the MBE chamber. The top mirror in the second-null devices was Be doped and aside from the aperture layer, it used an $\text{Al}_{0.7}\text{Ga}_{0.3}\text{As}/\text{GaAs}$ top mirror as opposed to the other devices which used $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}/\text{GaAs}$ top mirrors. (Unlike the first-null devices, the mirror in the second-null devices was grown entirely at 600°C so the Be doping profile was far from ideal.) Silicon doping was used in the bottom mirrors of all structures. Other details of the mirror doping and processing have been previously published [38, 46–47]. Table 4.1 lists some of the different parameters. In the subsequent analysis, changes in the broad-area device parameters will be accounted for so that the scaling is the salient feature.

4.2.2 OPTICAL CONFINEMENT

For tapered apertures, the degree of optical confinement depends critically not only on the position of the aperture in the standing wave, but also the gradient of the taper thickness which combine to give the gradient of the effective index step. From scanning electron microscope measurements of the tapers (like that shown in Fig. 4.3), one can determine the taper length, though the exact taper profile is harder to determine. Nevertheless, one can estimate the index gradient if one assumes that the taper thickness varies linearly and a plot of the effective index step, Δn_d , versus position is given in Fig. 4.4. The index step is calculated based upon the overlap of the index perturbation with the standing wave [48] which agrees with the index-step calculated based upon the change in the resonant wave-

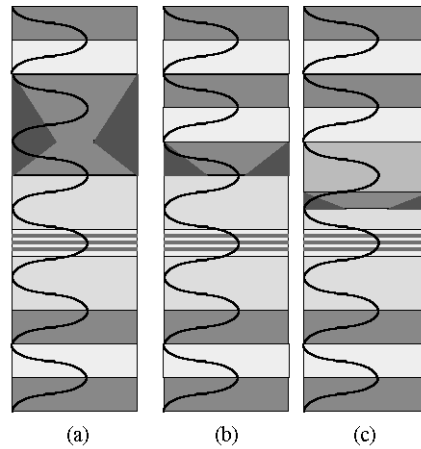


Figure 4.2. Schematic of the central portion of three VCSEL structures discussed in this work in which the aperture is positioned at (a) the second null from the active region (b) the first peak and (c) the first null. The black areas indicate the oxide aperture, the lightest regions represent GaAs, the gray regions indicate AlGaAs, except for the three stripes across the center which denote the three InGaAs quantum well active region. The oscillating line represents the standing-wave pattern. Grading of the interfaces is not shown.

length [19, 48] except in the thickest regions, where the perturbation of the oxide moves a second cavity resonance within the mirror stop-band. One can see that the taper positioned at the standing-wave peak has a much higher index gradient and this results in increased scattering losses. These losses are plotted in Fig. 4.5. This excess optical loss is calculated by using the slope efficiency to extract the total optical loss and then subtracting the loss found in large ($\sim 20\text{-}\mu\text{m}$ diameter) devices [35–36]. Calculating the total optical loss per pass,

$$L_T = T/2 \left(\frac{T_s \eta_i}{\eta_{ext}} - 1 \right), \quad (4.1)$$

relies upon calculating the output mirror transmission, T , estimating the injection efficiency, η_i , and the substrate transmission, T_s , which was lower for the highly doped n-substrate under the first-peak VCSELs. The substrate transmission can be directly measured by comparing the reflectivity in the stop band of the top of the DBR (effectively unity) with the reflectivity (again in the stop band) from the bottom of the sample. The injection efficiency is estimated to be 80% for the Be-doped devices and 90% for the first-peak carbon-doped device. (The injection efficiency is consistently measured higher for carbon doped in-plane lasers than for Be doped in-plane lasers with the same structure, presumably due to diffusing of the Be around the active region. This diffusion would increase the voltage to create inversion and the corresponding excess field could lower the conduction-band barrier that normally keeps electrons from entering the p-side.)

Aperture Position	Second Null	First Peak	First Null
# of quantum wells	3	3	3
Output transmission	0.7%	0.7%	0.7%
Broad-area loss	0.12%/pass	0.37%/pass	0.10%/pass
J_{TH} (broad-area)	810 A/cm ²	870 A/cm ²	600 A/cm ²
Est. doping between aperture and active	5×10^{17} cm ⁻³	1×10^{18} cm ⁻³	6×10^{17} cm ⁻³
Distance between aperture and active	210 nm	130 nm	50 nm
P-dopant	Be	C	Be
N-dopant	Si	Si	Si

Table 4.1. Parameters for the tapered-aperture VCSELs under investigation.

One also would like to correct for any injection efficiency changes with size. Unfortunately, this is difficult to do. However, as mentioned, a theoretical estimate of the size dependence of the injection efficiency due to current leakage indicates an assumption of constant-injection efficiency is adequate and measurements suggest this is correct. In addition, the excellent scaling of the slope efficiency of tapered apertured devices sets an upper bound on the drop of less than 10% at 2 μm . Despite this slight inaccuracy, our conclusions about the relative amounts of optical scattering from various aperture designs should be similar. We show the scattering loss in comparison to theoretical predictions [19] in Fig. 4.6. Figure 4.6(a) shows the comparison of the abrupt apertures data with the theoretical predictions, and Fig. 4.6(b) shows the comparison for the tapered apertured devices. Comparison shows that the measured loss is somewhat higher than predicted. Nevertheless, their coincidence on approximately the same scale suggests that scattering losses are significant in the devices and that the improvements observed agree with our understanding of optical loss. For the smallest size devices with tapered apertures, the comparison is trickier than for the abrupt apertures since it depends on the exact taper profile and the DBR stop band. The second-null curve remains low and then curves up at the smallest sizes, unlike the first-null curve. This effect has to do with the higher overall index step, which can confine smaller modes. (The mode radius for these sizes is predicted to be smaller than for the first-null curve.) The smaller modes at these sizes exceed the DBR angular stop band. The location of this rise is very sensitive to the taper length and the DBR angular stop band. The measured mode radius ($1/e^2$) point for the second-null device at 1- μm radius was $\sim 5 \mu\text{m}$, suggesting that the effective index may change more

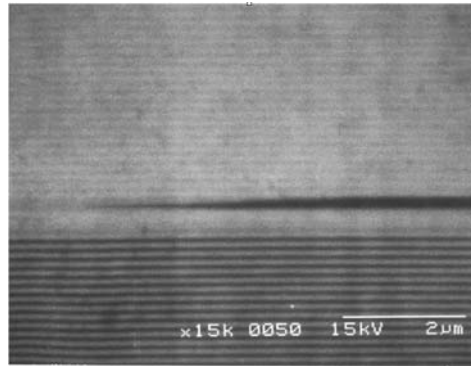


Figure 4.3. Scanning-electron-microscope image of a laterally oxidized tapered aperture (black region).

slowly around the aperture tip than estimated. The simulation for the first-null device suggests that because of the weaker index step, the DBR angular stop band should not be creating excess loss at the 1- μm - and 2- μm -diameter openings and this stipulation agrees with the measured mode size described below.

Figure 4.5 shows that the first-peak apertures have only slightly lower scattering loss than for devices with an abrupt 80-nm-thick oxide aperture [49]. The taper is providing some benefit to decrease the loss, but it is clear that the taper is more effective at standing-wave nulls where the effective index gradient is lower. In fact, the excess loss is so low that the 2- μm -diameter second-null VCSEL was multimode. (The excess loss in first-peak device is responsible for single-mode operation for the two- and three-micron-diameter devices because the waveguide index step of 0.06 is too strong to be a single-mode wave guide at that size.) On

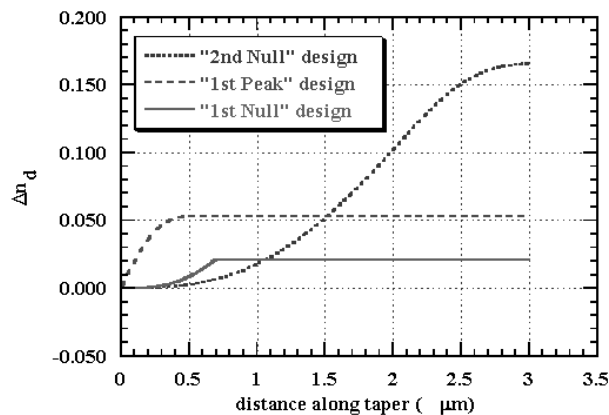


Figure 4.4. Effective index step versus position for the three different tapered-aperture designs calculated assuming that thickness varies linearly with position.

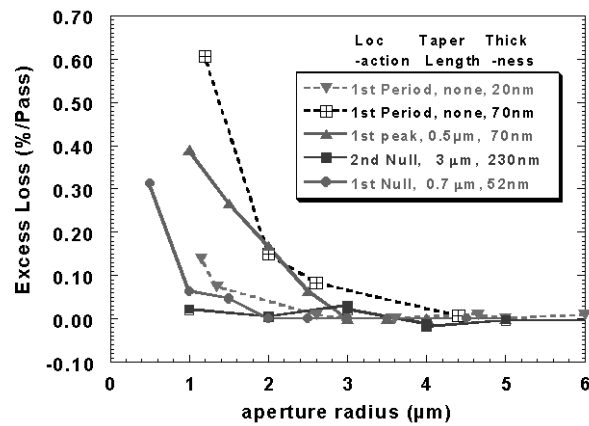


Figure 4.5. Summary of the excess optical “scattering” loss extracted from the change in slope efficiency with size for three tapered-aperture devices and two other VCSELs with abrupt apertures.

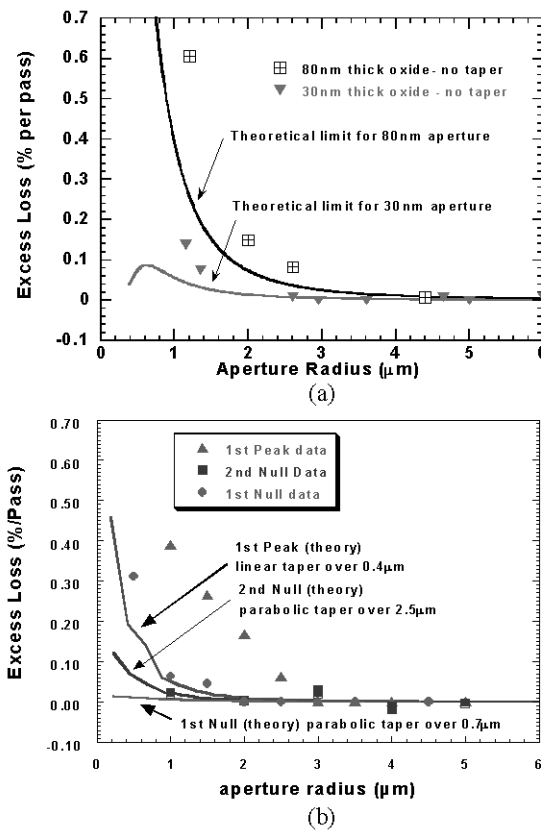


Figure 4.6. Comparison of measured and predicted loss for the same structures as Fig. 4.5 using calculations from [19] for the (a) abrupt apertures and (b) tapered apertures.

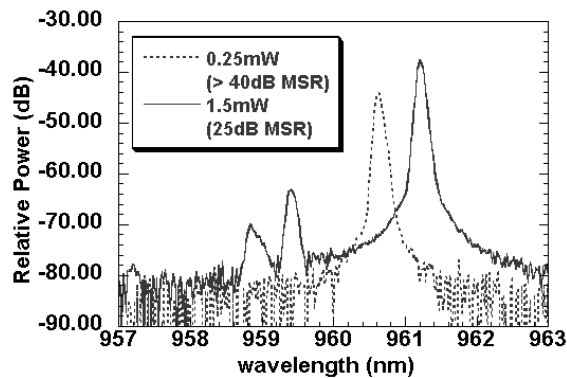


Figure 4.7. Optical spectrum from a 3- μm , first-null VCSEL showing single-mode operation at output powers below 1.5 mW.

the other hand, the weaker wave guiding for the first-null device promoted single-mode operation for the 2- and 3- μm -size devices. Optical spectra for the 3- μm first-null device is shown in Fig. 4.7.

The size given for the abrupt-aperture devices [37, 49] in Fig. 4.5 is calculated based upon the lateral mode separation. The size for the aperture opening for the tapered-aperture devices is determined by the closing of the aperture for pillars below a particular diameter. Since the pillar size increases in steps of 1 μm , a label of radius r_0 means the actual radius of the opening is between $r_0 - 0.5 \mu\text{m}$ and r_0 .

Due to the weaker guiding of both thinner apertures and tapered apertures, the mode is expected to be larger than the mode of an abrupt thick aperture of the same opening. Figure 4.8 shows a line scan of the near-field intensity in a first-null device with a 2- μm opening that shows about a $1/e^2$ diameter of 3 μm . The system resolution was not taken into account and even though a high 0.75-NA lens was used, sharp features like the boundary of an etched mesa had a boundary about 1 μm wide. Thus, the actual diameter may be even smaller, approximately 2.8 μm . Based on a vertically uniform waveguide analysis, one can calculate that for the mode guided by an 80-nm-thick abrupt aperture (effective index step, $\Delta n_d = 0.06$) to have a $1/e^2$ diameter of 3 μm , the radius of the aperture opening should be $\sim 1.8 \mu\text{m}$. But for this larger size, the excess optical losses of an abrupt, 80-nm-thick aperture are around 0.2%/pass, which still is much greater than those for the 2- and 3- μm -diameter tapered and thin apertures. Thus, even at the same mode size, thin or tapered apertures are preferable to quarter-wave-thick ones. This result also agrees with theoretical calculations of the scattering loss versus mode size [19].

One may also ask if tapered apertures are preferable to thin apertures when both are placed a standing-wave null. From the point of view of optical scattering losses for the lowest order mode there is little difference. However, there are situa-

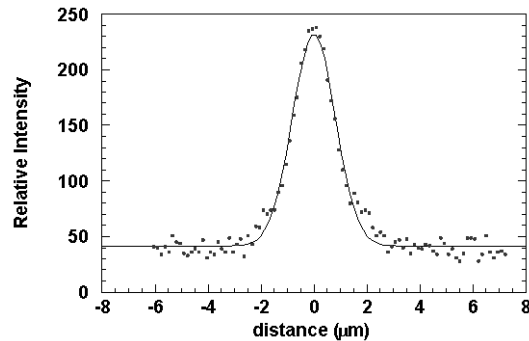


Figure 4.8. Near-field intensity versus position for a mode from a first-null VCSEL with a 2- μm -diameter opening.

tions where tapering is desirable. If one wishes to promote multimode operation in smaller VCSELs, use of a thick, tapered aperture is preferable to a thin aperture, which will not guide higher-order modes. Such a design is desired when making lower-threshold “medium-sized” ($\sim 5\text{-}\mu\text{m}$ to $10\text{-}\mu\text{m}$ diameter) VCSELs for coupling to multimode fiber. Furthermore, tapering to greater thickness fairly far ($> 3\text{ }\mu\text{m}$) from the optical mode is desirable to reduce the parasitic capacitance through the oxide. Finally, tapering may occur as a result of other design conditions; it is typically necessary to drive current through the layers within the aperture. Apertures at the first null are within the cavity, which is typically of high index, low aluminum content. However, if there is a large discontinuity in the aluminum composition, then one can incur a high drive voltage. This can be avoided with longer grading or surrounding the aperture with higher aluminum content that will frequently create a tapered aperture. So one must be mindful of the index gradient to avoid higher loss.

4.2.3 CURRENT CONFINEMENT

With size-dependent optical losses nearly eliminated, the next major challenge for improving the scaling of characteristics is to reduce the lateral current and carrier leakage. Previous studies have indicated that lateral current spreading between the aperture and the active region can be significant in apertured devices [35, 39, 41]. Figure 4.9(a) plots the threshold current versus aperture radius for the three different structures. Although there is some difference in the broad-area threshold current densities, most of the reduction in the threshold can be attributed to improved current confinement as the aperture is positioned closer to the active region. One sees that even the close position of $\sim 50\text{ nm}$ is still beneficial. For clarity, the ratio of the threshold current density to the broad-area threshold current density is also plotted in Fig. 4.9(b). Although the first-peak devices have higher scattering losses, the increase in the quantum-well threshold current is small

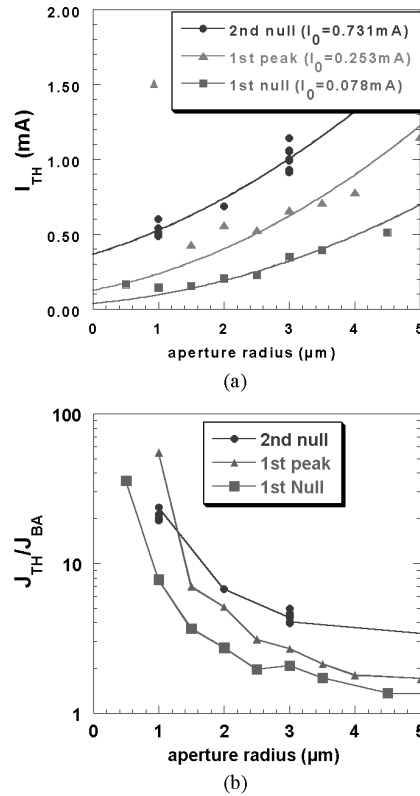


Figure 4.9. (a) Threshold current versus aperture radius for three different taper designs. (b) Same data normalized by the broad-area threshold current density.

(except at the smallest size) given the size of the active region. Therefore, most of the change in threshold current density relative to the broad-area threshold is due to current confinement only. To make this point clearer, it is useful to compare with modeling for the spreading when there is a purely resistive region between the aperture and an active region with a diode current density versus voltage characteristic [39]. In that model, I_{TH} (in the absence of excess loss) should scale as

$$I_{TH} = J_{BA}\pi R^2 + \frac{I_0}{2} + R\sqrt{\pi J_{BA}I_0}, \quad (4.2)$$

where

$$I_0 = 8\pi(\eta kT/q)\frac{t}{\rho},$$

for a layer of uniform resistivity, ρ , and thickness, t .

Figure 4.9(a) shows the fit we obtained for the parameter I_0 , using the broad-area threshold current density from 20- μm - to 30- μm -diameter devices. (For the

first-peak device, the smallest size is excluded from the fit because the threshold current increases due to scattering losses.) One can now compare the fitted I_0 parameter with that expected from the doping between the aperture and the active region. (Note that for the fitted values, the dominant term in the threshold is $I_0/2$ for the smallest sizes.) With the thickest region between the aperture and the active region, the second-null devices show the highest I_0 as expected, but it is still almost four times higher than if the region were uniformly doped at $5 \times 10^{17} \text{ cm}^{-3}$ as approximately designed. One possible reason is that relatively low Be doping combined with movement during growth at the heterointerfaces creates a much higher vertical resistance, forcing the current to spread further laterally. The carbon-doped first-peak devices are a better example.

These were doped around $1 \times 10^{18} \text{ cm}^{-3}$ between the aperture and active region yielding a result near the predicted value, $I_0 \sim 0.2 \text{ mA}$ [where a mobility $\sim 60 \text{ cm}^2/(\text{V sec})$ is expected for the higher doping in AlGaAs]. For the first-null device, the region between the aperture and the active region was not intentionally doped, but secondary ion mass spectroscopy (SIMS) of this region (which was grown at a substrate temperature of 600°C) revealed Be moved down during growth and reached a concentration of $\sim 6 \times 10^{17} \text{ cm}^{-3}$ over this region. Using this value, one obtains $I_0 \sim 70 \mu\text{A}$, which is a reasonable match given the typical error in concentration and mobility.

An additional effect to briefly consider is that the injected profile actually matches the aperture size in which case the injected current is smaller than the mode for the $2\text{-}\mu\text{m}$ -size device and the lateral confinement factor is 50–70% (depending upon the true aperture size). However, this effect would only imply a factor of 1.4–2.4 higher threshold current density (assuming a logarithmic gain model with three quantum wells) needed in the active region, which is still well below the observed factor of ten. We should note some reduction in confinement factor is not detrimental and, in fact, can make for a lower threshold current than one would achieve for a unity-confinement factor of a mode of the same size because the power density in the mode is not uniform.

4.2.4 SCALING OF RESISTANCE AND WALL-PLUG EFFICIENCY

In order to achieve high wall-plug efficiency at the low output powers (100–500 μW) one desires even lower power dissipated at threshold. Consequently, one must be able to maintain high differential efficiency as the device threshold (and size) is reduced. With low background losses, scattering losses, and improved current confinement, the first-null VCSELs presented can provide high wall-plug efficiency at low output powers. Figure 4.10 shows the wall-plug efficiency versus output power for several first-null devices described in this section. The $4\text{-}\mu\text{m}$ - and $3\text{-}\mu\text{m}$ -diameter devices have peak wall efficiencies of 26% and 22% at output powers of 0.5 mW and 0.34 mW, respectively. In addition, one slightly-higher-threshold (perhaps slightly larger) $5\text{-}\mu\text{m}$ device reached 30% wall-plug efficiency at 0.9 mW.

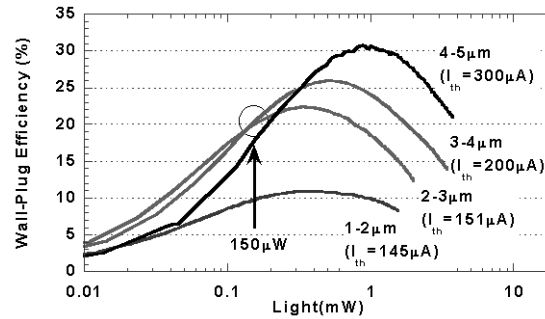


Figure 4.10. Wall-plug efficiency versus output power for several first-null VCSELs.

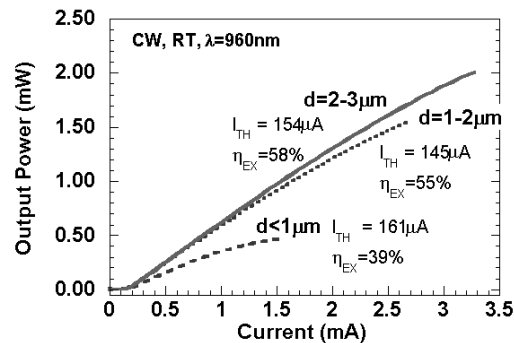


Figure 4.11. Light versus current characteristics for the smallest of the first-null devices.

At 150 μW , the 3- μm device shown has a wall-plug efficiency of 20%. Figure 4.11 shows the light-current curves for the smaller of these same devices. There is even a 1- μm -diameter device (as discussed, 1 μm is the upper bound on the diameter) that lases with a threshold of 161 μA and $\eta_{EX} = 39\%$.

Although the threshold may be lower for smaller devices, the peak wall-plug efficiency suffers because the devices have a resistance that increases approximately inversely with aperture radius, in agreement with other studies [34, 50]. Even if there was no current spreading, the increase in resistance causes the wall-plug efficiency (at a given output power) to have a maximum at a particular device diameter. However, as output power is lowered, removal of parasitic leakage becomes even more important for higher power conversion efficiency. In fact, if the resistance increases as $1/(\text{aperture radius})$, and there is no lateral leakage or optical losses, then the peak wall-plug efficiency will be higher in smaller devices. Of course, this peak keeps occurring at lower and lower powers. If one resigns oneself to having leakage current, then it may be desirable to make a larger size device single-mode. To illustrate these points, the size dependence of various

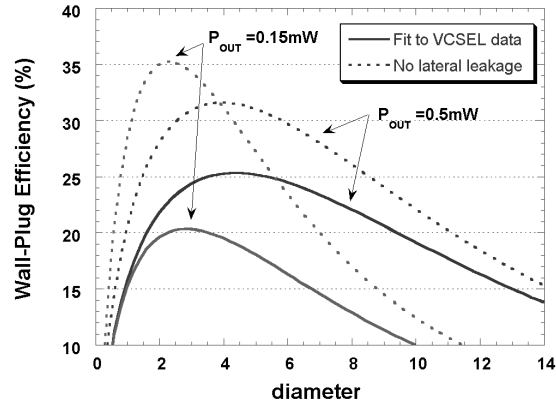


Figure 4.12. Extrapolated wall-plug efficiency versus device size with and without lateral leakage currents.

device parameters of the first-null devices have been fitted in order to extrapolate to the condition of no lateral leakage current. Figure 4.12 plots the wall-plug efficiency versus size at two different output powers of 0.5 mW and 0.15 mW based on the actual data and the case of no leakage current. As can be seen from the graph, when leakage is present the wall-plug efficiency at 0.5 mW is higher than at 0.15 mW, but if leakage is removed, the efficiency is actually higher at 0.15 mW than 0.5 mW. To generate the curves, both the size dependence of the current-voltage curves and the threshold current for the first-null devices was fit versus size and the slope efficiency was held constant (which only has a minor error at small sizes). The operating voltage fit to a current independent voltage, V_D , which varied as $V_D = 49 + 0.745d^{-1.56}$ (volts), where d is the aperture diameter in μm , and a series resistance, R_S , varying as $R_S = 2.51\text{k}\Omega(1/d)^{0.7}$. The size-dependence of the threshold fit well to $I_{TH} = 0.13 \text{ mA} + J_{BA}\pi(d/2)^2$, where $J_{BA} = 600 \text{ A/cm}^2$. The constant term was ignored to simulate no leakage current.

4.3 MONOLITHIC REFRACTIVE MICROLENSSES

Wafer-level integration of optoelectronic devices with microlenses will result in components that can be used directly in systems, without the need of aligning external optics. Combinations of VCSELs, LEDs, or detectors, with microlenses have applications in optical fiber systems, free-space optical computer interconnections [51], displays, printers, etc. Refractive lenses, etched in the backside of the substrate, have several advantages over diffractive lenses. Besides the simple and manufacturable fabrication technology, the lens optical efficiency is primarily limited by the absorption of the substrate, which is nominally transparent. Furthermore, these lenses are far less sensitive to wavelength variations. A process for fabricating refractive microlenses in GaAs and InP with the ability to control the

radius of curvature (ROC) within $\pm 10\%$, suitable for many practical applications is presented here. The process requires formation of a lens-shaped mask that is then transferred into the substrate by reactive ion etching (RIE) [52]. Microlenses can be formed from deposited materials, such as photoresist [53] or polyimide [54], on the surface of the substrate. However, by taking the process one step further and etching the lens in the substrate, a refractive index discontinuity at the lens-substrate interface is eliminated and the lens thickness reduced, since semiconductor materials have a high refractive index.

4.3.1 FABRICATION

This method of microlens fabrication requires transfer of a lens-like shape of patterned, reflowed photoresist into the semiconductor by anisotropic dry etching. Reactive ion etching was chosen over etching with inert ions [55] to achieve faster etch rates and a controllable ratio of the etch rate in the semiconductor to that in the erodable mask material. The latter allows precise control of the lens shape by adjusting the etch rate ratio during mask transfer.

Formation of a lens-shaped mask is the first part of the fabrication process. Since the lens ROC is primarily controlled by the shape of the erodable resist mask, the initial mask material thickness and diameter are of most importance. The polyimide polydimethylgluterimide (PMGI), used as an erodable mask, is spin-coated on the substrate and patterned into circles of desired lens diameter. This is done by patterning a thin layer of conventional photoresist and then exposing the PMGI to deep UV light and developing. PMGI can also be patterned in O_2 plasma or RIE, which prevents undercutting and loss of its volume. Conventional Novolak photoresist is removed with acetone. In case thick layers are needed, several layers of PMGI are coated and baked at 200°C on a hot plate after each spin. The initial thickness of PMGI is roughly set by the spin speed and then adjusted by controlled etchback in O_2 RIE after patterning. The semiconductor in the field is etched to form a pedestal, prior to reflow of the mask. The edge of the pedestal, typically 0.1 to $0.5\text{ }\mu\text{m}$, creates surface tension that prevents the PMGI from spreading during reflow, and results in preserving the lithographically defined diameter. Dry etching better preserves the edge of the circle. However, wet etching, though in some cases crystallographic, may also be applied for a shallow pedestal etch without much distortion of the lens shape, especially in case of large-diameter lenses ($> 100\text{ }\mu\text{m}$).

Next, the patterned PMGI cylinders are reflowed into lenses. The reflowed PMGI shape depends on the time and temperature of reflow as well as on the feature's aspect ratio. PMGI melts just below 300°C , and after ~ 30 minutes forms a spherical shape. The reflowed shape is well fitted by a paraboloid whose volume and diameter remain the same as the original cylinder before reflow, resulting in the peak height after reflow being twice the cylinder height. This is valid for height to diameter ratios greater than 0.015. Very thin layers of PMGI remain pla-

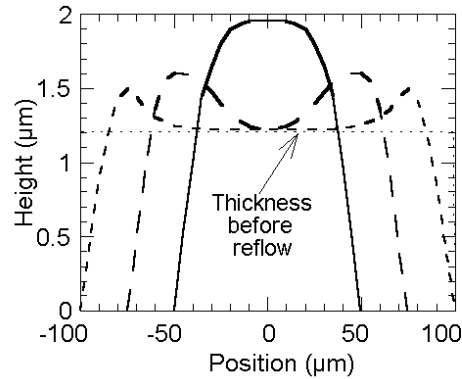


Figure 4.13. Surface-profilometer scans of PMGI after reflow at 290°C for 20 min for three different diameters and aspect ratios.

nar in the central region and deform only around the perimeter due to surface tension, as shown in Fig. 4.13. Hence, to achieve semiconductor surfaces with very large ROC, a thicker mask should be chosen and the pattern transfer in subsequent RIE done with a small semiconductor to mask etch rate ratio.

The lens-like shape of the mask is transferred to a semiconductor substrate in RIE. The final lens shape strongly depends on the ratio of etch rate in the semiconductor to that in the PMGI. Two in situ optical monitors in the RIE reactor makes it possible to control the etch rate ratio, thereby achieving a desired curvature of the surface and correcting for aberrations. PMGI etch rate is monitored by measuring the intensity of 632.8-nm HeNe laser light reflecting from a sample coated with a known initial thickness of the material. Due to interference between the top reflection and the reflection from the PMGI-substrate boundary, the signal exhibits periodic variations as the PMGI surface etches, with a period proportional to the etch rate, material's refractive index, and incidence angle. Etch rate of bulk semiconductors is monitored by measuring a diffracted beam from the surface patterned with parallel lines. As the exposed semiconductor etches, the diffracted signal exhibits periodic variations with a period proportional to the etch rate and incidence angle. The lines are masked with nickel and sputtered aluminum oxide that have much smaller etch rates than the semiconductors.

The etch rate of GaAs increases with pressure, due to the increasing chemical nature of the process, whereas the PMGI etch rate remains fairly constant, as shown in Fig. 4.14. The bias voltage proportionately increases the etch rate of both materials. The etch rate of GaAs is about three to five times faster than PMGI. Under typical conditions (2 mT and 350 V bias), the etch rates of GaAs and PMGI are 220 and 110 nm/min, respectively. At pressures higher than 7 mT the GaAs etch rate further increases but the semiconductor undercuts and the etch become

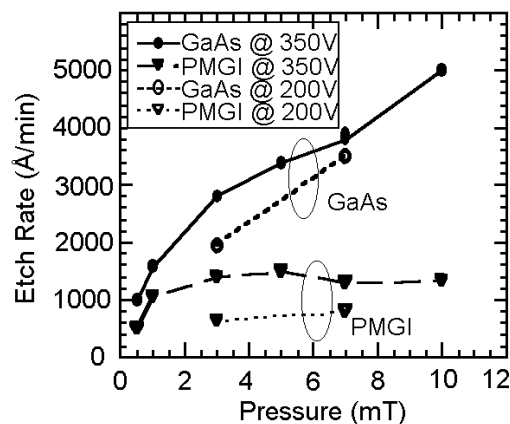


Figure 4.14. Etch rates of GaAs and PMGI as a function of pressure and bias voltage in Cl_2 RIE.

crystallographic due to ion scattering and an abundance of reactive species. The initial mask is made thinner than the final GaAs microlens, as illustrated by profilometer scans shown in Fig. 4.15.

InP etches very slowly in Cl_2 at room temperature due to low volatility of indium chloride reaction products. Under similar RIE conditions as above, the InP etch rate is only 40 nm/min while the PMGI etch rate is ~ 110 nm/min. This requires very thick mask layers and long etch times to form microlenses. In a different RIE reactor, we have used BCl_3/Cl_2 chemistry to etch InP microlenses. The pressure was kept low (~ 2 mT) to improve the volatility of the reaction products, and the gas composition was varied to increase the etch rate of InP relative to PMGI. At lower RF powers (200–350 W) the InP to PMGI etch rate ratio can be

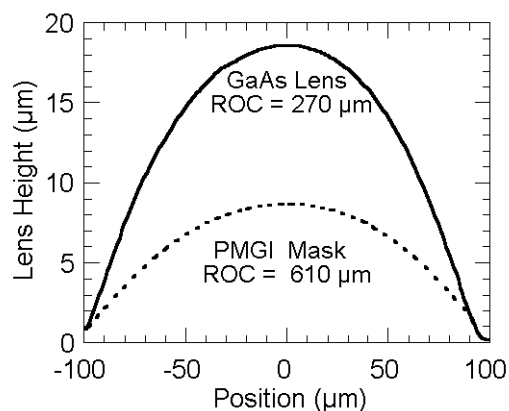


Figure 4.15. Surface-profilometer scans of the lens-shaped mask before RIE and the GaAs lens after transfer of the pattern in Cl_2 RIE.

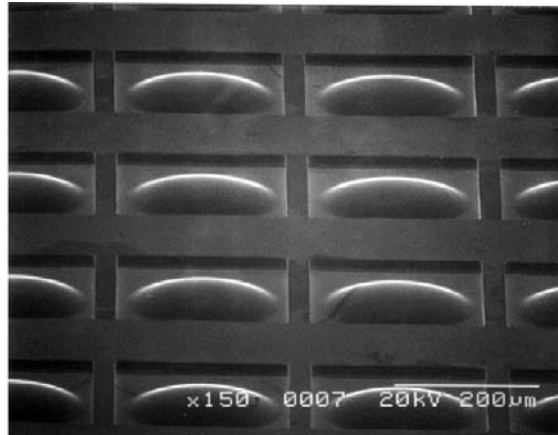


Figure 4.16. Scanning-electron-microscope image of GaAs microlenses recessed in the substrate.

varied from 0.6 to 1.4 by changing the flow rate ratio of BCl_3 to Cl_2 from 1 to 4. The etch rates are in the 100 nm/min range. By increasing RF power to 465 W (900 V bias), at flow rate ratio of BCl_3 to Cl_2 equal to 3, the etch rate of InP increased to 182 nm/min while the PMGI etch rate increased to 150 nm/min. With these conditions, microlenses were etched in InP substrates with high-quality, optically smooth surfaces. The higher power provides enough energy to reactively etch InP without the need to heat the substrate. In the same conditions, conventional Novolak photoresist erodes at a rate ~ 80 nm/min.

For integration with devices, the microlens is recessed into the semiconductor wafer, as shown in Fig. 4.16, for ease of handling, lens surface protection, and contact metallization. This is achieved by protecting some areas of the surface with thick photoresist, patterned after lens mask formation, prior to RIE. As an example, microlens requirements for collimation of a VCSEL beam exiting through a 500- μm substrate are $\text{ROC} \approx 360 \mu\text{m}$ (focal length $\sim 142 \mu\text{m}$ in air). The diameter requirements depend on the original divergence of the laser. To ensure minimum diffraction from the lens perimeter, for this application, lenses of diameters between 100 and 200 μm are typically produced.

4.3.2 INTEGRATION WITH DEVICES

As a method of evaluation of the performance, microlenses were integrated with VCSELs. The lenses were aligned to the laser features in an IR mask aligner. Because the distance between the lens and the laser is fixed by the substrate thickness ($\sim 500 \mu\text{m}$), the output beam divergence depends not only on the lens focal length but also on the initial laser diameter. The far-field beam divergence angles were measured for several types of single-mode laser and lens combinations. The

laser diameter (μm)	laser beam half-angle in air (no lens)	lens diameter (μm)	peak lens height (μm)	lens focal length in air (μm)	half-angle in air with lens
7.0	6.5°	100	8.8	60	12.3°
7.0	6.5°	150	8.8	125	2.2°
7.0	6.5°	200	8.8	225	1.9°
6.0	8.1°	200	17.4	117	1.1°
3.1	13.0°	200	13.8	151	1.0°

Table 4.2. Far-field half-angles of VCSELs integrated with refractive microlenses.

half-angles of $1/e^2$ of beam intensity were measured by laterally scanning a small aperture across the beam and confirmed by angularly scanning a single-mode fiber. A summary of these results is shown in Table 4.2. The first four cases are of etched post VCSELs. The output beams can be either collimated or focused down, depending on the lens focal length. The first case is of the beam focusing at about 100 μm outside the lens and diverging even stronger than the original beam. The last case shows that for very-small-diameter lasers, achieved by using dielectrically apertured structures [37], the beam can be collimated down to 1° half-angle, due to the initial strong divergence of $\sim 13^\circ$, which allows the beam to fill the lens.

4.4 MONOLITHICALLY INTEGRATED VCSELS AND DETECTORS WITH MICROLENSSES

The two-dimensional nature of free-space optical interconnects places demanding requirements on the devices to be used, especially for the transmitter. To avoid problems with heat dissipation, the devices need to be efficient and operate at low power [7, 56]. Devices that have low drive voltages and currents are also necessary since CMOS electronics continue to scale down to smaller sizes and drive voltages. In addition, to realize large data throughput, the devices need to have large bandwidths at these low powers. Recent advances in VCSEL technology have yielded devices that meet these requirements: VCSELs with low threshold currents [57–58] and voltages [59], as well as high power efficiencies [44–45] and large bandwidths, including large bandwidths at low drive currents, have been demonstrated [40, 60].

To fully utilize the two-dimensional nature of these optical interconnects it is desirable to have arrays of integrated VCSELs and photodetectors. This minimizes the distance between the electronics and the optical emitters and receivers in a bi-directional interconnect system. Figure 4.17 shows a schematic of a simple bi-directional FSOI consisting of arrays of integrated VCSELs and detectors. Many

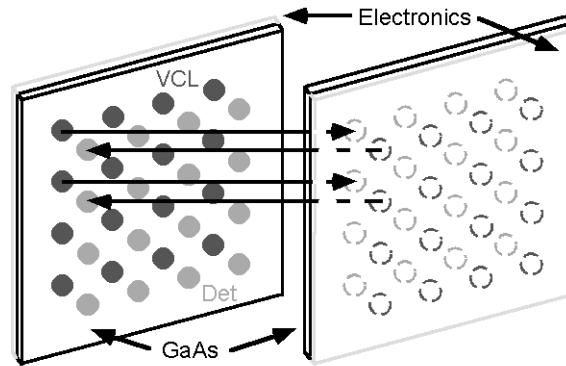


Figure 4.17. Schematic of a bidirectional free-space optical interconnect consisting of arrays of VCSELs and detectors packaged onto electronics.

groups have demonstrated monolithic integration of VCSELs and photodetectors with adequate performance [24–27]. However, for free-space optical interconnects to compete with electrical interconnects, the additional requirement of low cost is imposed. This means that not only must the devices be inexpensive, but they also need to be compatible with simple packaging techniques. Flip-chip bonding is the method of choice for packaging 2D arrays of optoelectronics with electronics. This requires that both the optical input and output be on one side of the sample (preferably the substrate side) and the electrical inputs and outputs on the other side of the sample (preferably the top side). However, all previous integration techniques have used top-emitting and detecting architectures, requiring vias through the substrate in order to bring the electrical contacts to the substrate side. Such complex processes as this are undesirable because of the detrimental effect they have on yield, especially for large arrays, and therefore cost. Integration of micro-optics that eliminate the need for or relax the tolerances on external optics is the other key to low-cost packaging. For top-emitting/detecting devices, this means fabricating the micro-optics on a separate superstrate, which is subsequently attached to the device array. This involves a critical planarity alignment between the optoelectronics and microlenses.

If the substrate is transparent at the wavelength of operation, devices can emit/detect through the substrate and photolithographically aligned micro-optics can be integrated directly on the backside of the substrate [21–23]. For GaAs, this means devices that operate at wavelengths greater than ~ 900 nm. Using such substrate-emitting/detecting devices, flip-chip bonding can be performed directly without complex via processes. Figure 4.18 shows a schematic illustrating the inherent advantages of the through-the-substrate architecture. The VCSEL emits through the substrate and the detector is illuminated through the substrate allowing

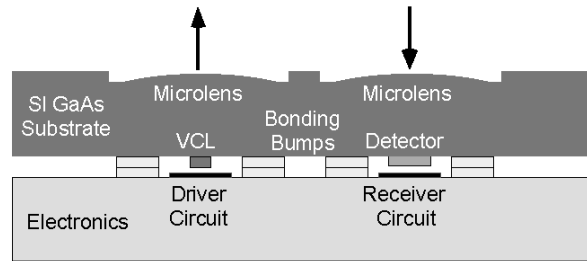


Figure 4.18. Schematic illustrating the inherent packaging advantages of substrate-emitting/detecting devices. The integrated VCSELs and detectors are flip-chip bonded directly to electronics without the need for via processes or substrate removal. Microlenses are integrated on the backside of the wafer to eliminate or relax tolerances for external optics.

straightforward integration of micro-optics, such as refractive microlenses, on the backside of the substrate. As mentioned above, these lenses can eliminate the need for or relax tolerances on external optics, and would typically collimate the VCSEL output and focus a collimated beam on the detector. In this section, we present monolithic integration of high-performance VCSELs and photodetectors on a single epitaxial growth structure with a through-the-substrate emission/detection compatible with cheap packaging and substrate-side micro-optics.

4.4.1 DEVICE DESIGN

The simplest integration technique uses different technologies for the emitter and detector, normally a VCSEL used as the emitter and an integrated MSM detector as the detector [24]. In this approach, the detector is grown on top of the VCSEL structure and then simply removed when fabricating the VCSELs. Unfortunately, this technique is not compatible with substrate emission/detection since the detector would have to be grown on the bottom and thus cannot be removed. Another integration technique is to use patterned regrowth where one device structure is grown first and then the epilayers in the field are selectively removed before the other device structure is grown. This, however, reduces the device yield and increases the complexity of the fabrication, leading to a substantial increase in cost. The most obvious integration technique is to try to use the same technology for both the VCSEL and detector: resonant-cavity enhanced photodetectors [61]. Although a resonant-cavity detector and a VCSEL can have a similar epitaxial structure, they have very different requirements for the input/output mirror. Using a VCSEL as a detector is possible, but since this detector would have an input-mirror reflectivity, R_{in} , larger than 99%, the absorption would be low and the optical bandwidth, defined as the full width at half maximum (FWHM), would be narrow. A wide optical bandwidth is required to tolerate vari-

ations in wavelengths over a wafer and from wafer to wafer, as well as differences in operating temperatures between different arrays. The peak quantum efficiency, η , and FWHM optical bandwidth, $\Delta\lambda_{1/2}$, of a resonant cavity detector are given by [61–62]

$$\eta = \frac{(1 - R_{in})(1 - R_{back}e^{-\Gamma_{enh}\alpha d})(1 - e^{-\Gamma_{enh}\alpha d})}{(1 - \sqrt{R_{in}R_{back}}e^{-\Gamma_{enh}\alpha d})^2} \quad (4.3)$$

$$\Delta\lambda_{1/2} = \frac{\lambda}{2\pi n_{\lambda s}} \frac{1 - \sqrt{R_{in}R_{back}}e^{-\Gamma_{enh}\alpha d}}{(R_{in}R_{back})^{1/4}e^{-\Gamma_{enh}\alpha d/2}} \quad (4.4)$$

where R_{in} and R_{back} are the input- and back-mirror reflectivities, respectively. α is the absorption coefficient, λ the resonance wavelength of the detector, d the total thickness of the absorbing material, $n_{\lambda s}$ the effective cavity length in multiples of the resonant wavelength, and Γ_{enh} the standing-wave enhancement factor. Figure 4.19(a) shows the relationship between quantum efficiency, optical bandwidth, and input-mirror reflectivity for a detector with $R_{back} = 100\%$, $\alpha = 6,000 \text{ cm}^{-1}$, $n_{\lambda s} = 3.5$, $\Gamma_{enh} = 83$ (the absorbing quantum wells [QWs] are centered around a standing-wave maximum), and $d = 240 \text{ \AA}$ (three $80\text{-\AA}$ InGaAs QWs). It can be seen in Fig. 4.19(a) that operating a VCSEL as a detector would yield a device with $\sim 40\%$ quantum efficiency and $\sim 1 \text{ nm}$ of optical bandwidth. By reducing the input-mirror reflectivity to about 95% , the detector becomes matched, i.e. no light is reflected back. A matched detector has the maximum possible quantum efficiency, but still has a fairly narrow optical bandwidth of $\sim 2 \text{ nm}$. Note that for this case (100% back-mirror reflectivity) the absorption is 100% , but $\sim 4\%$ of the incident light is absorbed outside the cavity. To obtain a wider optical bandwidth the input-mirror reflectivity must be reduced even further, thus trading quantum efficiency for optical bandwidth. For a FSOI, the bandwidth should be at least $5\text{--}6 \text{ nm}$, which can be obtained with an input mirror reflectivity of $\sim 80\%$, with a corresponding peak quantum efficiency of 60% . Thus, with the proper choice of input-mirror reflectivity, the detector can have both a reasonably large optical bandwidth and adequate absorption. Figure 4.19(b) shows an example of detector spectra for three different detector designs [shown with dashed lines in Fig. 4.19(a)]. In a top-emitting/detecting configuration, the required difference in mirror reflectivities between the VCSEL and detector can easily be achieved by etching away mirror periods (as can be seen in Fig. 4.27) from the VCSEL's top mirror [26–27], but for a through-the-substrate approach this is obviously not possible. Therefore, these different requirements for the bottom mirror have made the integration of substrate-emitting/detecting devices with adequate performance prohibitively difficult.

Using a novel technique [63], substrate-emitting VCSELs and substrate-illuminated resonant-cavity detectors [64] have been successfully integrated. The

devices share the same epitaxial structure and achieve the different bottom-mirror reflectivities by selective oxidation of a few $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ layers [65–66] for the VCSELs, eliminating the need for regrowth. Thus, the grown detector structure is converted into a VCSEL structure by changing the properties of the bottom mirror by postgrowth processing. A schematic of the epitaxial structure, designed to be resonant at 990 nm, is shown in Fig. 4.20. The right side of the Fig. 4.20 shows the structure when used as a detector, or as grown. The left side shows the structure when converted into a VCSEL by the oxidation of the bottom mirror. Figure 4.21 shows a scanning-electron-microscope image of an epitaxial structure for a VCSEL (after oxidation). Note that this is not the same epitaxial structure that was used to fabricate the devices presented in this chapter. This structure uses different compositions in the aperture and mirror layers to enhance the difference in oxidation distances, whereas the structure used for the devices presented in this paper uses a two-oxidation process. Nevertheless, this single-oxidation structure illus-

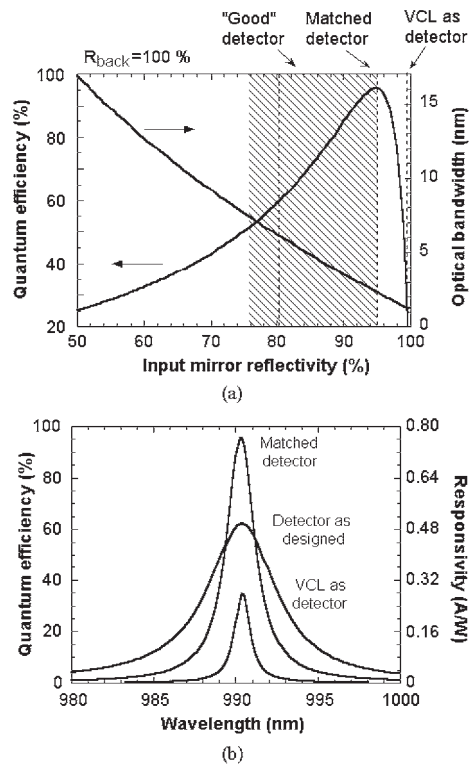


Figure 4.19. (a) Quantum efficiency and optical bandwidth versus input mirror reflectivity for a resonant-cavity detector, resonant at 980 nm, with three 80-Å InGaAs QWs, an absorption coefficient of 6000 cm^{-1} , an effective cavity length of 3.5 wavelengths, an enhancement factor of 1.83, a back-mirror reflectivity of 100%, and AlGaAs/GaAs mirrors. (b) Absorption spectrums of several different possible detector designs.

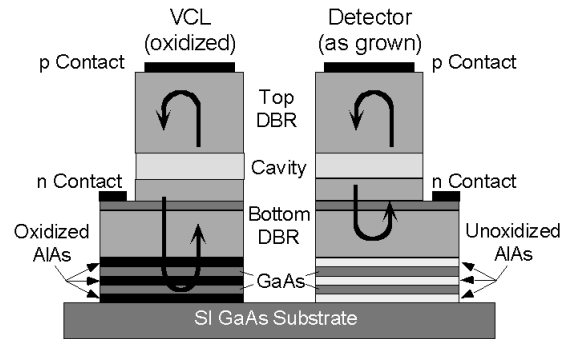


Figure 4.20. Schematic of the device structure showing the as-grown detector (right) and when converted to a VCSEL by increasing the reflectivity of the bottom mirror by selective oxidation of the three $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ layers (left).

trates the final epistructure (after the two oxidations) of the VCSEL presented in this section. From Fig. 4.21, it can be seen that the three bottom $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ layers have oxidized about $15\ \mu\text{m}$ (the oxide is black), while the oxide-aperture layer used for current confinement has oxidized about $5\ \mu\text{m}$. The rest of the $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}$ layers in the mirrors have not oxidized significantly ($< 2\ \mu\text{m}$).

The three $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ layers are placed below the detector's bottom mirror (i.e., below the detector). When these layers are oxidized, Al_xO_y is formed, thereby reducing the index of refraction of these layers from ~ 3.0 to ~ 1.57 [67]. Therefore, they are out of phase with the rest of the bottom mirror when not oxidized and thus do not affect the reflectivity or resonance wavelength of the detector significantly. This is illustrated in Fig. 4.22 where the spectral response for a detector with and without these (unoxidized) $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ layers is plotted.

When these layers are oxidized they come into phase with the rest of the mirror, increasing the reflectivity of the mirror significantly, thereby converting the as-

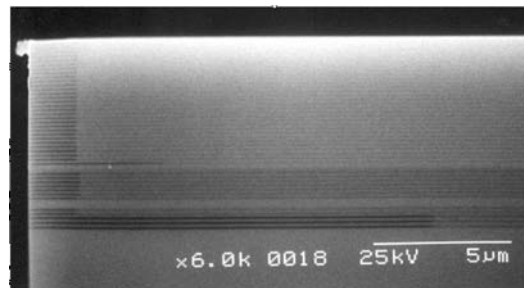


Figure 4.21. Scanning-electron-microscope image showing the layer structure of the VCSEL after oxidation. The bottom three $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ are oxidized $\sim 15\ \mu\text{m}$ while the aperture layer is oxidized $\sim 5\ \mu\text{m}$ and the $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}$ layers in the rest of the DBR are oxidized $< 2\ \mu\text{m}$.

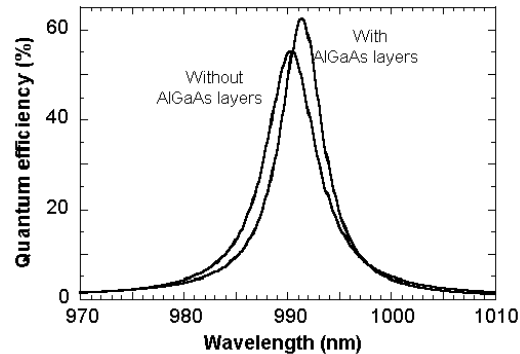


Figure 4.22. Quantum-efficiency spectrum for a detector with and without the “extra” $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ layers, showing that these layers do not affect the detector design significantly.

grown detector structure to a VCSEL structure. Because of the large index difference (high-contrast mirror) of this $\text{Al}_x\text{O}_y/\text{GaAs}$ mirror, only a few periods of this oxide mirror are needed to reach $> 99\%$ reflectivity [68].

To achieve the desired input mirror reflectivity of 80% for the detector, nine periods of quarter-wave $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}/\text{GaAs}$ are required. The additional reflectivity needed to form the VCSEL is generated by 2.5 periods of $\text{Al}_x\text{O}_y/\text{GaAs}$. Since quarter-wave layers of these periods would result in a reflectivity that is too high, the thicknesses of the oxidized layers were detuned from quarter wave. This fine-tunes the reflectivity by lowering the reflectivity of each half-wave period. Using oxidized layers of approximately $1/8 \lambda$ and GaAs layers of approximately $3/8 \lambda$ gives 99.3% reflectivity. From modeling, less than 1-nm difference in resonance wavelength for the detector and VCSEL is expected. Thus, using this technique, VCSELs and resonant-cavity photodetectors can be individually designed without regrowth or compromising performance in either device. Using a similar epitaxial structure, top- and bottom-emitting VCSELs can be monolithically integrated [69].

4.4.2 DESIGN ISSUES

The position of the $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}/\text{GaAs}$ mirror periods is important since these periods are badly detuned when they are left unoxidized. By placing them at the bottom of the mirror, they are essentially outside of the detector’s cavity. If they were placed on top of the mirror, next to the cavity, the effective cavity length of the detector would be increased and from Eq. (4.4), it can be seen that this results in a reduced optical bandwidth. The second reason for this placement is that it the lasing wavelength of the VCSEL becomes more tolerant to variation in these layers. This is important since it is desirable to have the VCSEL and detector cavities at the same wavelength. When using oxide mirror layers there are several possible uncertainties in the material properties that could affect

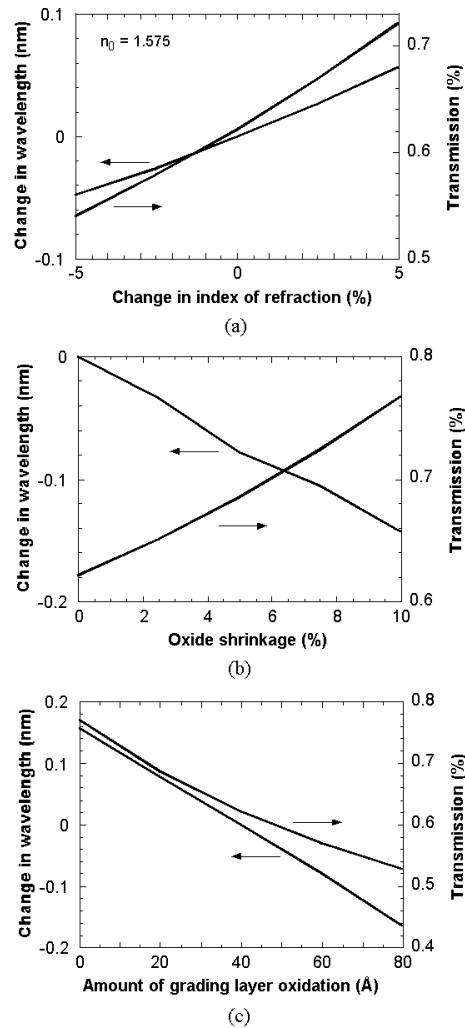


Figure 4.23. VCSEL-lasing wavelength and output-mirror transmission variations with respect to (a) changes in the index of refraction in the oxide, (b) shrinkage of the oxide, and (c) how much of the 160-Å grading layer, from GaAs to $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$, oxidizes.

device characteristics [67, 70–72]: the index of refraction of the oxide, shrinkage of the oxide with respect to the AlGaAs layers, and the amount that the 160-Å grading layer from GaAs to $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ oxidizes. The latter is perhaps the most important since it is difficult to control and this layer acts as a tuning layer similar to multiwavelength VCSEL arrays with oxide layer tuning [73]. Figure 4.23 shows the effects on lasing wavelength and output-mirror transmission (which affects the differential efficiency, η_d , and the threshold current, I_{th}) for the three issues discussed above: (a) variations in the index of refraction of the oxide of $\pm 5\%$ from the assumed value of 1.575, (b) oxide shrinkage of 0–10%, and (c) the

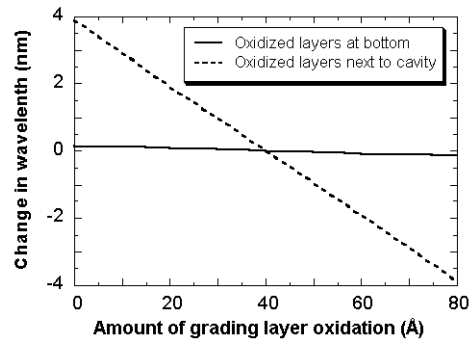


Figure 4.24. A comparison of how the grading layer oxidation affects the lasing wavelength in the design discussed here (solid) and a design where the oxide layers are next to the cavity (dashed).

amount the grading layer oxidizes. From Fig. 4.23, it can be seen that for all these possible variations, the lasing wavelength changes less than ± 0.2 nm and the output mirror transmission is between 0.5% and 0.8%, not significantly affecting the VCSEL characteristics.

Figure 4.24 shows what would happen to the lasing wavelength if these GaAs/ Al_xO_y layers were placed next to the cavity for the same variations as in Fig. 4.23(c). The variation increases from an insignificant ± 0.2 nm to a detrimental ± 4 nm. The variations in mirror transmission are similar in both cases.

Figure 4.25 shows how the mirror transmission and (round-trip) optical losses (L) affect the differential efficiency of a VCSEL (the injection efficiency, η_i is assumed to be 85%). The three curves represent different optical losses, ranging from 0.14% (which is the lowest optical loss reported [80]) to 0.3% (which is larger than the expected losses). From this one can conclude that for the expected optical

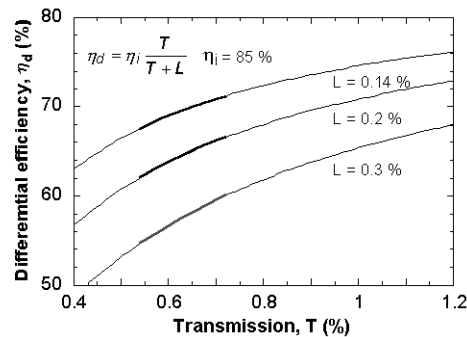


Figure 4.25. Plot of the differential efficiency, η_d , versus output-mirror transmission, T , for a VCSEL with an internal quantum efficiency, η_i , of 85%, and optical losses of 0.14%, 0.2%, and 0.3%. The highlighted part corresponds to the variations in transmission in Fig. 4.23(c).

losses, between 0.14% and 0.3%, the variations in transmission in Fig. 4.23 have little effect on the differential efficiency. The threshold gain, and thus the threshold current density, is affected similarly.

Equation (4.4) shows that the optical bandwidth of the detector is inversely proportional to the effective cavity length. The penetration depth in GaAs/AlGaAs mirrors is about $5/4 \lambda$. Using two GaAs/AlGaAs mirrors and a $1\text{-}\lambda$ cavity thus results in an effective cavity length of $\sim 3.5 \lambda$. A more aggressive approach is to use high-contrast GaAs/ Al_xO_y for the back mirror. This would reduce the effective cavity length to $\sim 2.5 \lambda$ since the penetration depth in the oxide mirrors is only about $1/4 \lambda$, increasing the bandwidth by $\sim 40\%$. Figure 4.26(a) shows the optical bandwidth versus effective cavity length for the three detector designs in Fig. 4.19, while Fig. 4.26(b) shows the expected detector response for the detector as designed and a detector with an oxidized top mirror. Another advantage with the high-contrast mirror is that the number of mirror periods needed to reach a certain reflectivity is reduced significantly. This is shown in Fig. 4.27 where the reflectiv-

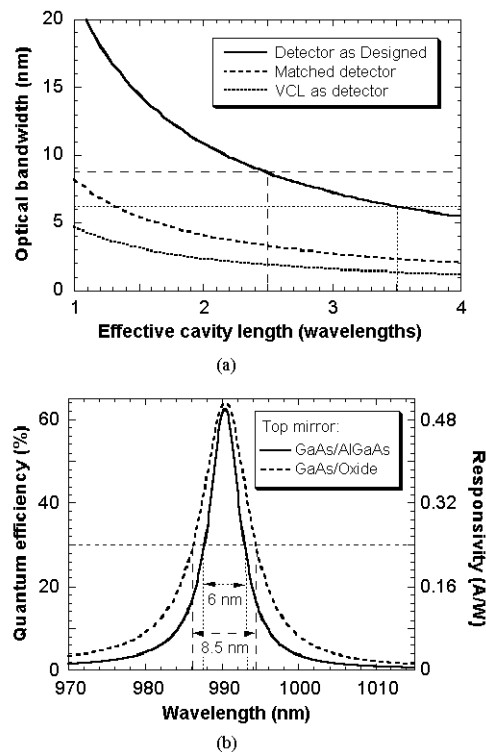


Figure 4.26. (a) Optical bandwidth as a function of effective cavity length for the three detector designs in Fig. 4.19. (b) Detector spectra for a detector with $\sim 60\%$ quantum efficiency with 3.5λ (GaAs/AlGaAs DBR, solid) and 2.5λ (GaAs/Oxide DBR, dashed) effective cavity length.

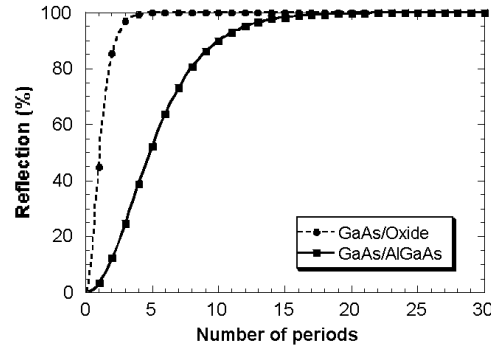


Figure 4.27. Reflectivity as a function of number of quarter-wave periods for a GaAs/ $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}$ DBR (solid) and a GaAs/Oxide DBR (dashed) showing that only four periods of the oxide mirror are needed to reach reflectivities above 99% and six periods are needed to reach 99.95% versus 17 and 25 periods for the semiconductor mirror.

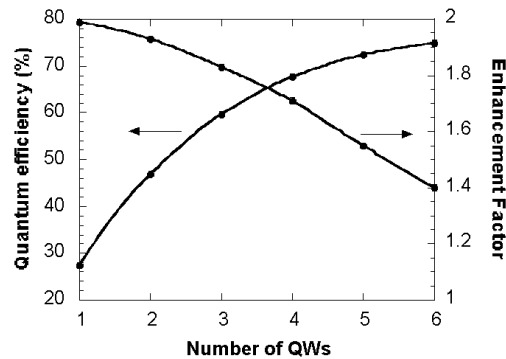


Figure 4.28. Quantum efficiency and standing-wave enhancement factor as a function of the number of 80-Å QWs (with 80-Å GaAs barriers) at 80% input-mirror reflectivity for the detector presented in Fig. 4.23 above. As the number of QWs is increased, the increase in total absorption length, αd , is offset by less overlap between the standing wave and the QWs.

ity as a function of number of mirror periods is plotted for a semiconductor mirror and an oxidized mirror. To reach 99.9%, 23 mirror periods are needed for the semiconductor mirror compared to only 5 for the oxidized mirror. The main disadvantages with oxidized mirrors are the more difficult processing (double intracavity contacts and more etches to expose the surfaces to be oxidized), the uncertainties reported in Fig. 4.23 above, and the risk of strain-induced delamination of the oxide layers.

Another way to increase the optical bandwidth and/or the absorption is to increase the absorption-length product, αd . One way of doing this is to increase the number of QWs in the cavity. But, as the number of quantum wells increase, the standing-wave enhancement factor, Γ_{enh} , decreases because of the total thick-

ness of the QWs. Figure 4.28 illustrates this problem. By using four or five QWs the absorption can be increased quite significantly, but beyond that there is not much to gain by increasing the number of QWs. Of course, this means that the In-content in the QWs has to be lowered or narrower wells must be used, to avoid strain-induced defect formation.

The absorption coefficient, α , can also be modified. In the calculations above, an absorption coefficient of 6000 cm^{-1} was used. For InGaAs QWs, α can be much larger than this, typically 10000 cm^{-1} with an exciton peak of $\sim 15000\text{--}20000 \text{ cm}^{-1}$. However, due to the quantum-confined Stark effect, this value is reduced and the absorption edge is redshifted when the QWs are placed in an electric field. At low fields, α can exceed 15000 cm^{-1} , but is very wavelength dependent ($> 10000 \text{ cm}^{-1}$ over $< 10 \text{ nm}$). However, for very large fields α is reduced to about $6000\text{--}8000 \text{ cm}^{-1}$ with a fairly flat spectral response ($\pm 15\%$ over 30 nm) [75]. Typically, non-resonant detectors have $\sim 1\text{-}\mu\text{m}$ intrinsic regions, resulting in a built-in electric field that is less than $2 \text{ V}/\mu\text{m}$. However, in a $1\text{-}\lambda$ cavity VCSEL design the intrinsic region is typically only about $1000 \text{ \AA}$, resulting in built-in fields exceeding $10 \text{ V}/\mu\text{m}$. In our design we used an intrinsic region of $1800 \text{ \AA}$, corresponding to a built-in field of about $8 \text{ V}/\mu\text{m}$ (assuming that the total depletion region on the p and n sides is $\sim 500 \text{ \AA}$). By expanding the intrinsic region, the built-in field could be reduced, leading to higher absorption coefficients. Figure 4.29 summarizes the effects this would have on the quantum efficiency and optical bandwidth. The three curves are for a αd of 0.0144 (three $80\text{-}\text{\AA}$ QWs and 6000 cm^{-1}), 0.024 (three QWs and 10000 cm^{-1}), and 0.032 (four QWs and 10000 cm^{-1}). Thus, with an

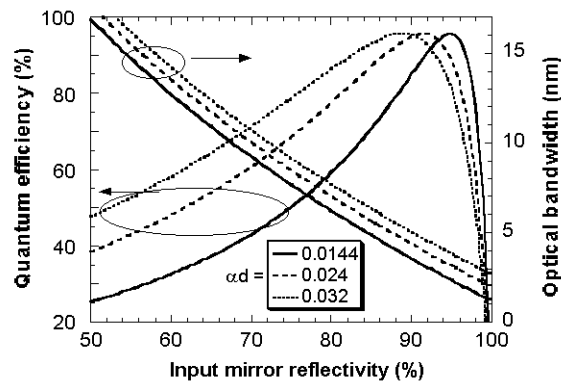


Figure 4.29. Quantum efficiency and optical bandwidth for several values of the absorption-length product, αd , versus input-mirror reflectivity for a resonant-cavity detector with 100% back-mirror reflectivity. All other parameters are the same as in Fig. 4.19 above. The different values of αd represent three QWs with $\alpha = 6000 \text{ cm}^{-1}$ (0.0144), three QWs with $\alpha = 10000 \text{ cm}^{-1}$ (0.024), and four QWs with $\alpha = 10000 \text{ cm}^{-1}$ (0.032). In all cases, the QWs are assumed to be $80 \text{ \AA}$ wide.

aggressive, but straightforward, design it is possible to simultaneously achieve quite large quantum efficiencies and optical bandwidths.

4.4.3 DEVICE FABRICATION

The device structure was grown by solid-source molecular-beam epitaxy using Be and Si dopants for p- and n-doping, respectively. The p-doped top mirror consists of a contact layer, 27 periods of $\lambda/4$ $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}/\text{GaAs}$, and three periods of $\lambda/4$ $\text{Al}_{0.75}\text{Ga}_{0.25}\text{As}$ (next to the cavity). A 200-Å $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ layer is included in the bottom $\text{Al}_{0.75}\text{Ga}_{0.25}\text{As}$ layer to be used as an oxide current-confinement layer for the VCSELs. All interfaces between GaAs and AlGaAs are linearly graded over 160 Å to reduce the mirror resistance. The $1\text{-}\lambda$ cavity consists of three 80-Å $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$ QWs (photoluminescence peak at 980 nm) with 80-Å GaAs barriers, placed at the peak of the standing wave, clad by 100 Å of GaAs and 1100 Å of $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ on either side of the multiple-quantum-well region. The intrinsic region of the cavity is 1800 Å. The bottom mirror consists 2.5 periods of n-doped $\lambda/4$ $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}/\text{GaAs}$, a $5/4\text{-}\lambda$ n^+ -contact layer, six periods of intrinsic $\lambda/4$ $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}/\text{GaAs}$, and finally 2.5 periods of intrinsic $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}/\text{GaAs}$, with the $\text{Al}_{0.98}\text{Ga}_{0.02}\text{As}$ being 752 Å thick and the GaAs layers 787 Å thick. Again, all interfaces were linearly graded over 160 Å.

Monolithically integrated VCSELs and detectors were fabricated using standard photolithographic techniques. The fabrication started with Ti/Pt/Au p-contact formation and subsequent mesa etching down to the n-contact layer using Cl_2 RIE. Next, Ti/Au/Ge/Ni/Au n-contacts were formed adjacent to the pillars and SiN_x was deposited to protect the detector pillars during oxidation of the VCSEL current-aperture layer. After this oxidation, a large mesa was etched to the substrate by Cl_2 reactive ion etching, exposing the mirror periods to be oxidized. Once again SiN_x was deposited to protect the detectors and VCSEL aperture layer from oxidizing. Both oxidations were done at 425°C under N_2 gas bubbled through deionized water at 80°C. This step also served to anneal the contacts. Figure 4.30 shows the layout of devices with detectors (rows 1 and 3), on a 500- μm pitch, and interleaved VCSELs (rows 2 and 4) offset 250 μm both horizontally and vertically from the detectors. Note that the large circles above the contact pads are dummy bumps for the flip-chip bonding.

4.4.4 MICROLENS INTEGRATION AND FLIP-CHIP BONDING

Refractive microlenses designed to collimate the VCSEL output or focus a collimated beam on the detector are monolithically integrated on the substrate side of the devices. The photolithographically aligned microlenses are fabricated by transferring the shape of a reflowed polyimide mask into the GaAs substrate by Cl_2 reactive ion etching [23]. Areas between the microlenses are masked with photoresist for the etch, yielding protective walls that surround the subsequently

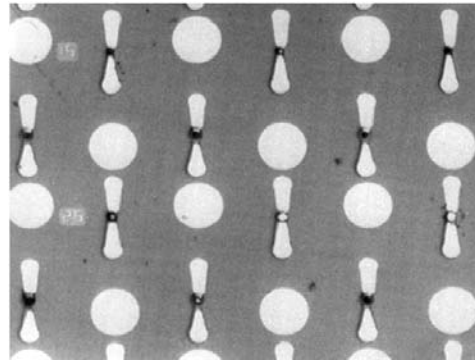


Figure 4.30. Microscope picture of the device layout where rows 1 and 3 are detectors and rows 2 and 4 are VCSELs offset 250 μm by 250 μm from the detectors. The large circles are dummy bumps for flip-chip bonding.

recessed microlenses. In addition, the microlenses are antireflection coated with SiN_x . Figure 4.31 shows a scanning-electron-microscope image of the recessed substrate side GaAs microlenses.

The device chip is diced into one-dimensional arrays containing eight VCSELs and eight detectors. These arrays are flip-chip bonded onto semi-insulating GaAs submounts with fan-out patterns of 50- Ω coplanar transmission lines. The lasers and detectors each have two indium bonding bumps per device, one for the p contact and one for the n contact. The indium bumps are 40 μm in diameter and approximately 6 μm in height. Since the device arrays only have 32 bonding bumps (16 devices, each with two bumps), the overall area of the bonding bumps is quite small. This small bonding bump area means the bonding force required is

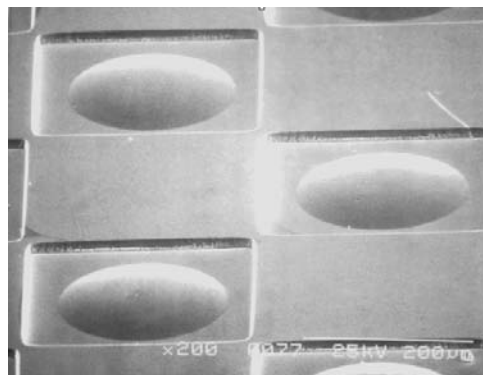


Figure 4.31. Scanning-electron-microscope image of substrate side, showing the recessed microlenses.

fairly low, resulting in poor reproducibility and mechanical support. To achieve better reproducibility, a larger bonding force is desirable. For this reason, large dummy bonding bumps were put near each device to increase the overall bonding bump area by about a factor of five. The submounts have matching Indium bonding bumps of about the same height on them. Since the devices have a nonplanarity of about $7.5\text{ }\mu\text{m}$, holes were etched into the submounts to ensure the device pillars would not be damaged.

The flip-chip bonding was performed using a Research Devices M-8A flip-chip aligner/bonder to precisely align the corresponding Indium bumps on the device array and submount. The two pieces are then pressed together and the pressure is maintained for about one second. A bonding mass of $\sim 1000\text{ g}$ was used, which corresponds to a bonding pressure of $\sim 5\text{ kg/mm}^2$. After flip-chip bonding, the device array is underfilled with an epoxy for mechanical support and heat spreading.

Figure 4.32 shows a scanning-electron-microscope image of a flip-chip bonded array. The recessed, substrate-side microlenses can be seen on the top of the device array. The integrated VCSELs and detectors are on the bottom of the array, electrically connected to the transmission lines through the bonding bumps. The devices in the top row are detectors and the devices in the bottom row are VCSELs. The microlenses are $200\text{ }\mu\text{m}$ in diameter and the device spacing within a row is $500\text{ }\mu\text{m}$. The devices in the adjacent row are shifted both vertically and horizontally by $250\text{ }\mu\text{m}$.

4.4.5 DC DEVICE CHARACTERISTICS

Initial results from these devices were previously reported [64]. Figure 4.33 shows the output power versus current characteristics of several different size

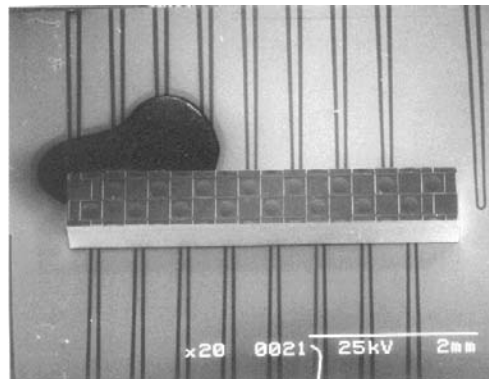


Figure 4.32. Scanning-electron-microscope image of an array of monolithically integrated VCSELs and detectors flip-chip bonded to coplanar transmission lines.

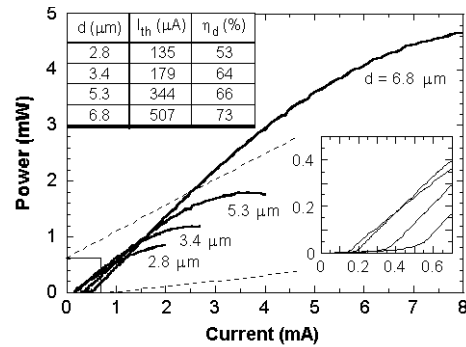


Figure 4.33. Light-current characteristics for several different size VCSELs. The inset shows the characteristics around threshold. The smallest VCSEL has a threshold current of 135 μA while maintaining a differential efficiency above 50% and the largest VCSEL has a differential efficiency above 70% with 0.5 mA threshold current.

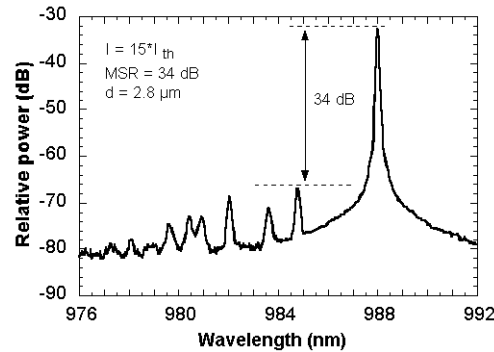


Figure 4.34. Lasing spectrum at 15 times threshold ($\sim 2 \mu\text{A}$) for the 2.8- μm VCSEL presented in Fig. 4.33 showing that even at this high bias current the device is single-mode with a mode suppression ratio of 34 dB.

VCSELs. The smallest, 2.8- μm -diameter VCSEL has the lowest threshold current ever reported for a three-QW VCSEL, 135 μA , while maintaining a high differential efficiency of 53%. The threshold voltage for this device was 2.5 V. The slightly larger, 3.4- μm -diameter VCSEL has a differential efficiency of 64% with a threshold current of 179 μA . The largest, 6.8- μm -diameter VCSEL has a differential efficiency above 70% with ~ 0.5 -mA threshold current and ~ 2 -V threshold voltage. Figure 4.34 shows the lasing spectrum at 15 times threshold for the 2.8- μm VCSEL. The device is single-mode with a mode suppression ratio of 34 dB. Based on the threshold currents and differential efficiencies of the VCSELs and Fig. 4.25 above, it is believed that the optical losses (round-trip) are fairly low, $\sim 0.2\%$, and the output mirror transmission is $\sim 1\%$.

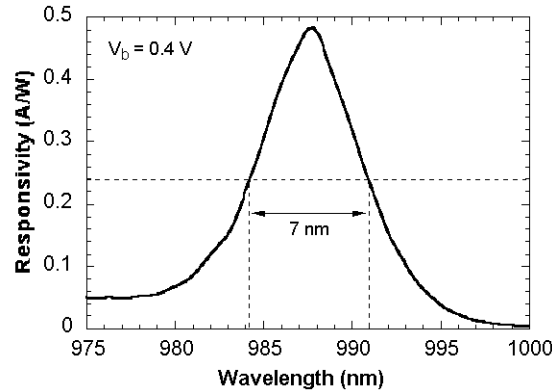


Figure 4.35. Detector responsivity spectrum at 0.4 V bias voltage for the 50- μm -diameter detector adjacent to the VCSEL in Fig. 4.34. The peak responsivity is 0.48 A/W, corresponding to a quantum efficiency of 60%, and the optical bandwidth is 7 nm. Note that the resonance wavelength of the detector is the same, to within measurement tolerances, as the lasing wavelength of the adjacent VCSEL.

Figure 4.35 shows the spectrum of the detector adjacent to the VCSEL in Fig. 4.34. The bias voltage for the detector was 0.4 V. The peak responsivity is 0.48 A/W, corresponding to a quantum efficiency of 60% and a FWHM optical bandwidth of 7 nm. This indicates that the input-mirror reflectivity of the detector is slightly less than expected and the absorption coefficient slightly higher than assumed. Note that the lasing wavelength of the VCSEL and resonance wavelength of the adjacent detector is the same to within measurement tolerances. The detectors exhibited high reverse-breakdown fields of 80 MV/m and low dark currents. The dark current of a 29- μm -diameter detector was ~ 80 pA at 1 V reverse bias, corresponding to a dark-current density of ~ 12 $\mu\text{A}/\text{cm}^2$. As mentioned above, the built-in field in these devices is quite large. Therefore the responsivity varies very little with reverse bias and is actually increased with a slight forward bias due to the associated reduction in the built-in field. For reverse biases between 0 and 5 V, the responsivity stays fairly constant, at ~ 0.35 A/W.

The above results clearly show that there is little, if any, penalty in device characteristics from using this design approach with a hybrid mirror for the VCSEL and having some additional (detuned) layers beneath the detector. In fact, the 135- μA threshold current is the lowest reported for a three-QW VCSEL and the detector characteristics are very close to the theoretically expected values. Using carbon doping instead of beryllium in the p-mirror and/or designing the interfaces between GaAs and $\text{Al}_{0.9}\text{Ga}_{0.1}\text{As}$ to minimize resistance, the threshold voltages (and dynamic resistances) of the VCSELs could be reduced significantly. In addition, by using tapered oxide apertures, the scattering losses can be further reduced, allowing the VCSELs to be scaled to smaller sizes [19, 46, 76].

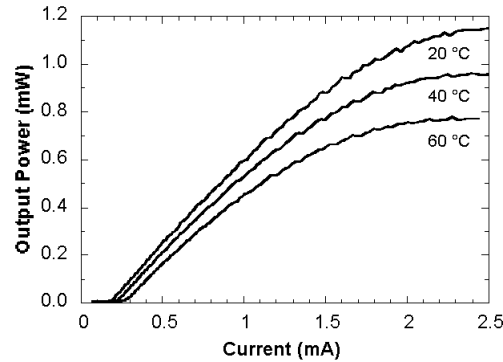


Figure 4.36. Light-current characteristics of a flip-chip bonded, 3.4- μm VCSEL at several temperatures.

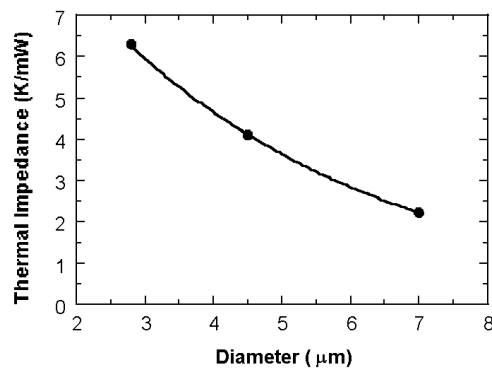


Figure 4.37. Thermal impedance as a function of VCSEL diameter.

Figure 4.36 shows the light-current characteristics at several temperatures of the 3.4- μm VCSEL used in the modulation and free-space link experiments discussed in the following sections. The VCSEL remained single-mode for all currents up to the rollover point. At 20°C, the VCSEL exhibited a threshold current of 179 μA , threshold voltage of 2.43 V, and a differential efficiency of 64%. At 60°C, the VCSEL shows only a modest decrease in performance with a threshold current of 265 μA , threshold voltage of 2.37 V, and differential efficiency of 56%. The voltages are fairly high since the structure did not include uniparabolic doping in the p mirror and Be was used instead of C.

Figure 4.37 shows the thermal impedance of flip-chip bonded VCSELs as a function of device size (diameter of aperture). The method for measuring the thermal impedance is discussed in more detail elsewhere [68]. The thermal impedances measured are similar to results previously reported for both VCSELs with oxidized mirrors [70] and VCSELs with all epitaxial mirrors [44].

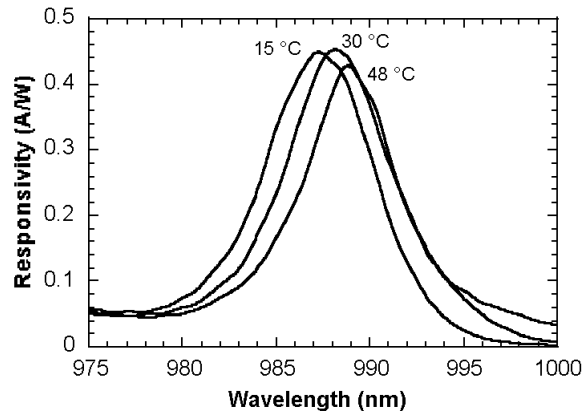


Figure 4.38. Responsivity spectrum at 0.4 V bias of a flip-chip bonded detector at several temperatures.

Figure 4.38 shows the responsivity spectrum at several temperatures for the 50- μm detector used in the modulation and free-space link experiments discussed in later sections. The detector exhibited a peak responsivity of 0.48 A/W (60% absorption) at 20°C. This responsivity stayed fairly constant for lower temperatures, but experienced a modest drop-off at temperatures above $\sim 40^\circ\text{C}$. The optical bandwidth, defined as the full width at half maximum, is a fairly broad 7.0 nm. This relatively broad optical bandwidth is important since it allows for variations in temperature or epitaxial growth between arrays. The wavelength of the peak responsivity of the detector is nearly the same, to within < 0.4 nm, as the lasing wavelength of adjacent VCSELs. The detectors exhibited low dark currents of less than 250 pA at 1 V reverse bias.

The rate at which the peak responsivity of the detector moves with temperature ($\partial\lambda/\partial T \approx 0.6 \text{ \AA}/^\circ\text{C}$) is nearly the same as the rate at which the lasing wavelength of the VCSEL moves with temperature ($\partial\lambda/\partial T \approx 0.6 \text{ \AA}/^\circ\text{C}$). The VCSEL mode probably moves at a slightly faster rate due to an increase in thermal impedance at higher temperatures, which causes the temperature difference between the substrate and junction to increase slightly, yielding a somewhat larger rate of change.

4.4.6 MODULATION CHARACTERISTICS

Coplanar probes were used to probe the transmission lines on the submounts. The collimated beam from the microlensed VCSEL was coupled into a fiber that went to a fiber-coupled 25 GHz detector followed by two 20-GHz, 15-dB amplifiers. The modulation response (S_{12}) was measured using a HP 8510C network analyzer. The small-signal modulation response was measured at different bias currents and temperatures. Figure 4.39(a) shows the small-signal modulation response of a 3.4- μm , single-mode VCSEL at 20°C, while Fig. 4.39(b) shows the

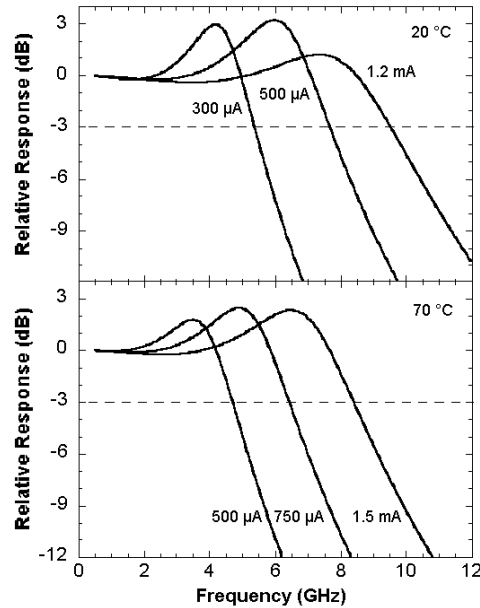


Figure 4.39. Small-signal modulation response of a 3.4-mm, single-mode VCSEL at (a) 20°C and (b) 70°C.

response at 70°C. Slightly higher bias currents are needed to reach a given 3 dB bandwidth at 70°C than at 20°C. The maximum 3-dB bandwidth is 9.5 GHz at 20°C and 8.4 GHz at 70°C.

These results represent the lowest reported bias currents required to reach bandwidths of up to 8 GHz. These results can be seen more clearly in Fig. 4.40(a), which shows a plot of 3-dB bandwidth versus bias current at both 20°C and 70°C. For dense, two-dimensional, free-space optical interconnects, power dissipation is a major concern. For this reason, VCSELs that can reach high enough bandwidths at very low biases are critical. The VCSEL in Fig. 4.39(a) reaches a 3-dB bandwidth of 7.8 GHz at a bias current of only 500 μA and a bandwidth of 5.3 GHz at 300 μA of bias. Smaller, lower-threshold-current VCSELs exhibited even higher bandwidths at low currents, with a 2.8- μm , 135- μA threshold VCSEL having a bandwidth of 6.3 GHz at 300 μA of bias current and 4.3 GHz at 200 μA of bias. Even at 70°C, sub-milliamp bias currents were needed to reach a 3-dB bandwidth of 7 GHz.

The modulation current efficiency factor (MCEF) is an effective way to compare the modulation characteristics of different VCSELs for free-space optical interconnects. The MCEF, defined as the slope of the linear region of the plot shown in Fig. 4.40(b), shows how well a device converts current above threshold into bandwidth. Thus, it is a good parameter for determining how well a device reaches high bandwidths at low bias currents. The 3.4- μm VCSEL has a MCEF of 15.3 GHz/ $\text{mA}^{1/2}$ at 20°C and a MCEF of 10.3 GHz/ $\text{mA}^{1/2}$ at 70°C. The 2.8- μm ,

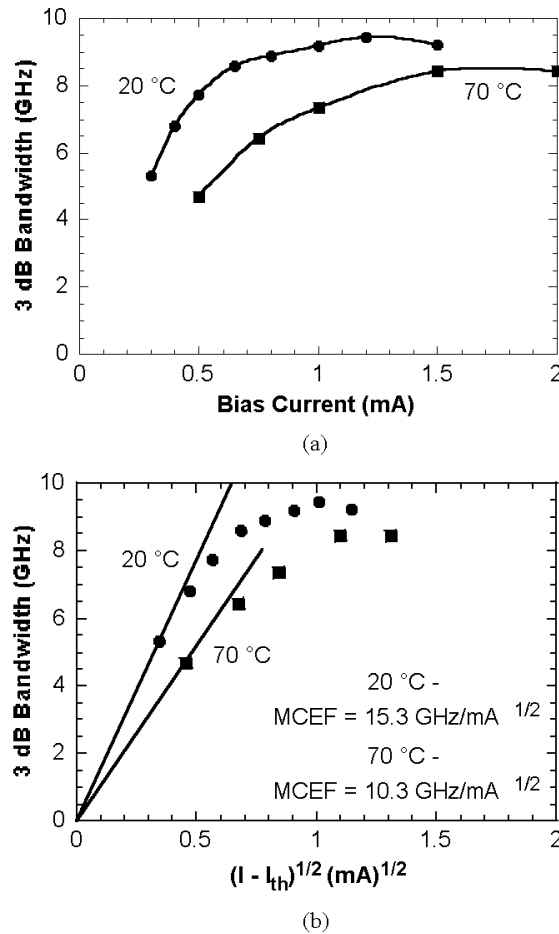


Figure 4.40. (a) 3-dB bandwidth versus bias current for a 3.4-mm VCSEL. (b) 3-dB bandwidth versus the square root of the bias current above threshold, which is used to determine the MCEF.

135- μ A-threshold-current VCSEL exhibited a MCEF of 16.8 GHz/mA^{1/2}, which ties the highest previously reported MCEF [60].

In order to understand the modulation results better and determine what is limiting the VCSEL performance, the modulation response at each bias current was fit with the three-parameter fit [77]

$$|H(f)|^2 = \frac{1}{1 + \left(\frac{f}{f_0}\right)^2} \cdot \frac{f_R^4}{(f_R^2 - f^2)^2 + \left(\frac{\gamma}{2\pi}\right)^2 f^2}, \quad (4.5)$$

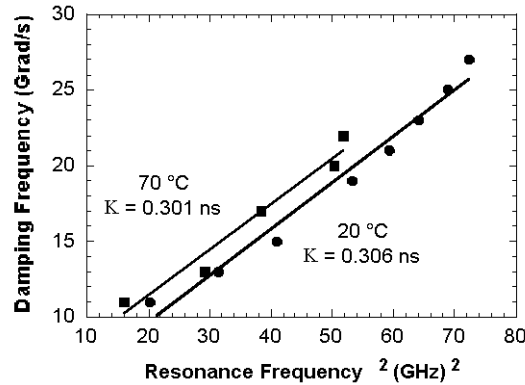


Figure 4.41. Damping frequency versus the square of the resonance frequency for a 3.4-mm VCSEL.

where f_0 is the parasitic frequency, f_R is the resonance frequency of the VCSEL, and γ is the damping frequency of the VCSEL. Figure 4.41 is a plot of the damping frequency versus the square of the resonance frequency at both 20°C and 70°C. The slope of this curve is defined as the K -factor of the laser and describes the damping of the response for large resonance frequencies. The maximum intrinsic 3-dB bandwidth, neglecting heating and parasitics, can be determined from the K -factor,

$$f_{3dB(max)} = \frac{2\pi\sqrt{2}}{K} \quad (4.6)$$

Using the K -factors obtained from Fig. 4.41, the maximum intrinsic 3-dB bandwidth, neglecting heating and parasitics, is found to be ~29 GHz at both 20°C and 70°C. This shows that the maximum bandwidth of the VCSELs is not intrinsically limited, and therefore must be limited by either heating or parasitics.

Although the K -factor indicates that the maximum intrinsic 3-dB bandwidth is well above the maximum bandwidths that were measured, it doesn't identify the mechanism limiting the performance. This limitation could be caused by either parasitics or heating which increases the threshold current and reduces the differential gain. Figure 4.42(a) is a plot of the 3-dB frequency versus the resonance frequency at both 20°C and 70°C. The bends on the right sides of the curves indicates that the limiting factor is parasitics [78]. This reduction in 3-dB bandwidth with increasing resonance frequency was observed for all devices tested.

Figure 4.42(b) shows a plot of the parasitic frequency versus bias current at 20°C and 70°C. The parasitic frequency varies between approximately 2.5 GHz and 5 GHz over these temperatures and bias currents. This parasitic frequency is caused by parasitic capacitances that lead to a RC roll-off. The VCSELs not only

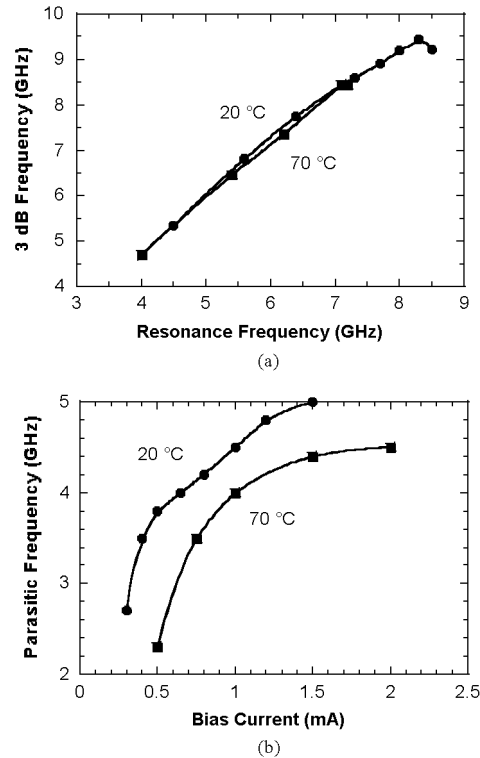


Figure 4.42. (a) VCSEL 3-dB bandwidth versus resonance frequency. The bend on the right indicates the bandwidth is parasitic limited. (b) Parasitic frequency as a function of the DC bias current for a 3.4-mm VCSEL.

have a capacitance associated with the diode, but also have capacitances due to the pad metals and oxide aperture. The VCSELs also have resistances distributed between these capacitances. These resistances comprise the differential (dynamic) resistance of the VCSEL, which changes significantly with bias current and temperature. These changes in differential resistance lead to changes in the RC roll-off (parasitic) frequency, and thus explain why the parasitic frequencies in Fig. 4.42(b) are not constant.

There are numerous ways to improve the parasitics of the VCSELs. Reducing the diameter of the VCSEL pillars would reduce the diode and aperture capacitances since the 22- μm diameter pillars used in this work are much larger than are necessary. An improved doping scheme could lower the differential resistance and improved isolation could lower the pad capacitances. Using a thicker aperture could also reduce the aperture capacitance. However, this would lead to higher optical scattering losses and diminish the performance of the VCSEL [19]. To get around this problem, a tapered aperture could be used [46]. A tapered

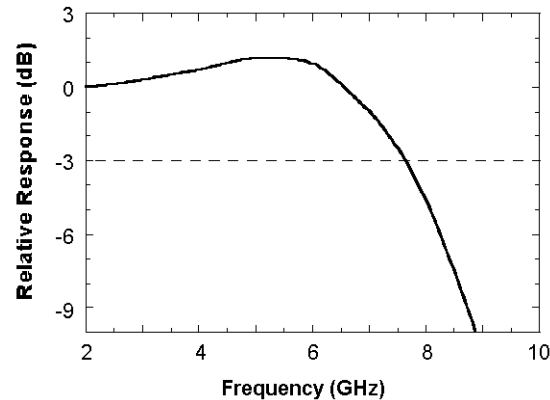


Figure 4.43. Large-signal modulation response of a 3.4-mm VCSEL at 20°C.

aperture has the advantage of not only having lower optical scattering losses than even thin apertures [46], but also has a lower capacitance due to the tapered oxide being significantly thicker away from the aperture-oxidation front, which is most of the area.

To reduce the bias currents required to reach a given bandwidth even further, smaller VCSELs are needed. However, smaller devices will not help if their parameters don't scale with their size. If size-dependent losses were not an issue, making a device with a higher MCEF would not be a problem, since the MCEF is inversely proportional to the VCSEL diameter. The thin-aperture devices presented here have low optical loss and scale well down to about 3.5 μm , at which point optical scattering losses from the aperture begin to have a significant effect. Therefore, although a 2.8- μm VCSEL demonstrated the highest reported MCEF (16.8 GHz/ $\text{mA}^{1/2}$), it could have been even higher (18.6 GHz/ $\text{mA}^{1/2}$) if size-dependent losses had not been present. Tapered apertures may hold the key to increasing VCSEL MCEF and lowering the required bias currents. Tapered apertures have been shown to have the lowest optical loss of any aperture [19] and lower optical losses can lead to higher differential gains and improved modulation characteristics. In addition, tapered apertures scale better than other apertures, allowing device characteristics to be scaled down to smaller sizes [46–47] and, as was previously discussed, tapered apertures can also be used to reduce the parasitic capacitance of VCSELs.

Although the small-signal modulation of VCSELs has been presented, the large-signal modulation is also of interest since that is what would be used in a real system. Figure 4.43 shows the large-signal modulation response of the same VCSEL as Fig. 4.39 at 20°C. For this measurement, the VCSEL was biased at 550 μA and was modulated $\pm 450 \mu\text{A}$. This resulted in a current swing from 100 μA to 1 mA. Since the threshold current of the VCSEL is 179 μA , it was modulated well

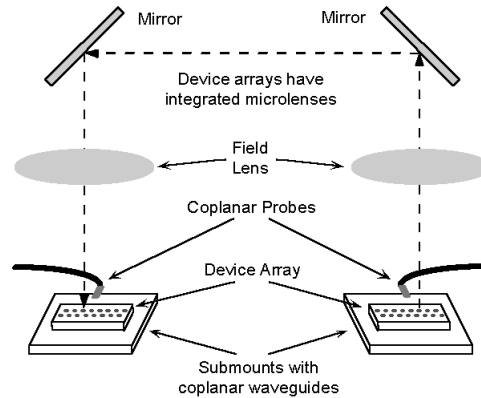


Figure 4.44. Schematic of the experimental free-space optical link.

below threshold. The resulting optical modulation was from essentially zero (spontaneous emission) to $600\ \mu\text{W}$. The 3-dB bandwidth of the VCSEL with this large-signal modulation was found to be 7.7 GHz, which is just slightly less than the small-signal bandwidth with the same average current ($550\ \mu\text{A}$).

The bandwidth of the detectors was also measured using this VCSEL as a source. At 0 V bias, a $50\text{-}\mu\text{m}$ -diameter detector was found to have a 3-dB bandwidth of ~ 2.5 GHz and a $25\text{-}\mu\text{m}$ -diameter detector was found to have a 3-dB bandwidth of ~ 9 GHz. These results agree well with values calculated from the intrinsic region thickness and area.

4.4.7 FREE-SPACE LINK

Figure 4.44 shows a schematic of the experimental free-space optical link setup. Two flip-chip bonded arrays were placed on different temperature stages with a VCSEL from one array and a detector from the other array contacted using high-speed coplanar probes. Field lenses are placed above the microlenses for both the VCSEL and detector to extend the collimated beams. Folding mirrors are used to direct the collimated beams across the optical table from one array to the other. Bias tees were connected to each coplanar probe, allowing the VCSEL and detector to be biased. Amplifiers were placed at the receiver end to boost the signal level for bit-error-rate (BER) measurements. The optical cross talk between adjacent receiver channels was determined by measuring the received power with the optical beam well aligned to the desired detector, and then measuring the received power at the adjacent detector without moving the optical beam. The optical cross talk between several different adjacent channels was measured and found to be -39 to -40 dB for all cases.

An optical attenuator was placed in the optical beam between the two folding mirrors so that the received optical power could be varied for the BER measure-

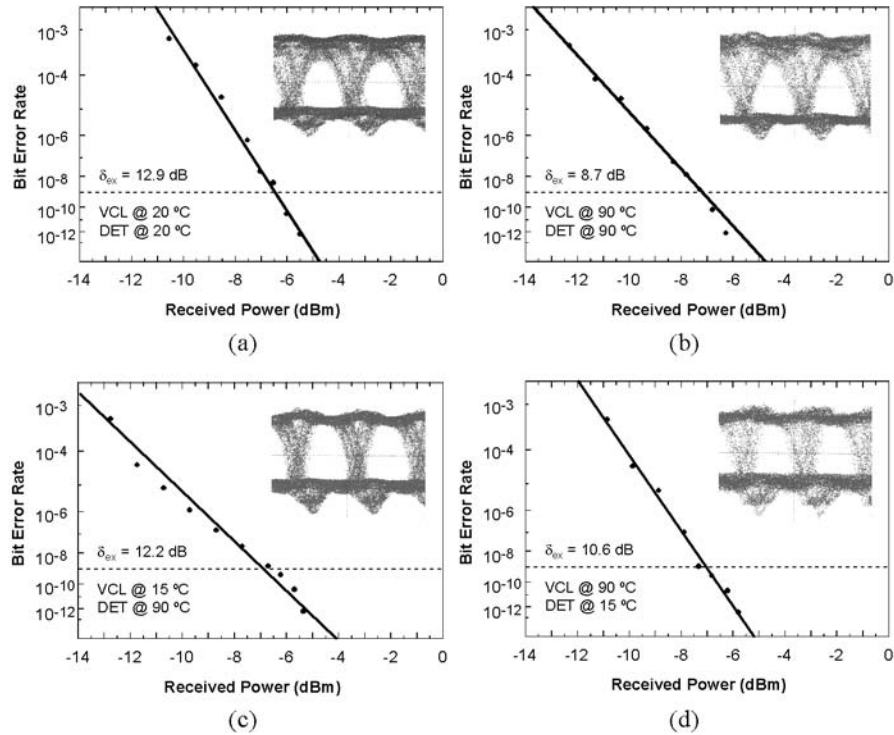


Figure 4.45. Bit error rate versus received optical power at 1.25 Gbit/s with (a) both the VCSEL and detector at 20°C, (b) both the VCSEL and detector at 90°C, (c) the VCSEL at 15°C and the detector at 90°C, and (d) the VCSEL at 90°C and the detector at 15°C. The inset eye diagrams are for error-free (BER < 10^{-12}) data transmission in each case and correspond to the same electrical power to the transmitter, not the same bit error rate. The extinction ratio power penalty, δ_{ex} , is also shown for each case.

ments. Figure 4.45 shows the BER versus received optical power at 25 Gbit/s for four different combinations of transmitter temperature and receiver temperature. The insets show the eye diagrams for error-free (BER < 10^{-12}) data transmission in each case. The eye diagrams were all taken with the same electrical power applied to the transmitter, but correspond to different bit error rates due to the different cases having different differential resistances and noise characteristics. Figure 4.45(a) shows both the VCSEL and detector at 20°C, Fig. 4.45(b) shows both the VCSEL and detector at 90°C, Fig. 4.45(c) shows the VCSEL at 15°C and the detector at 90°C, while Fig. 4.45(d) shows the VCSEL at 90°C and the detector at 15°C. In each case, error-free data transmission was demonstrated, with the system sensitivity, defined as the received optical power, at a BER of 10^{-9} , between -6.5 and -7.3 dBm. No error floor was observed for any combination of transmitter and receiver temperatures. The slope of the BER curves in Figs. 4.45(b) and

4.45(c) are not as steep as those in Figs. 4.45(a) and 4.45(d). This difference in the slope of the BER curves is attributed to the difference in the temperature of the detector, which leads to different noise characteristics. It is also important to note that no realignments or changes in biases were done between the measurements.

The system sensitivities are fairly low because the transmitted signal has a low extinction ratio, defined as $r_{ex} = P_1/P_0$, where P_1 is the power transmitted in a 1 bit and P_0 is the power transmitted in a 0 bit [79]. The extinction ratio was low because our source could only provide a small voltage swing and impedance mismatches between the transmission line and VCSEL led to reflections that reduced the electrical power incident on the VCSEL. This low extinction ratio leads to a significant power penalty (δ_{ex}) in the system sensitivities [79]. The extinction ratio varied between 1.11 and 1.32 for each of the temperature combinations. The extinction ratio varies in each case because in addition to changes in bias currents, at different temperatures the VCSEL has slightly different threshold currents and differential efficiencies, as well as different differential resistances. Therefore, although the bias current was ~ 1.2 mA in each case, the transmitted DC optical power varied between ~ 400 μ W and ~ 700 μ W. Changes in the differential resistance also caused the modulated power to vary between ~ 70 μ W and ~ 120 μ W. These changes lead to power penalties in each case that vary between 12.9 dB for the case where both the VCSEL and detector are at 20°C to 8.7 dB for the case where both the VCSEL and detector are at 90°C. Taking these power penalties into account, the receiver sensitivity with an ideal extinction ratio ($r_{ex} = \infty$) can be determined. The receiver sensitivity, with infinite extinction ratio, is the best for the case where both the VCSEL and detector are at 20°C where it is ~ -19 dBm. As expected, the worst receiver sensitivity is for the case where both the VCSEL and detector are at 90°C. In this case the receiver sensitivity is ~ -16 dBm, indicating a power penalty of about 3 dB due to the increased noise in the VCSEL and detector at the higher temperature. The system performance could be significantly improved by better impedance matching at the transmitter and providing higher input signal levels.

The tolerance of a free-space optical link to changes in temperature, including temperature differences between arrays, is a crucial aspect of free-space interconnect design and characterization. To realize a temperature tolerant link, not only must the devices themselves have good temperature characteristics, but the wavelengths of the detector and VCSEL must also be sufficiently aligned. With a temperature difference between arrays in the link, the lasing wavelength of the VCSELs will be different from that of the peak responsivities of the detectors. When the detectors are resonant cavity detectors, this wavelength mismatch results in a reduction in the responsivity. This leads to a reduction in receiver current and therefore a lower signal-to-noise ratio at the receiver. This is the reason that a broad optical bandwidth is required for temperature-tolerant systems. The fact that the link was able to tolerate temperature differences of $\pm 75^\circ\text{C}$ between

the VCSEL and detector indicates that the detectors have sufficient optical bandwidth for these temperature differences. In addition, it was demonstrated that both the VCSELs and detectors have adequate performance, including noise, at temperatures up to 90°C.

Although this free-space link used a 50- μm -diameter detector with a 3-dB bandwidth of ~ 2.5 GHz, this was not the bandwidth limitation for the link. The bandwidth of the free-space link was limited by the bandwidth of the amplifiers used. The combination of amplifiers in the system had a 3-dB bandwidth of ~ 1.2 GHz. Even with a 25- μm -diameter detector (3-dB bandwidth of ~ 9 GHz), the system would have still had approximately the same maximum bit rate. As shown in the previous section, the integrated VCSELs and detectors both have sufficient bandwidth for data transmission at bit rates well above what we demonstrated. To reach higher bit rates, only higher bandwidth amplifiers would be needed.

4.5 WDM PIE-SHAPED VCSEL ARRAYS

The revolution of microprocessors has drastically increased the capability of data processing, and now it gives rise to more severe requirements for computer interconnects and local area networks. The networks have to transmit the data much more quickly while costing less to build, manage, or reconfigure. Recent advancements in multimode fiber (MMF) links have proven that they are promising candidates to satisfy these demands [28]. These needs now attract increasing interest in WDM configurations for the short-haul networks [29–30]. WDM not only increases data capacity in the networks, but it also makes their construction and reconfiguration easier and cheaper. Multiple-wavelength VCSEL arrays for the WDM transmitters of the MMF networks [31–32] have been demonstrated. The array is designed to be coupled directly into a single MMF without any additional optics, so that in each array eight VCSELs are arranged within a 60- μm circle with a 3- μm spacing.

In such a densely packed array, cross talk among the VCSELs is an important criteria to be investigated because each VCSEL has to be operated independently in order to increase the aggregate transmission rate of one array or to provide eight independent channels in a WDM network. Three sources of cross talk have to be considered in the VCSEL arrays: thermal, electrical, and optical sources. The cross talk in VCSEL arrays with 3- μm pitches has been reported regarding the thermal cross talk at DC [80–81]. RF-cross-talk measurements have been performed between VCSELs separated by 1.6 mm in 16-wavelength VCSEL arrays at frequencies up to 10 GHz [82]. The observed cross talk is less than -35 dB below 1 GHz, and between -10 and -30 dB at 5 GHz. The optical as well as thermal cross talks in these VCSEL arrays in a frequency range from 1 Hz to 10 MHz have been investigated [83]. The optical cross talk between the adjacent VCSELs was

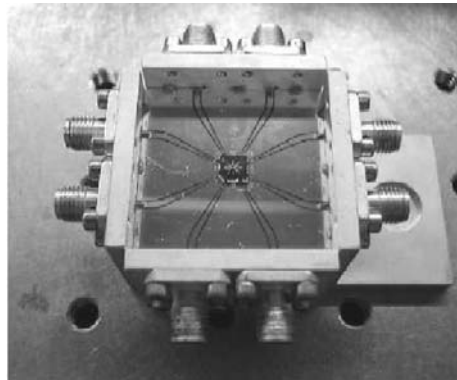


Figure 4.46. Photograph of the test fixture for the RF-cross-talk measurement. In the test fixture, a multiple-wavelength VCSEL array is flip-chip bonded onto a GaAs submount, which is put on an AlN submount.

reduced to less than -20 dB by biasing both of them near or above threshold, while thermal cross talk was found to be insignificant in our array. In this section, cross talk in multiple-wavelength VCSEL arrays under higher frequency RF modulation is investigated.

4.5.1 PACKAGING AND MEASUREMENT SETUP

The multiple-wavelength VCSEL array [31] is flip-chip bonded onto a semi-insulating GaAs submount with $50\text{-}\Omega$ coplanar lines. The GaAs submount is put on an AlN submount, which also has $50\text{-}\Omega$ coplanar lines, and their coplanar lines are wire bonded together. The GaAs submount is used to make exchanging VCSEL arrays easier. The flip-chip bonding eliminates series inductance due to bonding wires between the p-contacts of the VCSELs and the coplanar lines, and minimizes external electrical cross talk [84–85]. It also provides good heat sinking, which reduces thermal cross talk. The AlN submount is set on a test fixture, which has eight SMA connectors. Figure 4.46 shows a photograph of the test fixture with the VCSEL array mounted on the submounts.

Figure 4.47 shows the measurement setup employed in this experiment. The cross-talk measurement was performed by detecting the optical signals from VCSEL A when either VCSEL A or B is modulated. Both VCSELs are biased near or above threshold in order to minimize their optical cross talk due to modulated spontaneous emission [83], and they are modulated with RF signals from a network analyzer (HP 8510C). The optical signals are collected with a microscope objective lens, and coupled into a MMF. The MMF is connected to an optical spectrum analyzer (HP 70951A), which has the capability to output the optical signal through its monochromator. This provides the optical signal from only

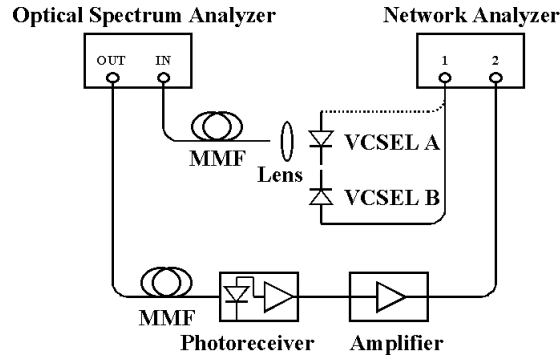


Figure 4.47. Schematic of the RF-cross-talk measurement setup. An RF signal from a network analyzer modulates either VCSEL A or VCSEL B, both of which are biased near or above threshold.

VCSEL A and cuts off the other. The lasing-wavelength spacing of the VCSELs is 3.0 nm on average, and with a 0.5-nm bandwidth of the monochromator, the isolation between the adjacent wavelengths is better than -50 dB in optical power.

The optical signal coming out of the OSA is detected with a high-speed photoreceiver and an amplifier, and is measured by the network analyzer. The photoreceiver (New Focus 1534) had a high conversion gain of 250 V/W and the amplifier (New Focus 1422) had a low noise figure of 8 dB. The measured noise floor of the network analyzer with the photoreceiver and the amplifier connected is less than -70 dBm below 2 GHz, and -100 dBm above that frequency. Another measurement was performed using a microwave spectrum analyzer (HP 8566B) to increase the dynamic range, especially in the frequency range below 2 GHz, while the network analyzer is still used as an RF signal source. The spectrum analyzer has a variable-resolution bandwidth, and the noise floor can be reduced by narrowing this resolution bandwidth. With its minimum bandwidth of 10 Hz, a -110 -dBm noise floor was achieved in the whole frequency range.

4.5.2 CROSS TALK

The light output versus current characteristics, lasing spectra, and RF-modulation responses are measured for all the VCSELs in the mounted array. Compared with the results for the unmounted array [31–32], no significant change is observed for the mounted one. The cross talk measurement is performed with two VCSELs, which are referred to as VCSELs #4 and #5. They are neighboring both in geometry and in wavelength, which are 976 nm and 980 nm for VCSELs #4 and #5, respectively. The threshold current is 2.0 mA for both VCSELs #4 and #5, and the 3-dB bandwidth and resonant frequency are 3.8 GHz and 2.4 GHz for VCSEL #4, and 3.1 GHz and 2.0 GHz for VCSEL #5 at a bias current of 5.0 mA.

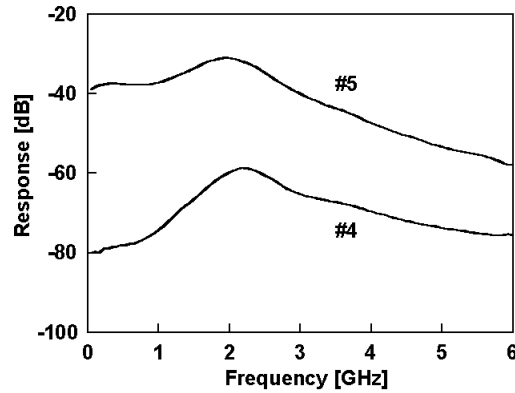


Figure 4.48. Modulation responses of VCSEL #5: one is the response when VCSEL #5 is modulated (labeled as #5), and the other one is for when VCSEL #4 is modulated (labeled as #4).

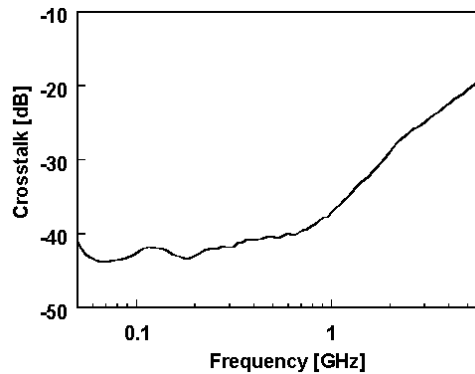


Figure 4.49. Cross talk from VCSEL #4 to VCSEL #5, which is derived from the results shown in Fig. 4.48.

Both VCSELs #4 and #5 are biased at 5.0 mA to achieve a 3-dB bandwidth higher than 3.0 GHz. The wavelength of the monochromator is set to the lasing wavelength of VCSEL #5, that is, 979.9 nm. Figure 4.48 shows two modulation responses of VCSEL #5 in the frequency range from 50 MHz to 6 GHz: one is the response when VCSEL #5 is modulated (labeled as #5), and the other one is that when VCSEL #4 is modulated (labeled as #4). Their difference is the cross talk from VCSEL #4 to VCSEL #5, which is shown in Fig. 4.49. It linearly increases with frequency above 1 GHz. It is less than -30 dB below 1.9 GHz, and is -25 dB at the 3-dB bandwidth of VCSEL #5. This value is comparable with the cross talk observed in the array with much larger spacing (1.6 mm) [82]. The VCSELs studied in this paper are isolated from each other by deep trenches, which are etched through the active regions, whereas previously-reported VCSELs were isolated by

proton implantation [82]. The trenches prevent electrical coupling between the adjacent VCSELs in the densely-packed array.

4.6 SUMMARY

The issues involved in scaling VCSELs to smaller sizes and the advantages of utilizing tapered apertures are discussed. Device and modeling results were presented for tapered apertures at the first null, first peak, and second null of the standing wave. Experimental results show that tapered apertures nearly eliminate optical scattering losses from the aperture, even for very small devices ($\sim 1\ \mu\text{m}$). When the tapered aperture is placed at the first null of the standing wave, current spreading can also be greatly reduced, yielding small, yet high-performance devices.

A process for the formation of refractive microlenses in GaAs and InP was developed which relies on transfer of a lens-shaped erodable mask by RIE. The ability to control the ratio of the etch rate in the semiconductor to that in the erodable mask material allows greater flexibility of the resulting radius of curvature of the microlens. Integration of microlenses with VCSELs and with back-illuminated photodiodes for focusing onto the detector [51] demonstrates their applicability in optoelectronics.

The design, fabrication, packaging, and characterization of monolithically integrated arrays of VCSELs and detectors were discussed. The integrated arrays of devices are flip-chip bonded and have substrate-side refractive microlenses etched into the substrate to collimate the VCSEL outputs and focus collimated beams on the detectors. The VCSELs exhibit low threshold currents and high efficiencies. Devices with threshold currents as low as $135\ \mu\text{A}$ were obtained in single-mode devices with differential efficiencies of 53%. The detectors had responsivities of $0.48\ \text{A/W}$ with 7-nm optical bandwidths. Modulation and free-space link results were also presented, demonstrating the potential of the integrated devices for free-space board-to-board interconnects.

Cross talk was measured in multiple-wavelength VCSEL arrays with a $3\text{-}\mu\text{m}$ spacing which are designed to be coupled into a single MMF. In the densely-packed VCSEL arrays, the cross talk is below $-20\ \text{dB}$ at the 3-dB bandwidth of the VCSELs. This means every VCSEL in each array can be operated independently up to its 3-dB bandwidth in relatively low dynamic range applications. Digital communication is one example, and the multiple-wavelength VCSEL array studied here can increase the data capacity or provide other benefits of WDM networks.

REFERENCES

- [1] M. R. Feldman, S. C. Esener, C. C. Guest, and S. H. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations," *Appl. Opt.*, Vol. 27, pp. 1742–1751 (1988).
- [2] A. V. Krishnamoorthy and D. A. B. Miller, "Scaling optoelectronic-VLSI circuits into the 21st century: a technology roadmap," *IEEE J. Sel. Topics Quantum Electron.*, Vol. 2, pp. 55–76 (1996).
- [3] F. A. P. Tooley, "Challenges in optically interconnecting electronics," *IEEE J. Sel. Topics Quantum Electron.*, Vol. 2, pp. 3–13 (1996).
- [4] G. A. Betzos and P. A. Mitkas, "Performance evaluation of massively parallel processing architectures with three-dimensional optical interconnects," *Appl. Opt.*, Vol. 37, pp. 315–325 (1998).
- [5] M. W. Haney and M. P. Christensen, "Performance scaling comparison for free-space optical and electrical interconnection approaches," *Appl. Opt.*, Vol. 37, pp. 2888–2894 (1998).
- [6] G. I. Yayla, P. J. Marchand, and S. C. Esener, "Speed and energy analysis of digital interconnections: comparison of on-chip, off-chip, and free-space technologies," *Appl. Opt.*, Vol. 37, pp. 205–227 (1998).
- [7] D. M. Cutrer and K. Y. Lau, "Ultralow power optical interconnect with zero-biased, ultralow threshold laser—How low a threshold is low enough?," *IEEE Photon. Technol. Lett.*, Vol. 7, pp. 4–6 (1995).
- [8] D.-S. Kim, J. C. Dries, M. R. Gokhale, and S. R. Forrest, "Optoelectronic InP-InGaAs smart pixels for optical interconnections and computing," *IEEE J. Quantum Electron.*, Vol. 33, pp. 1407–1416 (1997).
- [9] F. B. McCormick, F. A. Tooley, T. J. Cloonan, J. M. Sasian, H. S. Hinton, K. O. Mersereau, and A. Y. Feldblum, "Optical interconnections using microlens arrays," *Opt. Quantum Electron.*, Vol. 24, pp. 465–477 (1992).
- [10] E. M. Strzelecka, D. A. Loudereback, B. J. Thibeault, G. B. Thompson, K. Bertilsson, and L. A. Coldren, "Parallel free-space optical interconnect based on arrays of vertical cavity lasers and detectors with monolithic microlenses," *Appl. Opt.*, Vol. 37, pp. 2811–2821 (1998).
- [11] D. S. Wills, W. S. Lacy, C. Camperi-Ginestet, B. Buchanan, H. H. Cat, S. Wilkinson, M. Lee, N. M. Jokerst, and M. A. Brooke, "A three-dimensional high-throughput architecture using through-wafer optical interconnect," *IEEE J. Lightwave Technol.*, Vol. 13, pp. 1085–1092 (1995).
- [12] J. J. Yoo, J. E. Leight, C. Kim, G. Giaretta, W. Yuen, A. E. Willner, and C. J. Chang-Hasnain, "Experimental demonstration of a multihop shuffle network using WDM multiple-plane optical interconnection with VCSEL and MQW/DBR detector arrays," *IEEE Photon. Technol. Lett.*, Vol. 10, pp. 1507–1509 (1998).

- [13] K. Hirabayashi, T. Yamamoto, S. Matsou, and S. Hino, "Board-to-board free-space optical interconnections passing through boards for a bookshelf-assembled terabit-per-second-class ATM switch," *Appl. Opt.*, Vol. 37, pp. 2985–2995 (1998).
- [14] Y. Liu, B. Robertson, G. C. Boisset, M. H. Ayliffe, R. Iyer, and D. V. Plant, "Design, implementation, an characterization of a hybrid optical interconnect for a four-stage free-space optical backplane demonstrator," *Appl. Opt.*, Vol. 37, pp. 2895–2914 (1998).
- [15] D. V. Plant, B. Robertson, H. S. Hinton, M. H. Ayliffe, G. C. Boisset, W. Hsiao, D. Kabal, N. H. Kim, Y. S. Liu, M. R. Otazo, D. Pavlasek, A. Z. Shang, J. Simmons, K. Song, D. A. Thompson, and W. M. Robertson, " 4×4 vertical-cavity surface-emitting laser (VCSEL) and metal-semiconductor-metal (MSM) optical backplane demonstrator system," *Appl. Opt.*, Vol. 35, pp. 6365–6368, (1996).
- [16] J.-H. Yeh, R. D. Kostuk, and K.-Y. T. Tu, "Hybrid free-space optical bus system for board-to-board interconnections," *Appl. Opt.*, Vol. 35, pp. 6354–6364 (1996).
- [17] P. A. Mitkas, L. J. Irakliotis, J. F. R. Beyette, S. A. Feld, and C. W. Wilmsen, "Optoelectronic data filter for selection and projection," *Appl. Opt.*, Vol. 13, pp. 1074–1084 (1994).
- [18] A. C. Walker, T.-Y. Yang, J. Gourlay, J. A. B. Dines, M. G. Forbes, S. M. Prince, D. A. Baillie, D. T. Nelson, R. Williams, L. C. Wilkinson, G. R. Smith, M. P. Y. Desmulliez, G. S. Buller, M. R. Taghizadeh, A. Waddie, I. Underwood, C. R. Stanley, F. Pottier, B. Vögele, and W. Sibbett, "Optoelectronic systems based on InGaAs-complementary-metal-oxide-semiconductor smart-pixel arrays and free-space optical interconnects," *Appl. Opt.*, Vol. 37, pp. 2822–2830 (1998).
- [19] E. R. Hegblom, D. I. Babic, B. J. Thibeault, and L. A. Coldren, "Scattering losses from dielectric apertures in vertical cavity lasers," *IEEE J. Selected Topics in Quantum Electron.*, Vol. 3, pp. 379–389 (1997).
- [20] R. L. Naone, P. D. Floyd, D. B. Young, E. R. Hegblom, T. A. Strand, L. A. Coldren, "Interdiffused quantum wells for lateral carrier confinement in VCSELs," *IEEE J. Selected Topics in Quantum Electron.*, Vol. 4, pp. 706–714 (1998).
- [21] M. Ghisoni, J. Bengtsson, J. A. Vukusic, H. Martinsson, and A. Larsson, "Single- and multimode VCSELs operating with continuous relief kinoform for focused spot-array generation," *IEEE Photon. Tech. Lett.*, Vol. 9, pp. 1466–1468 (1997).
- [22] B. Dhoedt, P. D. Dobbelaere, J. Blondelle, P. V. Daele, P. Demeester, and R. Baets, "Monolithic integration of diffractive lenses with LED-arrays for board-to-board free space optical interconnects," *IEEE J. Lightwave Technol.*,

- Vol. 13, pp. 1065–1073 (1995).
- [23] E. M. Strzelecka, G. D. Robinson, L. A. Coldren, and E. L. Hu, “Fabrication of refractive microlenses in semiconductors by mask shape transfer in reactive ion etching,” *Microelectronic Eng.*, Vol. 35, pp. 385–388 (1997).
 - [24] M. K. Hibbs-Brenner, Y. Liu, R. Morgan, and J. Lehman, “VCSEL/MSM detector smart pixel arrays,” *Proc. LEOS Summer Topical Meeting on Smart Pixels*, pp. 3–4 (1998).
 - [25] P. Kazlas, J. Wasserbauer, J. Scott, D. Paananen, S. Swirhun, and D. Lewis, “Monolithic vertical-cavity laser/p-i-n photodiode transceiver array for optical interconnects,” *IEEE Photon. Technol. Lett.*, Vol. 10, pp. 1530–1532 (1998).
 - [26] S. Q. Luong, G. G. Ortiz, Y. Zhou, J. Lu, C. P. Hains, J. Cheng, H. Q. Hou, and G. A. Vawter, “Monolithic wavelength graded VCSEL and resonance-enhanced photodetector arrays for parallel optical interconnects,” *IEEE Photon. Technol. Lett.*, Vol. 10, pp. 642–644 (1998).
 - [27] G. G. Ortiz, C. P. Hains, J. Cheng, H. Q. Hou, and J. C. Zolper, “Monolithic integration of InGaAs VCSELs with resonance-enhanced quantum-well photodetectors,” *Electron. Lett.*, Vol. 32, pp. 1205–1207 (1996).
 - [28] L. Raddatz, D. Hardacre, I. H. White, R. V. Pentty, D. G. Cunningham, M. R. T. Tan and S. Y. Wang, “High bandwidth data transmission in multi-mode fibre links using subcarrier multiplexing with VCSELs,” *Electron Lett.*, Vol. 34, pp. 686–688 (1998).
 - [29] L. B. Aronson, B. E. Lemoff, L. A. Buckman, and D. W. Dolfi, “Low-cost multimode WDM for local area networks up to 10 Gb/s,” *IEEE Photon. Technol. Lett.*, Vol. 10, pp. 1489–1491 (1998).
 - [30] J. J. Yoo, J. E. Leight, C. Kim, G. Giaretta, W. Yuen, A. E. Willner, and C. J. Chang-Hasnain, “Experimental demonstration of a multihop shuffle network using WDM multiple-plane optical interconnection with VCSEL and MQW/DBR detector arrays,” *IEEE Photon. Technol. Lett.*, Vol. 10, pp. 1507–1509 (1998).
 - [31] S. Y. Hu, J. Ko, and L. A. Coldren, “High-performance densely packed vertical-cavity photonic integrated emitter arrays for direct-coupled WDM applications,” *IEEE Photon. Technol. Lett.*, Vol. 10, pp. 766–768 (1998).
 - [32] S. Y. Hu, S. Z. Zhang, J. Ko, J. E. Bowers, and L. A. Coldren, “5 Gbit/s/channel operation of multiple-wavelength vertical-cavity photonic integrated emitter arrays for low-cost multimode WDM local-area networks,” *Electron. Lett.*, Vol. 34, pp. 768–769 (1998).
 - [33] R. L. Thornton, “Vertical cavity lasers for printing,” *Proc. LEOS Summer Topical Meeting*, p. 71 (1997).
 - [34] K. L. Lear and R. P. Schneider, “Uniparabolic mirror grading for vertical cavity surface emitting lasers,” *Appl. Phys. Lett.*, Vol. 68, pp. 605–607

- (1996).
- [35] B. J. Thibeault, E. R. Hegblom, Y. A. Akulova, P. D. Floyd, J. Ko, R. Naone, and L. A. Coldren, "Electrical and optical losses in dielectrically-apertured vertical-cavity lasers," *Proc. SPIE Photonics West*, Vol. 3003, pp. 86–99 (1997).
 - [36] B. J. Thibeault, T. A. Strand, T. Wipiejewski, M. G. Peters, D. B. Young, S. W. Corzine, L. A. Coldren, and J. W. Scott, "Evaluating the effects of optical and carrier losses in etched post vertical cavity lasers," *J. Appl. Phys.*, Vol. 78, pp. 5871–5876 (1995).
 - [37] B. J. Thibeault, E. R. Hegblom, P. D. Floyd, R. Naone, Y. Akulova, and L. A. Coldren, "Reduced optical scattering loss in vertical-cavity lasers using a thin (300Å) oxide aperture," *IEEE Photon. Tech. Lett.*, Vol. 8, pp. 593–596 (1996).
 - [38] E. R. Hegblom, B. J. Thibeault, R. L. Naone, and L. A. Coldren, "Vertical Cavity Lasers with tapered oxide apertures for low scattering loss," *Electron. Lett.*, Vol. 33, pp. 869–879 (1997).
 - [39] E. R. Hegblom, N. M. Margalit, B. J. Thibeault, L. A. Coldren, and J. E. Bowers, "Current spreading in apertured vertical cavity lasers," *Proc. SPIE Photonics West*, Vol. 3003, pp. 176–180 (1997).
 - [40] B. J. Thibeault, K. Bertilsson, E. R. Hegblom, E. M. Strzelecka, Y. Akulova, and L. A. Coldren, "High speed characteristics of low optical loss oxide-apertured vertical cavity lasers," *IEEE Photon. Technol. Lett.*, Vol. 9, pp. 11–13 (1997).
 - [41] N. M. Margalit, J. Piprek, S. Zhang, D. I. Babic, K. Streubel, R. P. Mirin, J. R. Wesselmann, J. E. Bowers, and E. L. Hu, "64°C continuous-wave operation of 5 μm vertical cavity laser," *IEEE J. Selected Topics in Quantum Electron.*, Vol. 3, pp. 359–365 (1997).
 - [42] K. D. Choquette, W. W. Chow, G. R. Hadley, H. Q. Hou, and K. M. Geib, "Scalability of small-aperture selectively oxidized vertical cavity lasers," *Appl. Phys. Lett.*, Vol. 70, pp. 823–825 (1997).
 - [43] A. E. Bond, P. D. Dapkus, and J. D. O'Brien, "Aperture placement effects in oxide-defined vertical cavity surface-emitting lasers," *IEEE Photon. Tech. Lett.*, Vol. 10, pp. 1362–1363 (1998).
 - [44] K. L. Lear, K. D. Choquette, R. P. Schneider, S. P. Kilcoyne, and K. M. Geib, "Selectively oxidised vertical cavity lasers with 50% power conversion efficiency," *Electron. Lett.*, Vol. 31, pp. 208–210 (1995).
 - [45] B. Weigl, M. Grabherr, C. Jung, R. Jager, G. Reiner, R. Michalzik, D. Sowada, and K. J. Ebeling, "High performance oxide-confined GaAs VCSELs," *IEEE J. Selected Topics in Quantum Electron.*, Vol. 3, pp. 409–415 (1997).
 - [46] E. R. Hegblom, N. M. Margalit, A. Fiore, and L. A. Coldren, "Small, efficient vertical-cavity lasers with tapered oxide apertures," *Electron. Lett.*, Vol. 34, pp. 895–897 (1998).

- [47] E. R. Hegblom, R. L. Naone, N. M. Margalit, and L. A. Coldren, "Comparison of tapered apertures in vertical cavity lasers," *Proc. LEOS '97*, Vol. 2, pp. 350–351 (1997).
- [48] G. R. Hadley, "Effective index model for vertical-cavity surface-emitting lasers," *Opt. Lett.*, Vol. 20, pp. 1483–1485 (1995).
- [49] P. D. Floyd, B. J. Thibeault, E. R. Hegblom, J. Ko, and L. A. Coldren, "Comparison of optical losses in dielectric apertured vertical-cavity lasers," *IEEE Photon. Technol. Lett.*, Vol. 8, pp. 590–592 (1996).
- [50] K. L. Lear, H. Q. Hou, J. J. Banas, B. E. Hammons, J. Furioli, and M. Osinski, "Vertical cavity lasers on p-doped substrates," *Electron. Lett.*, Vol. 33, pp. 783–784 (1997).
- [51] E. M. Strzelecka, G. B. Thompson, G. D. Robinson, M. G. Peters, B. J. Thibeault, M. Mondry, V. Jayaraman, F. H. Peters, and L. A. Coldren, "Monolithic integration of refractive lenses with vertical cavity lasers and detectors for optical interconnections," *Proc. SPIE Photonics West*, Vol. 2691, pp. 43–53 (1996).
- [52] E. M. Strzelecka, G. D. Robinson, M. G. Peters, F. H. Peters, and L. A. Coldren, "Monolithic integration of vertical-cavity laser diodes with refractive GaAs microlenses," *Electron. Lett.*, Vol. 31, pp. 724–725 (1995).
- [53] Z. D. Popovic, R. A. Sprague, and G. A. N. Connel, "Technique for monolithic fabrication of microlens arrays," *Appl. Opt.*, Vol. 27, pp. 1281–1284 (1988).
- [54] O. Blum, S. P. Kilcoyne, M. E. Warren, T. C. Du, K. L. Lear, R. P. Schneider, Jr., R. F. Carson, G. Robinson, and F. H. Peters, "Vertical-cavity surface-emitting lasers with integrated refractive microlenses," *Electron. Lett.*, Vol. 31, pp. 44–45 (1995).
- [55] H. Sankur, E. Motamedi, R. Hall, W. J. Gunning, and M. Khoshnevisan, "Fabrication of refractive microlens arrays," *Proc. SPIE Photonics West*, Vol. 2383, pp. 179–83 (1996).
- [56] L. P. Chen and K. Y. Lau, "Regime where zero-bias is the low-power solution for digitally modulated laser diodes," *IEEE Photon. Technol. Lett.*, Vol. 8, pp. 185–187 (1996).
- [57] D. L. Huffaker, L. A. Graham, H. Deng, and D. G. Deppe, "Sub-40 μ A continuous-wave lasing in an oxidized vertical-cavity surface-emitting laser with dielectric mirrors," *IEEE Photon. Technol. Lett.*, Vol. 8, pp. 974–976 (1996).
- [58] K. L. Lear, K. D. Choquette, R. P. Schneider, and S. P. Kilcoyne, "Modal analysis of a small surface emitting laser with a selectively oxidized waveguide," *Appl. Phys. Lett.*, Vol. 66, pp. 2616–2619 (1995).
- [59] K. D. Choquette, R. P. Schneider, K. L. Lear, and K. M. Geib, "Low threshold voltage vertical-cavity lasers fabricated by selective oxidation," *Electron.*

- Lett.*, Vol. 30, pp. 2043–2044 (1994).
- [60] K. L. Lear, A. Mar, K. D. Choquette, S. P. Kilcoyne, R. P. Schneider, and K. M. Geib, “High frequency modulation of oxide confined vertical cavity surface emitting lasers,” *Electron. Lett.*, Vol. 32, pp. 457–458 (1996).
 - [61] M. S. Ünlü and S. Strite, “Resonant cavity enhanced photonic devices,” *J. Appl. Phys.*, Vol. 78, pp. 607–639 (1995).
 - [62] K. Kishino, M. S. Ünlü, J.-I. Chyi, J. Reed, L. Arsenault, and H. Morkoç, “Resonant cavity-enhanced (RCE) photodetectors,” *IEEE J. Quantum Electron.*, Vol. 27, pp. 2025–2034 (1991).
 - [63] O. Sjölund, D. A. Louderback, E. R. Hegblom, J. Ko, and L. A. Coldren, “Novel integration technique of vertical cavity lasers and resonant photodetectors,” *Appl. Phys. Lett.*, Vol. 73, pp. 1–3 (1998).
 - [64] O. Sjölund, D. A. Louderback, E. R. Hegblom, J. Ko, and L. A. Coldren, “Monolithic integration of microlensed resonant photodetectors and vertical cavity lasers,” *Electron. Lett.*, Vol. 34, pp. 1742–1743 (1998).
 - [65] J. M. Dallesasse, N. Holonyak, A. R. Sugg, T. A. Richard, and N. Elzein, “Hydrolyzation oxidation of AlGaAs-AlAs-GaAs quantum well heterostructures and superlattices,” *Appl. Phys. Lett.*, Vol. 57, pp. 2844–2846 (1990).
 - [66] M. H. MacDougall, G. M. Yang, A. E. Bond, C.-K. Lin, D. Tishinin, and P. D. Dapkus, “Wide-bandwidth distributed Bragg reflectors using oxide/GaAs multilayers,” *Electron. Lett.*, Vol. 30, pp. 1147–1149 (1994).
 - [67] K. D. Choquette, K. L. Lear, R. P. Schneider, and K. M. Geib, “Cavity characteristics of selectively oxidized vertical-cavity lasers,” *Appl. Phys. Lett.*, Vol. 66, pp. 3413–3415 (1995).
 - [68] M. H. MacDougall, P. D. Dapkus, A. E. Bond, C.-K. Lin, and J. Geske, “Design and fabrication of VCSELs with Al_xO_y -GaAs DBRs,” *IEEE J. Sel. Top. Quantum Electron.*, Vol. 3, pp. 905–915 (1997).
 - [69] A. E. Bond, “Monolithically integrated surface and substrate emitting vertical cavity lasers for smart pixels,” *Appl. Phys. Lett.*, Vol. 73, pp. 19–21 (1998).
 - [70] K. D. Choquette, K. M. Geib, H. C. Chui, B. E. Hammons, H. Q. Hou, T. J. Drummond, and R. Hull, “Selective oxidation of buried AlGaAs versus AlAs layers,” *Appl. Phys. Lett.*, Vol. 69, pp. 1385–1387 (1996).
 - [71] M. H. MacDougall and P. D. Dapkus, “Wavelength shift of selectively oxidized Al_xO_y -AlGaAs-GaAs distributed Bragg reflectors,” *IEEE Photon. Technol. Lett.*, Vol. 9, pp. 884–886 (1997).
 - [72] J.-H. Kim, D. H. Lim, K. S. Kim, G. M. Yang, K. Y. Lim, and H. J. Lee, “Lateral wet oxidation of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ depending on its structure,” *Appl. Phys. Lett.*, Vol. 69, pp. 3357–3359 (1996).
 - [73] A. Fiore, Y. A. Akulova, J. Ko, E. R. Hegblom, and L. A. Coldren, “Low-threshold multiple-wavelength vertical-cavity laser arrays obtained by post-

- growth wet oxidation," *Electron. Lett.*, Vol. 34, pp. 1857–1858 (1998).
- [74] M. H. MacDougal, G. M. Yang, A. E. Bond, C.-K. Lin, D. Tishinin, and P. D. Dapkus, "Electrically-pumped vertical-cavity lasers with Al_xO_y -GaAs reflectors," *IEEE Photon. Technol. Lett.*, Vol. 8, pp. 310–312 (1996).
 - [75] M. Ghisoni, O. Sjölund, A. Larsson, S. M. Wang, J. Thordsson, T. G. Andersson, and L. Hart, "A comparative study of strain relaxation effects on the performance of InGaAs quantum well based heterojunction phototransistors," *IEEE J. Sel. Topics Quantum Electron.*, Vol. 3, pp. 768–779 (1997).
 - [76] R. L. Naone, E. R. Hegblom, B. J. Thibeault, and L. A. Coldren, "Oxidation of AlGaAs for tapered apertures in vertical cavity lasers," *Electron. Lett.*, Vol. 33, pp. 300–301 (1997).
 - [77] L. A. Coldren and S. W. Corzine, *Diode lasers and photonic integrated circuits*, Wiley, New York, pp. 201–204 (1995).
 - [78] J. E. Bowers, "High speed semiconductor laser design and performance," *Solid State Electron.*, Vol. 1, pp. 1–11 (1987).
 - [79] G. P. Agrawal, *Fiber-optic communication systems*, Wiley, New York, pp. 170–171 (1992).
 - [80] D. L. Huffaker and D. G. Deppe, "Multiwavelength, densely-packed 2×2 vertical-cavity surface-emitting laser array fabricated using selective oxidation," *IEEE Photon. Technol. Lett.*, Vol. 8, pp. 858–860 (1996).
 - [81] C. L. Chua, R. L. Thornton, D. W. Treat, and R. M. Donaldson, "Independently addressable VCSEL arrays on 3- μm pitch," *IEEE Photon. Technol. Lett.*, Vol. 10, pp. 917–920 (1998).
 - [82] M. W. Maeda, C. J. Chang-Hasnain, A. C. Von Lehmen, H. Izadpanah, C. Lin, M. Z. Iqbal, L. Florez, and J. Harbison, "Multigigabit/s operation of 16-wavelength vertical-cavity surface-emitting laser array," *IEEE Photon. Technol. Lett.*, Vol. 10, pp. 863–865 (1991).
 - [83] S. Y. Hu, J. Ko, O. Sjolund, and L. A. Coldren, "Optical cross talk in monolithically-integrated multiple-wavelength vertical-cavity laser arrays for multi-mode WDM local-area networks," *Electron. Lett.*, Vol. 34, pp. 676–677 (1998).
 - [84] G. Heise, "Cross talk investigation of laser-diode pairs," *IEEE Photon. Technol. Lett.*, Vol. 2, pp. 97–99 (1990).
 - [85] T. C. Banwell, A. C. Von Lehmen, and R. R. Cordell, "VCSE laser transmitters for parallel data links," *IEEE J. Quantum Electron.*, Vol. 29, pp. 635–644 (1993).

Large-Scale Heterogeneous Optoelectronic Integration

John A. Trezza, Jeffry Powell, Richard J. Williams, Ronald J. Olson,
Mani Sundaram, Keith Kang
Sanders, A Lockheed Martin Company

5.1 INTRODUCTION TO LARGE-SCALE HETEROGENEOUS INTEGRATION

Heterogeneous integration can refer to the combination of many types of devices. Silicon, GaAs, or SiGe electronics, micro-electro-mechanical devices, optics, optoelectronic devices, and packaging can be combined. This chapter will focus on the construction and use of devices incorporating very-large-scale integration with a particular emphasis on the integration of silicon electronics and optoelectronic devices.

Despite the availability today of leap-ahead integration capability, the fundamental physical limitations of very-large-scale heterogeneous integration have not been remotely approached. For example, given today's wafer uniformity and flatness, integration of pieces that are 3" can be easily accomplished. With inexpensive optical-lithography equipment, bump bonds of 2 microns can be easily produced. If one assumes 2-micron spacing around the bumps, and a 1-micron pixel spacing, this gives a 7-micron pixel pitch. In a 3" × 3" area, then, this allows almost 120 million pixels per hybridization step. If one allows more exotic lithographic techniques (such as x-ray or e-beam lithography) pixel pitches of 1 micron are possible. This would allow billions of devices to be flip-chip bonded in a single step. In addition, today's capabilities could allow such things as the flip-chip bonding of micromachined structures and advanced optical and electrical packaging to permit the use of such large numbers of devices. While such huge numbers of elements have never been attempted, it is clear that there remains much room for progress in the area of heterogeneous integration and that the technology is ripe to attempt such ultra-large-scale possibilities.

The extent of heterogeneous integration, while much less grandiose than the possibilities described above, is still quite impressive. Products incorporating very-large-scale heterogeneously-integrated optoelectronic devices are already a

reality. For example, heterogeneously integrated arrays with over 300,000 individually controllable optical elements with common device yields of over 99.99% are currently commercially available. In addition, wafer-at-a-time integration is available for high-volume production. If commercial demand required, arrays with $4,000 \times 4,000$ arrays could be made readily available this year. Compared with pin counts in traditional commercial applications (where 500 I/O seems enormous), technology available for commercialization is far in excess of what is commonly used in the commercial world.

The real value of heterogeneous integration, though, is not in being able to demonstrate high device densities or laboratory-level enhancements. Integrating large numbers of devices is useless unless the resulting modules have an application-specific use and are used in application-oriented systems. This chapter will outline some of the integration techniques being developed today. These include beam-lead integration, ball-grid-array development, and flip-chip bonding as well as some of the device-specific processing steps required. In addition, we will discuss resulting optoelectronic modules and performance in more mature system-level demonstrations, such as an infrared imaging (sensing), hyperspectral data processing, beam steering, and printing.

5.2 DEVICES AND DEVICE INTEGRATION TECHNIQUES

5.2.1 BACKGROUND

In the past, the technologies required to develop semiconductor devices and the means by which they were interconnected were treated mostly as separate entities. The design and development of a typical Si or GaAs semiconductor (e.g., an IC or discrete component) would rely on the omnipresent wire bonder to effect what will be defined as the first-level interconnection from the IC to the outside world (e.g., to a chip carrier). Generally, a good approximation of total cost could be assessed by simply adding the separate costs for device fabrication: the semiconductor die-to-lead-frame wire bonding and molding to create, for example, a dual in-line package (DIP), and finally the cost of the assembly of the printed wiring board (PWB). Because of the desire for higher I/O counts (driven by higher circuit density and increased functionality) and cost, adaptive and more efficient interconnection techniques were created. Some of these alternative first-level methods are beam lead interconnection, solder-bumps for flip-chip (FC) assembly, ball-grid arrays (BGA) for FC, and fine-pitch wire bonding. This section will review the progress on these technologies and their use in applications.

The term “first level” designates that the interconnection is made to the semiconductor die, whether for an optical, electrical, or optoelectronic device. In this context, a second-level interconnect can be envisioned that would attach to the

output of the first-level. An example may be several high-power ICs that have a priority on heat removal, or a modular device that integrates pretested subassemblies to a motherboard. Other second-level interconnection schemes may be required to match a part with coarse dimensions to one with finer dimensions, e.g., an ASIC to a PWB. Still other applications may focus only on the requirement to shrink the total area (volume) of the assembly. This second-level connection must also be compatible with the requirements imposed by mating materials with differing coefficients of thermal expansion (CTE) or those that require unique temperature control. The present discussion will be limited to a third-level interconnection, defined as the last connection required. An example of this includes the interconnection of the output of a daughter board to a PWB.

The decision of which combinations of first-, second-, or third-level packaging will be used is a multifaceted problem. For example, high material costs of the individual components that make up a composite assembly may be a deciding factor as to which functions are performed on which component. Such costs typically come into play when yield-limited availability of materials precludes the dedication of a material for an interconnect purpose. A specific case may include using all of the available area of an MCT (mercury-cadmium-telluride) substrate solely for detection of infrared radiation as opposed to the possibility of performing electronic interconnect functions on that material. Conversely, when using more widely available materials such as Si or GaAs, the designer may have more leeway to use the surface to handle some of the more basic functions such as interconnection. When material costs are low, more flexible packaging designs can be used. For example, the higher CTE of GaAs, compared to Si, may improve the matching to still higher CTE material. With the greater flexibility afforded by these more common materials, the complete elimination of the second-level interconnect may be realized. In this way the costs associated with the design, fabrication, and added assembly of the entire second-level may be avoided completely. While the trade-offs experienced in packaging are numerous, it should be noted that, in today's climate, cost is the overriding factor. The primary reason that today heterogeneous integration is attractive compared with conventional techniques, such as wirebonding, is that heterogeneous integration allows industry to achieve high I/O counts at a lower cost than conventional approaches. Should someone discover a way to reduce the cost of wirebonding by a hundredfold, the equation could change considerably.

5.2.2 TYPES OF INTERCONNECTS

5.2.2.1 *Silicon beam lead interconnect*

A particular type of first-level interconnect, the silicon beam lead interconnect (SBLI) has been used for a variety of applications and with a great deal of versatility. The basic process flow consists of first photolithographically defining the

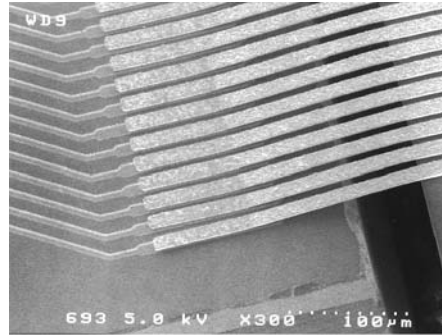


Figure 5.1. SEM photograph of SBLI used here to change elevation and heterogeneous materials.

desired beam pattern on the Si substrate. The beams are electroformed to a thickness of about 0.001 inch. Next the backside of the wafer is patterned with “windows” that are then etched through—cantilevering the portion of the beams that extend into this window area. The finished device must then be precisely aligned to the bond pads of the mating device or devices and then welded (by parallel-gap welding equipment). Any number of components may be interconnected by this method and can be bent into a variety of configurations. Thought of another way, they are rudimentary micromachined structures that allow high-density two- and three-dimensional tiling of electronic chips. Thousands of connections per chip are currently made this way; this allows systems such as imaging arrays or transceiver arrays to break the barrier of CMOS reticle sizes. As shown in Fig. 5.1, the cantilevered Au beams can be fabricated as fine-pitch interconnects (75- μm pitch shown) to interconnect dissimilar materials. Some of the uses for this technology are: to change the pitch of one component’s I/Os to match another, to bridge a gap between substrates, to provide stress relief, and to interconnect opposing geometrical planes.

The SBLI can replace conventional wire bonding for fine-pitch applications. This allows for tighter packing density of components. With a dielectric applied to a first-tier SBLI, a second tier can be inverted and placed on top of the first. This allows a stack assembled with even greater interconnect count per linear distance. This type of device can be fabricated to provide robust, over-the-edge interconnection so that devices can be formed around corners or even folded back completely to effect peripherally connected, stacked-chip configurations (see Fig. 5.2). This figure also shows how the SBLI can be used to change the orientation of a device, e.g., from circuit side up to circuit side down.

The costs of fabrication and assembly of this type of interconnect must of course be weighed against the costs of other choices that address similar intercon-

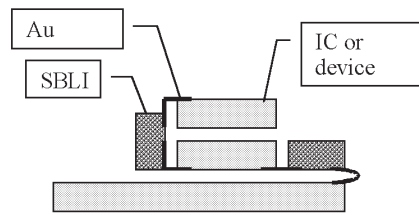


Figure 5.2. Illustration of SBLI devices used as over-the-edge interconnects to facilitate chip stacking.

nection challenges. If the total cost of all related steps is not competitive with competing technologies then there must be other compelling reasons for its inclusion.

5.2.2.2 *Ball-grid arrays*

There apparently exists a healthy level of debate pertaining to the naming of the different interconnection methods. It appears to some [1] that the ball-grid array (BGA) designation should refer to packages with ball pitches larger than 1.0 mm and that those with smaller pitches should be known as chip-scale packages (CSP).

To avoid confusion, this chapter will draw a distinction between the traditional BGA that refers to a surface-mount area-array package and a “bga.” The bga can be considered purely a type of interconnection as opposed to a package type, more of an array of (metal) balls distributed on a grid. It can apply to either fine- or coarse-pitch arrays as they might pertain to respectively, a thin-film device (metallization thickness less than 0.5 μm) or to a thick-film device. With this definition in mind, it is easy to conceive of a bga that simply interconnects dissimilar materials, e.g. an IC directly to a PWB, and not only the IC to a surface mount area array package. Typical thin-film parts may be ICs and optoelectronic devices, whereas the thick-film parts could be PWB, patterned ceramic, etc. The continued use of the traditional, peripherally located I/O pads for some applications may be unnecessarily prolonging the use of an older technology while adding extra cost. If one views the acceptance of BGA designs for flip-chip (FC) as a paradigm shift and takes advantage of all of the benefits, total costs will be reduced. Full acceptance would mean that, for example, I/O pad drivers from electronic chips could be spread throughout the 2D area of the chip rather than being located near the periphery of the chips. This is in contrast to near-periphery placement done in most current flip-chip-bonded commercial products such as the latest generation of microprocessors. Of course, full 2D contact placement demands that all of the engineering disciplines adapt to the new design rules in order maximize the benefits of FC technology. Figure 5.3 shows a coarse-pitch, linear bga and Fig. 5.4

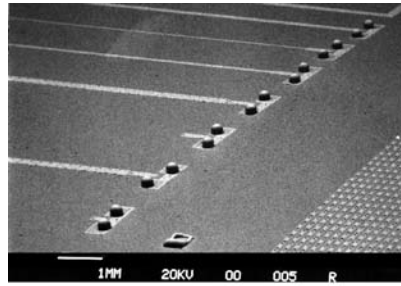


Figure 5.3. Coarse bumps (500- μ m pitch) for flip-chip on a diamond substrate.

shows a two-dimensional (2D), fine-pitch array. A defining characteristic of the bga is that a higher number of interconnections can be made compared to their wire-bonded counterparts. Still, in the coarse-bonding approach (typical of what is used in commercial electronics) wiring (on the ASIC) to the ball must be made. While this does consume space on the chip, typical precision and multilevel metallizations of modern semiconductor fabrication tools and processes can alleviate the impact on cost. On fine-pitch arrays, dispersion of I/O throughout the electronic chip is much more conveniently realized. The impact of I/O throughout the chip is typically the increased complexity of circuit layout and signal routing in high-speed areas of the chip. High-speed signals must route around the I/O areas and this can affect signal skew and decrease noise margins. In coarse-pitch arrays, the impact on layout is often too great. In fine-pitch arrays, however, the chip area needed for I/O in critical-path portions of the chip is reduced to the point where circuit layout is not significantly impacted.

Today, bga devices with 24- μ m pitches are routinely hybridized on 640×480 arrays. These devices integrate a GaAs, quantum-well infrared photodetector (QWIP) to a Si, readout integrated circuit (ROIC) to create a QWIP focal-plane array (FPA). Adding to the challenge of this composite assembly is the fact that these parts must be operated at cryogenic temperatures (< 100 K) while maintaining interconnection of the 307,200 bumps, packed into a size of roughly 12×15 mm. At this tight pitch, operability of $> 99.99\%$ has been achieved, which includes the operation of the QWIP device (each pixel), each readout cell of the ROIC, and the bump-to-bump integrity—for all of the 307,200 interconnections. The hybridization process for this array size has been qualified by successfully completing 5000 temperature cycles between room temperature and 77 K—with no loss in operability.

GaAs is the predominant material for device fabrication and therefore the main material requiring interconnection. The previously mentioned SBLI, bump deposition on Si ROIC (for flip-chip attach) and traditional wire bonding are standard methods used for this purpose. It is also conceivable that an SBLI device

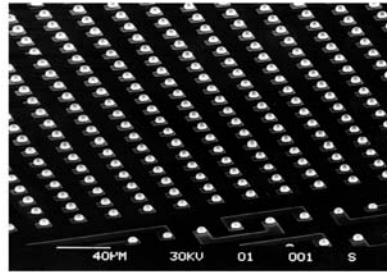


Figure 5.4. Fine-pitch bumps (25- μm pitch) for flip-chip interconnection of optoelectronic devices on a silicon substrate.

could be formed on another substrate such as GaAs, if required, for example, to better match the CTE. Likewise the bump deposition and flip-chip attach process can be performed to a variety of substrates.

Microvia and other multilayered boards are available to integrate different materials and pitches. With proper CTE properties, these boards can be used to integrate low CTE components, such as Si, to high CTE components, such as PWBs. A principal advantage of these boards is that they have higher density capability compared to their PWB counterparts which further enables their use for flip-chip attach.

5.2.3 FLIP-CHIP DETAILS

5.2.3.1 Background

Typical room-temperature hybridization, compression bump-bonding processes are actually effecting cold welding a subcategory of solid-phase welding. In this process, arrays of metal bumps on opposing substrates are aligned and then brought together with enough force to cause deformation. In order to achieve dependable cold welds [2], the metals must be deformed by 50–95%, depending on the particular metal. This method of interconnection is highly reliable provided that there is adequate bonding force available. This becomes more of a challenge for larger arrays (more bumps) and harder metals because the precision alignment of the parts must be maintained during the entire flip-chip sequence—under terrific loads. This becomes of concern, especially when bonding comparatively fragile substrates like MCT.

For each process or assembly step, the engineer must obviously consider the temperature limitations of the preceding operations' components. In general the higher-temperature operations need to be done first so that overheating of the subsequent, vulnerable assemblies is avoided. This can lead to challenges when integrating advanced devices. In order to increase the production worthiness of certain temperature-sensitive, advanced devices, alternatives to the safer, descending-tem-

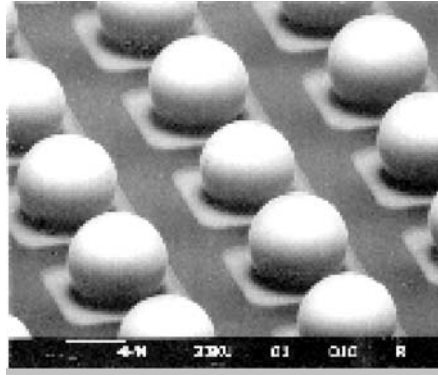


Figure 5.5. Close-up view of reflowed solder bumps.

perature profile must be invoked. If the final assembly includes a solder reflow of a PCB, the high temperatures of this process must be accounted for by some means. One method to maintain high connectivity with an acceptable bonding force is to include limited heating during the FC cycle. Even by adding heat that is below the melting temperature, the magnitude of the bonding force may be sufficiently reduced. This way the higher-melting-temperature metal can be used earlier in the process.

Before a proper bond can be made between some metals, e.g., Pb-based solders, the native oxide may require removal if consistent interconnect is to be achieved. Heating parts in the presence of oxygen can increase the thickness of the oxide. Leaving the oxide in place can degrade mechanical performance of the interconnection or create electrical instabilities if it is nonconductive. There are a number of methods available to negate the undesirable effects of these oxides, such as: removing them with fluxes or by bonding in a “oxygen reducing” environment, mechanically disrupting them or even storing the bumped die in an inert environment (avoiding the oxide growth in the first place).

5.2.3.2 Other techniques

For some high-speed devices, it may prove advantageous to reflow the bumps either before hybridization or after. By reflowing the bumps, the sharp edges are reduced or eliminated, which may lessen electromagnetic “radiating points,” and improve the noise figure of the circuit. This is especially important in many of the MMIC chips produced where speeds of up to 200 GHz are employed. Reflow can be accomplished prior to hybridization as shown in Fig. 5.5 or after if there are high temperature limits for a mating nonbumped component.

By placing the bumps on a smaller pad of “wetable metal” and reflowing, the bumps can be made to self-align to the pads. This wettable metal pad would have

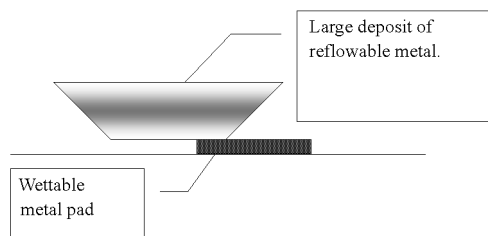


Figure 5.6. As-deposited bump before reflow.

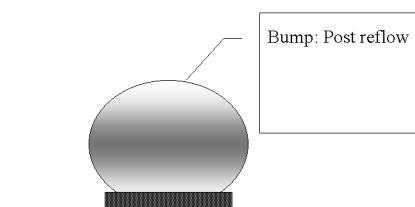


Figure 5.7. Bump after reflow.

a much higher affinity for the molten bump than the surrounding areas. An example of this might include a gold pad that is surrounded by silicon nitride. Applications of this technique might include a low-precision (low-cost tool) alignment of a metal bump pattern or bump height modification. This latter use may apply when the deposited metal's height is process limited; a thin metal consuming a large area can be reflowed to form a bump on a smaller area with increased height. Figure 5.6 shows a single, low, misaligned pad/bump before reflow and Fig. 5.7 the resulting changes to the height and placement of the bump.

This same mechanism can be performed over an entire FC-bonded BGA and has been used in various forms throughout the industry. When reflowed, the whole array can be shifted en masse to self-align two parts. Observing the progression of this technique in real time can be interesting as viewed by optical microscopy; when the melting temperature of the bumps is reached, the top chip “snaps” into alignment.

5.2.4 REACTIVE-ION ETCHING (RIE) AND MOLECULAR-BEAM EPITAXY (MBE)

For the typical case of the QWIP or spatial light modulator (SLM, detailed in a later section), after it has been hybridized, the substrate (the starting GaAs wafer) needs to be removed without harming the epitaxial layers. This process is referred to as backside thinning and allows the optimization of the device's elec-

trical, optical, and mechanical properties. Backside thinning also enables “staged” or repeated integration of other devices on a single ASIC and is done to create devices for such things as hyperspectral sensors or transceiver arrays. For example, in transceiver arrays one device type is hybridized to an ASIC (driver) and thinned so that a second device type can be hybridized to the same driver. In this way the two GaAs devices are interdigitated on one driver in a two-stage hybridization. Using this same approach, the heterogeneous integration of several components is quite reproducible.

Though the backside-thinning process is similar to the RIE processes used to define features (such as electrical isolation of the individual pixels) on the front side of the wafer, there are notable differences. In order to create the fine front side features the etch needs to be highly anisotropic. The pixel-to-pixel isolation etch must be capable of etching trenches that are, as an example, 5 μm deep and 0.1 μm to 1 μm wide. In addition, these features must be highly uniform across the entire wafer. Conversely, the backside process must be very isotropic with high etch rates and low damage.

This ability to thin QWIP, modulator, VCSEL, p-i-n detector, and similar devices virtually assures the success of integrating parts with increased active areas. With QWIP devices, the thin, low-mass detector material tends to conform to the changing surface of the ROIC as the temperature varies. On arrays of two-terminal devices that can be fabricated without a continuous contact layer, there is virtually no limit to the size of the array. Low-cost processing of heterogeneous devices has been realized through advances in dry-etch techniques coupled with judicious use of MBE grown layers. This RIE-MBE combination yields a highly selective process with exceptional latitude and high device performance. These attributes define a successful, production-worthy process.

Specific electro-optical improvements brought about by the backside-thinning operation are better photoresponse and reduced pixel-to-pixel cross talk. Using the example of QWIP technology, because the actual device layers are so thin (typically less than 5 μm thick) the thermal mismatch between the focal-plane array (FPA—a visible or infrared imaging sensor array) and ROIC is dramatically reduced. As the RIE process allows the stop-etch to occur with exacting precision, vertical integration of subsequent devices is facilitated. Again using the example of hyperspectral sensors or transceiver arrays, the first device needs to be made thin enough that the second device has room to be placed in close proximity or interstitially in arrays. At the same time, the first device’s layers must be left thick enough so as not to damage the active device layers. The second hybridization is mating to a pseudoplanar surface.

Backside thinning aids in the manufacturing of higher-power chips, higher-frequency devices and high-efficiency optoelectrical devices [3]. In the case of monolithic microwave ICs, backside-substrate removal allows efficient heat

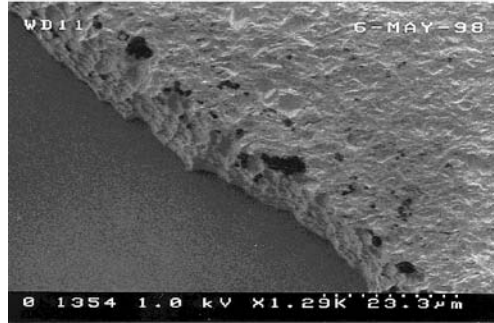


Figure 5.8. A typical post-RIE surface on the left side, with the rougher, pre-etch, lapped surface on the right. The lapped side was preserved by masking during the etch process.

removal from the device and reduces transmission-line impedance—both of which improve stability and electrical performance.

The conventional method, now superseded, of backside-substrate removal is chemical-mechanical polishing [4, 5, 6]. This method suffers from poor control of the substrate removal which can lead to erosion and damage (lapping-induced stress) of sensitive epilayers (such as GaAs/AlGaAs quantum wells) grown on the substrate. Chemical-mechanical polishing requires the presence of *thick* (~1000 nm) AlGaAs etch-stop layers between active device layers and substrate. Such thick epilayers require longer MBE growth times, which adds cost, and also results in proportionately higher growth-defect densities. These growth defects are responsible for reducing pixel operability in QWIP FPAs.

A dry-etch process has been developed to completely remove the GaAs substrate, revealing the device-processed epilayers [6]. The process is uniform, reproducible, and preserves the properties and yield of the processed devices. Infrared

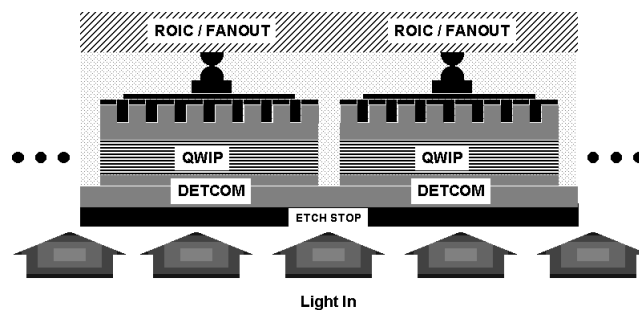


Figure 5.9. Cross section of QWIP FPA bump-bonded to a ROIC/Fanout. The photosensitive QWIP layers are sandwiched between doped contact layers. A metallized grating etched into the top contact couples light incident from the bottom. The AlGaAs etch-stop layer is shown at the bottom; it is removed prior to packaging.

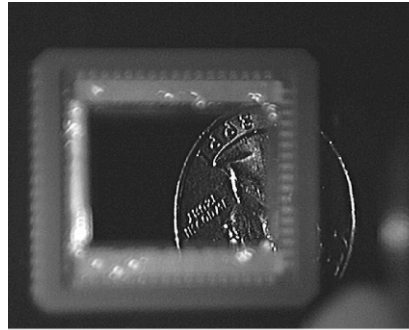


Figure 5.10(a). Digital image of the reflection of a quarter from a thinned FPA.

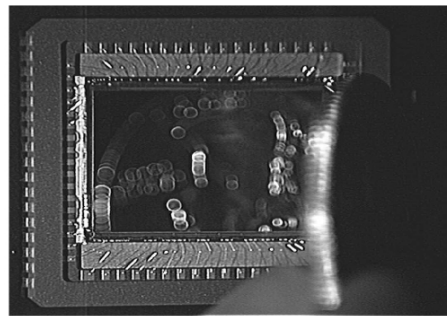


Figure 5.10(b). Digital image of the complete integrated package.

images produced by FPAs thinned using this alternate backside process, directly and vividly demonstrate its high quality. The SEM photograph (Fig. 5.8) shows the roughness of the lapped, pre-RIE GaAs substrate and the smoothness of the epilayer as revealed by the RIE process.

A cross section of a completed QWIP FPA (fabrication details are presented in a later section) is illustrated in Fig. 5.9. The Si ROIC (top) is shown interconnected by solder bumps to the QWIP. Figure 5.10(a) demonstrates the specularity of the etched backside surface of a packaged QWIP FPA by imaging the reflection of a quarter (coin). Figure 5.10(b) shows the complete integrated package more clearly with the wirebonds interconnecting the ROIC to the leadless chip carrier (LCC). The technique's success in producing a smooth and uniform surface (following substrate removal) over a large chip area (typically $16 \text{ mm} \times 12 \text{ mm}$) is directly evident in the high quality of the infrared imaging the FPAs demonstrate (Fig. 5.11). A typical noise-equivalent temperature difference (NETD) [16] of approximately 20 mK is obtained with these FPAs; the NETD includes both tem-



Figure 5.11. An image from a 640×480 8.6-mm (long-wave IR) QWIP FPA thinned using the technique presented in this chapter. Target temperature differences as small as 20 mK can be resolved.

poral and spatial noise and is strongly affected by the quality of the backside thinning and its uniformity. Imaging with QWIP FPA represents the most pragmatic proof of the success of the thinning technique. Figure 5.11 shows one such image produced by this process. These routine results attest to the technique's quality, uniformity, and reproducibility.

5.2.5 FOCAL-PLANE ARRAY (FPA)-SPECIFIC PROCESSING

FPA fabrication is carried out in a wafer-level process using stepper lithography and reactive-ion etching (RIE) for all etching steps. The one-color FPA is processed as follows: grating metal is first defined and a grating pattern etched into the top layer (the grating is necessary to diffract the normally incident light and cause it to be absorbed by intersubband transitions in the valence band of the quantum wells). Formation of the grating is followed by etching of the pixel mesas using ohmic metal as a mask. (This metal serves as both pixel electrical contact as well as grating reflector). Ohmic metal is then evaporated to contact the bottom contact layer at the FPA boundaries. All ohmic contacts are annealed. An array of solder bumps (on top of each pixel and at the FPA boundaries) is then evaporated.

The wafer is diced to separate the FPAs; individual dies are selected and hybridized (i.e., In bump-bonded) to matching Si ROIC dies. Each hybrid is thinned to have the entire GaAs substrate removed through RIE (the etch-stop layer aids in complete removal of any residual substrate at this step). Substrate removal mitigates the problems of thermal mismatch between the GaAs-based detector and the Si ROIC, improves optical coupling (and therefore quantum efficiency), and cuts pixel-to-pixel optical cross talk. The result of all the processing is a $\sim 5\text{-}\mu\text{m}$ -thick epitaxial

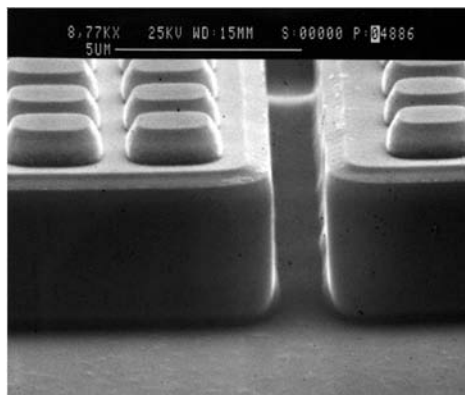


Figure 5.12. Cross-section of one-color QWIP FPA pixel. Note narrow trench, vertical walls, and high fill-factor.

membrane glued to the Si ROIC. The hybrid is then packaged and wire-bonded in a 68-pin LCC (with a calibrated temperature sensor) for imaging. Illumination is through the bottom ohmic contact or “Detcom” layer (so called because it is the electrical contact layer common to all detector pixels). One die from each wafer is selected and hybridized to a fan out for direct electrical and optical measurements. The 320×256 FPA has $28 \times 28\text{-}\mu\text{m}$ pixels on a $30 \times 30\text{-}\mu\text{m}$ pitch (i.e. fill factor $\sim 87\%$); there are 30 detector arrays on a three-inch wafer. The 640×480 FPA has $22 \times 22\text{-}\mu\text{m}$ pixels on a $24 \times 24\text{-}\mu\text{m}$ pitch (i.e. fill factor $\sim 84\%$); there are 12 detector arrays on a three-inch wafer. The control and precision of the fabrication process are readily seen in the image of a typical pixel shown in Fig. 5.12.

The two-color FPA process (the function of the two-color detector is described in Section 5.3) is similar to the one-color process. However, there are extra steps required to create and contact via holes to the middle and bottom ohmic contact layers in each pixel and to run metal lines (on an insulating layer to prevent shorting) from the two vias to the pixel top. The trenches between the pixels are etched deep enough (to the top etch-stop layer) to achieve complete electrical isolation between pixels. The bottom etch stop aids in substrate removal from the backside. The $1\text{-}\mu\text{m}$ -thick GaAs layer between the two etch stops serves to physically hold the array together (in addition to the In bumps and the epoxy); this layer is thin enough that complete optical isolation between the pixels is secured. The 256×256 FPA has a $40 \times 40\text{-}\mu\text{m}$ pitch; each pixel has two $9 \times 9\text{-}\mu\text{m}$ via cutouts on a $38 \times 38\text{-}\mu\text{m}$ frame, resulting in $\sim 80\%$ fill factor for the top (“red”—longer wavelength) QWIP and $\sim 85\%$ fill factor for the bottom (“blue”—shorter wavelength) QWIP. There are 16 detector arrays on a three-inch wafer. Figure 5.13 shows top views of a two-color pixel before and after In bump deposition: the narrow trench-

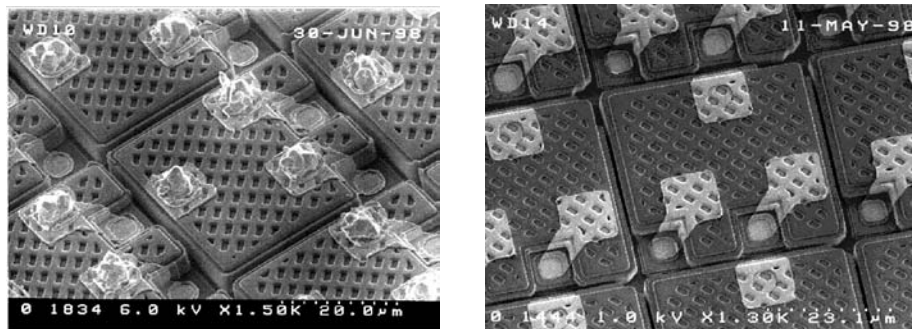


Figure 5.13. Top view of two-color QWIP FPA pixel showing metal lines running from mid and bottom vias to pixel top (left), and after solder bump deposition (right).

es (between pixels), small vias, and vertical sidewalls testify to the advanced degree of process control.

The grating to couple light to the two spectral bands in the two-color FPA is selected after a careful optimization of all grating parameters (pitch, duty cycle, polarity, orientation, depth, and choice of metal) in control experiments on corresponding one-color arrays. The optimum choice turns out to be one where the blue QWIP absorbs both light polarizations whereas the red QWIP absorbs light of only one polarization. It is possible to make the FPA polarization insensitive in both bands at the expense of slightly less quantum efficiency in the blue band.

5.2.6 SPATIAL-LIGHT-MODULATOR-SPECIFIC PROCESSING

As their name implies, spatial light modulators (SLMs) are devices developed to produce spatially varying optical patterns in a controllable way. Such devices are particularly useful in optical signal processing—correlation, triple product processing, Fourier filtering—as well as for beamsteering, data encoding, and wavefront correction. Spatial light modulators have been created in other technologies, such as liquid crystal. These have been hampered by slow switching time. For example, ferroelectric liquid crystal pixels have a maximum operating rate of about 4 KHz. Individual GaAs pixels can be operated at tens of GHz, seven orders of magnitude faster!

One of the keys to GaAs-based modulators' usefulness is that individual pixels can be tailored to be amplitude, phase, or direction modulators, detectors, or light-emitting diodes. Arrays that are close-packed, but act in detection mode are referred to as “smart detectors.” When arrays are more sparsely populated and contain modulators, emitters, and detectors, they are referred to as “smart pixel arrays.” In each case, the optical chip is connected at the pixel level to an electronic driver chip consisting of a custom, “mixed-signal” silicon ASIC.

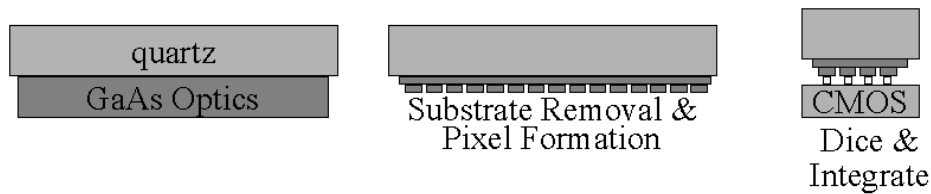


Figure 5.14. Summary of SLM-type process flow.

The process flow for spatial light modulators is somewhat different from that of the focal-plane-array process. Unlike the focal-plane-array process in which the GaAs wafer is processed, then hybridized and then thinned, in the spatial-light-modulator process, the wafer is first attached to a quartz carrier, then the substrate of the GaAs wafer is removed; the devices' optical cavity can then be tuned, and the arrays pixelated. Subsequently, the wafer is attached to CMOS chips and diced (the dicing and hybridization steps can occur in either order). Unlike the FPA process, in which the top of the epitaxial, as-grown wafer is closest to the CMOS chip in the final product, in the SLM process, the top of the as-grown epitaxial material is farthest from the CMOS in the final product. The major reason this is preferred is because it allows one to more easily tune the optical cavity of the modulator structures after substrate removal but before integration with CMOS. The FPA process would require CMOS integration prior to substrate removal and device tuning. The added processing steps required after CMOS integration—such as tuning of optical cavities—adds risk of damaging CMOS driver chips, which could negatively impact cost of manufacturing. This is less of an issue in structures where optical-cavity tuning is not required after growth.

5.2.7 TRANSCEIVER-SPECIFIC PROCESSING

Unlike focal-plane arrays or spatial-light-modulator arrays in which one type of device is used (detectors in the case of focal planes and modulators in the case of spatial light modulators), transceivers use two types of devices—transmitting devices and detecting devices. Two basic types of transceiver-array modules have been created in the past: modulator-based and emitter-based.

Modulator-based transceivers are still the most mature transceiver technology developed. In addition, they are the cheapest to produce, the highest yield, and the most tolerant to environmental conditions. Often, amplitude modulators are employed in such modulator-based transceiver links. An advantage of amplitude modulators—those types of modulator devices which switch between being highly reflective or highly absorptive—is that they also act as detectors when in their absorptive phase. In such a transceiver array, while two types of *logical* devices exist in the array (transmitter and detector), only one type of *physical* device exists

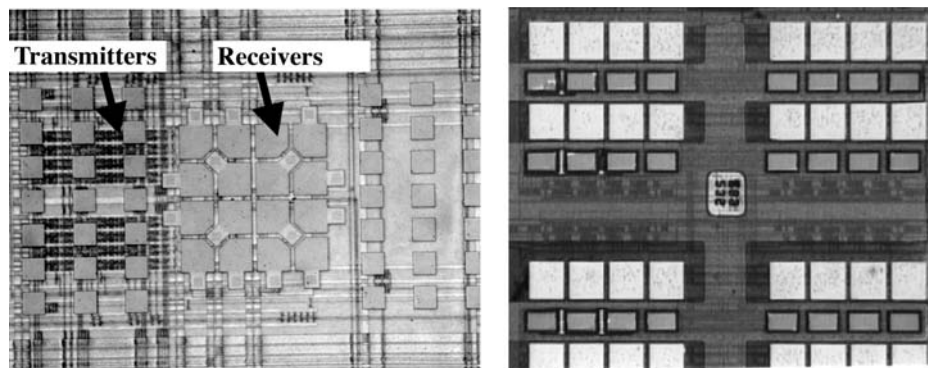


Figure 5.15. Close-up of sections of a fully CMOS-integrated transceiver array (left) created with one type of device (amplitude modulators) and (right) with two different types of devices (emitters and detectors) for transmitter and receiver. Modulator/detector-based arrays up to 256×256 , LED/detector-based arrays up to 256×256 and VCSEL/detector arrays up to 32×32 have been made.

(the amplitude modulator). The determination of whether individual devices act as modulators or detectors at any given moment is determined by the electronic driver circuit. Fabrication of modulator-based transceiver devices then follows the same type of integration processes that either focal planes or spatial light modulators follow—because only one type of physical device is present.

The second type of transceiver-based device contains two types of physical devices—one optimized as an emitter and one optimized as a detector. Emitters are usually comprised of either light-emitting diodes, resonant-cavity light-emitting diodes, or vertical-cavity surface-emitting lasers. Detectors are usually either p-i-n- or MSM-based. Regardless of the specific technologies used, the integration process is essentially the same. One of two approaches can be used. The first approach is based upon the focal-plane-array development. In this approach, the first of the two device types is hybridized to the electronic chip, then the substrate of the optoelectronic device is removed, and the second device type is hybridized in the intervening spaces. The second approach follows the spatial-light-modulator type of development. In this approach, the two types of optical devices are attached to a transparent carrier substrate—such as quartz, SiC, or sapphire. After this integration on the transparent carrier, the transceiver arrays are attached to the electronic driver chip in a single step. Both steps have been employed. Wafer-at-a-time processing is used in both techniques.

More complex arrangements with multiple types of devices on the same substrate (e.g., a single array with emitters, visible [red, green, and blue], near-, mid-, and long-wave IR, detectors, and modulators) can be fabricated by taking the transceiver-based approaches to their logical conclusion.

5.3 APPLICATIONS

5.3.1 INTRODUCTION

The ability to heterogeneously integrate optical and electronic devices together by itself is of no real value. The value is realized when this manufacturing capability allows dramatic increases in the performance of electronic systems or permits the creation of whole new applications that would be impossible to realize without heterogeneous integration. Figure 5.16 depicts some of the application space that has been investigated as a result of heterogeneous integration capabilities. Section 5.3 will discuss in detail some of these applications and the results of devices created using the types of techniques outlined in Section 5.2.

5.3.2 FOCAL-PLANE ARRAYS

5.3.2.1 Introduction

Long-wavelength infrared (LWIR) imaging (particularly in the 8–12- μm atmospheric window) is useful for security and surveillance, navigation and targeting, industrial process control, medical imaging, the astronomy of cold objects, the study of atmospheric chemistry (e.g., monitoring and measuring the compositions and levels of pollutants and gases in the air), etc. Quantum-well infrared photodetectors (QWIPs) have become the leading candidates for sensitive and cost-effective LWIR focal-plane arrays (FPAs). The benefits of QWIPs include:

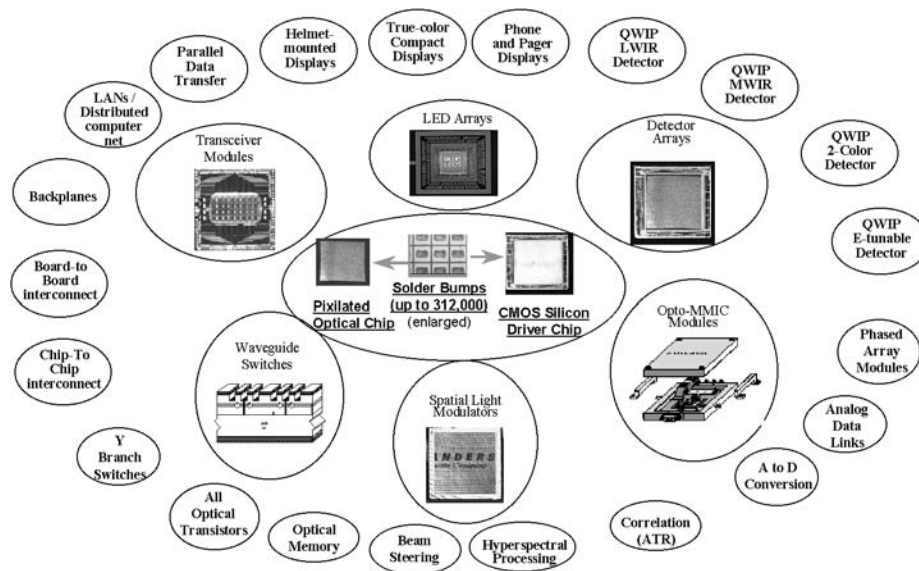


Figure 5.16. Application space.

InGaAs/GaAs/AlGaAs growth and processing technologies whose maturity and uniformity can be exploited for large-format 2D (640×480) array fabrication, tailorable spectral peak and width, laser and radiation hardness, negligible $1/f$ noise, and extension to spatially- and temporally-multicolor imaging. One-color LW QWIP FPAs in formats up to 640×480 and with excellent uniformity, operability, and highly sensitive imaging have been made by several groups [7, 8]. The technology has advanced enough that the FPAs can be consistently and reproducibly made, a requirement for commercialization. Performance summaries are presented of one-color 320×256 and 640×480 LW FPAs, devices now commercially produced (for insertion into infrared cameras such as the Inframetrics Milcam) and that represent the state of the art of one-color QWIP FPA technology in the market today.

Simultaneous and spatially-registered imaging in multiple infrared spectral bands is expected to provide or enhance clutter rejection, camouflage detection, false-alarm reduction, decoy discrimination, countermeasures resistance, target-ID enhancement, and absolute gray-body target-temperature measurement. Current multisensor systems require cumbersome optics precisely aligned to split and direct the incoming infrared light to two or more one-color staring FPAs, an approach that aggravates the cost, weight, performance, and reliability risk of the systems. A single staring FPA that provides simultaneous and spatially-registered imagery in multiple infrared bands is an attractive alternative that promises to mitigate such problems. No special optics are needed. Thanks to advances in material and fabrication technologies such FPAs can be made today.

Certain properties unique to the QWIP make it a natural choice for multicolor FPAs. Its sharp spectral response ($\Delta\lambda/\lambda \sim 10\%$) and tailorable peak ($\lambda_{\text{Peak}} \sim 3\text{--}18 \mu\text{m}$) mean that several detectors can be grown stacked on top of each other without the need for filters to prevent spectral cross talk. Also, one or more of the detectors can be an interband p-i-n diode for near infrared detection.

A two-color LW/LW QWIP FPA with pixel-registered and simultaneous imagery at $8.6 \mu\text{m}$ and $11.2 \mu\text{m}$ was recently reported [8]. This architecture has since been applied to MW/LW and MW/MW dualband combinations.

5.3.2.2 QWIP FPA architecture

Architectures of a one-color and two-color FPA are shown in Fig. 5.16. The one-color FPA has one solder bump/pixel which is bonded to the source of an input-switching FET in the corresponding unit cell of a matching Si readout-integrated circuit (ROIC) array. A layer (called Detcom in Fig. 5.17, top) common to all the pixels serves as the other contact to the FPA and is connected to the ROIC via bumps at the array boundaries. The ROIC applies a bias across the FPA by separately adjusting Detcom and VG (the voltage applied to the gate of the input FET); the photocurrent from each pixel is fed into an integrating capacitor C_{int} in

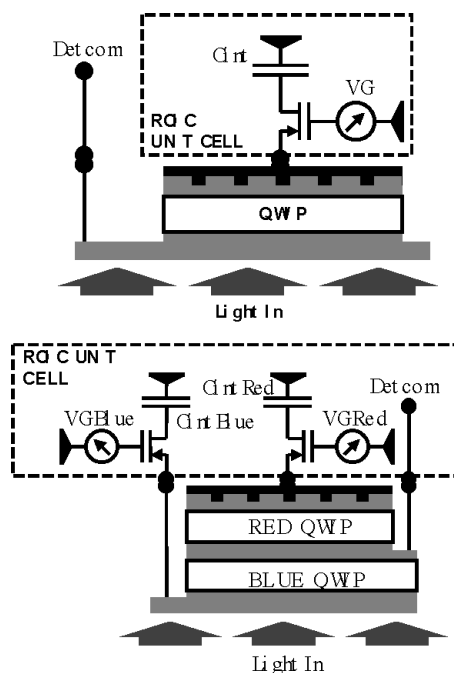


Figure 5.17. One-color (top) and two-color (bottom) FPA architectures showing one QWIP pixel In-bump-bonded to a ROIC unit cell.

the corresponding ROIC unit cell; this integrated charge is sampled and read out to generate the image at a frame rate of 100 Hz.

The two-color FPA architecture (Fig. 5.17 bottom) uses three solder bumps/pixel to put two vertically stacked QWIPs (grown in series) in parallel electrically. Detcom, VGred, VGblue are adjusted to separately bias the “red” (longer wavelength) and “blue” (shorter wavelength) QWIPs; the red and blue photocurrents are fed to two separate integrating capacitors in each unit cell which are then read out to generate the two simultaneous images. The vertical stacking of the detectors (together with the small contact geometries) ensures exact spatial registration of the two images. Such an architecture is readily extended to larger-format arrays and to at least four simultaneous colors (the number of colors being limited by the maximum number of bumps and contact vias that can reside on each pixel). Note that this architecture calls for complete electrical isolation between pixels: the red photocurrent flows from the top contact and the blue photocurrent from the bottom contact. Also note that while in the one-color architecture Detcom is common to all pixels, in the two-color architecture Detcom (the middle contact layer) is common only to the two colors in each pixel.

The two-color detector material features a red QWIP stack on top of a blue QWIP stack with three ohmic contact layers (the first below the blue stack, the sec-

ond between the blue and red stacks, and the third above the red stack) grown by molecular beam epitaxy (MBE). Each stack is either a LW QWIP realized with quantum wells of GaAs sandwiched between AlGaAs barriers, or a MW QWIP realized with deeper wells of strained InGaAs sandwiched between AlGaAs barriers; the contact layers consist of doped GaAs. Two etch-stop layers sandwiching a 1- μm layer of undoped GaAs are grown between the substrate and the device layers (their purpose is explained in the process description below). QWIP designs similar to the ones used here have been extensively reviewed in the literature [7]. Following growth, each wafer is x-rayed and the (004) x-ray rocking curve analyzed (in conjunction with the MBE growth parameters) to precisely yield the thickness and composition of every layer in the device structure to ensure compliance with design tolerances. The one-color detector material is similarly grown and characterized, the only difference being that there is a single QWIP stack with top and bottom ohmic-contact layers, and a single etch stop layer between the substrate and device layers.

5.3.2.3 One-color LW QWIP FPA performance

The performance of the commercial 320×256 one-color LW QWIP FPA is graphed in Figs. 18 and 19. At a typical operating bias of -3 V (the voltage applied to the top/grating contact with respect to the bottom/Detcom contact) the spectrum peaks at $8.6\text{ }\mu\text{m}$ (Fig. 5.17) with a 500 K black-body responsivity of 30 mA/W , a peak responsivity of 325 mA/W (Fig. 5.18), a quantum efficiency of 15%, and a photoconductive gain of 0.3. A peak detectivity D^* of 1.5×10^{10} Jones is

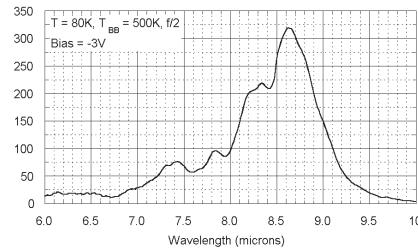


Figure 5.18. Spectral response of 320×256 LW QWIP FPA.

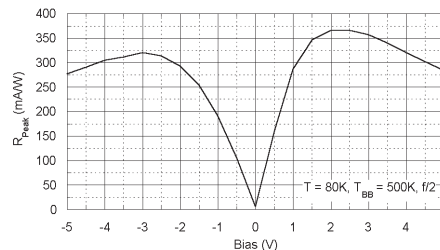


Figure 5.19. Bias dependence of the peak responsivity.

achieved at 80 K, limited by the shot noise of the dark current at 80 K. Cooling freezes out the dark current (dominated by thermionic emission at $T > 40$ K) exponentially, with a corresponding increase in D^* . In fact, the FPAs are usually operated at slightly colder temperatures, between 65 K and 77 K. These results are typically and consistently obtained and represent the best performance today with GaAs/AlGaAs quantum wells and a simple 2D periodic-grating coupler.

In the imaging arrays, a pixel operability (defined as every pixel that responds from 50% to 150% of the median FPA response) of $> 99.5\%$ is routinely obtained. Quite a number of the FPAs have operability in excess of 99.99%. The uncorrected response and dark current are extremely uniform with $\sigma/\text{mean} < 2\%$, testifying to the high uniformity in material growth and processing. In imaging radiometric measurements, the FPAs show a noise-equivalent temperature difference (NETD) < 20 mK when operated at 65 K with bias between -1 V and -3 V and integration time (T_{int}) between 4 ms and 10 ms in a system with $f/2$ optics, 295 K background, and 100 Hz frame rate. This low NETD value is proof of the sensitivity and uniformity of the FPA. An image from one such commercial FPA operating at a temperature of 78 K in a pour-fill LN2 dewar is shown in Fig. 5.20. Note the high image quality even at this relatively high temperature of operation.

The 640×480 FPA is identical in performance to the 320×256 FPA, with the added scene resolution afforded by its smaller pixel pitch and larger number of pixels.

5.3.2.4 Two-color QWIP FPAs

Simultaneous imagery from three separate two-color FPAs—LW/LW ($11.2 \mu\text{m}/8.6 \mu\text{m}$), LW/MW ($8.5 \mu\text{m}/5.1 \mu\text{m}$), and MW/MW ($4.7 \mu\text{m}/4.0 \mu\text{m}$)—are shown in Fig. 5.21. Note differences in each image pair due to target spectral differences.

The spectral response of the MW/LW array is graphed in Fig. 5.22. It shows detection in two distinct and separate bands peaked at $5.1 \mu\text{m}$ and $8.5 \mu\text{m}$.

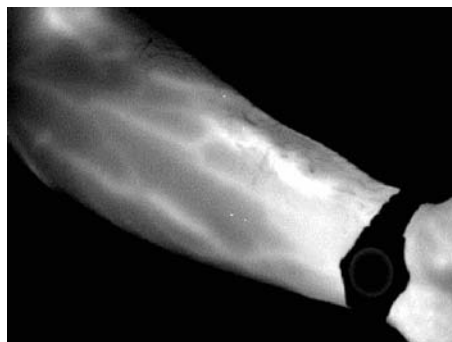


Figure 5.20. 8.6-mm image from 320×256 FPA with NETD < 20 mK. The FPA is at ~ 78 K in a pour-fill LN2 dewar.

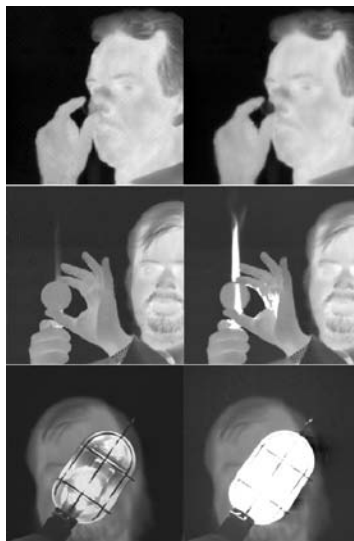


Figure 5.21. Simultaneous images from LW/LW (top), LW/MW (middle), and MW/MW (bottom) two-color QWIP FPAs.

Spectral cross talk is minimal. The fringes in the red band result from Fabry Perot resonances of the diffracted light waves in the pixel cavity due to the partial (to $\sim 25\ \mu\text{m}$) thinning of the array in this case. Complete thinning removes these fringes.

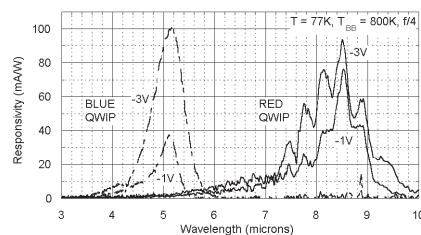


Figure 5.22. Spectral response of MW/LW QWIP FPA.

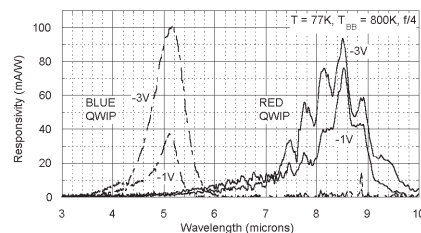


Figure 5.23. Bias dependence of peak responsivities.

The 800 K black-body responsivity RBB approaches ~ 10 mA/W for the blue QWIP (the highest ever measured for a MW QWIP FPA) and ~ 14 mA/W for the red QWIP. A typical operating bias of -2 V secures $R_{\text{Peak}} \sim 80$ mA/W and 110 mA/W for the blue and red bands, respectively (Fig. 5.22). The similar response in the two bands allows for balanced operation with the same bias across the two QWIPs, an attractive feature from the point of view of ROIC operation.

The best two-color FPAs in each of the three formats feature $> 99\%$ pixel operability, NETD better than 35 mK ($f/3$ optics, 295 K background, and 60–100 Hz frame rate), and excellent uniformity ($\sigma/\text{mean} < 4\%$) in both colors. While falling short of the $> 99.5\%$ operability and $\text{NETD} < 20$ mK routinely obtained in one-color LW QWIP FPAs, these numbers are excellent for a first attempt.

The two-color FPA is typically operated with a bias between -1 V and -2 V and tint ~ 4 –10 ms (the two charge wells in the ROIC are sized differently to allow the same tint for both colors for truly simultaneous integration). The LW/LW FPA runs at 40 K, the MW/LW (mid-wave/long-wave) FPA at 70 K, and the MW/MW FPA at 100 K, the operating temperature in each case being limited by the corresponding Red (longer-wavelength) QWIP.

Having demonstrated the feasibility of the two-color FPA, current efforts are aimed at improving the process to increase operability and uniformity, optimizing material and optical coupling designs to boost response and operating temperature, and making medium-format (640×512) two-color arrays. It is believed that substantial progress can be achieved in all these areas.

5.3.2.5 Tunable QWIP FPAs

An infrared focal plane array with electronically tunable spectral response offers a powerful way of performing both imaging as well as target spectroscopy—in other words, to collect the classic $x\lambda$ image data cube. Traditional systems use a broadband-detector array with a dispersive element (grating/prism) at the entrance slit to disperse the spectral bands along, say, the x direction. A 2D array secures $y\lambda$ data this way; the motion of the plane/satellite is then used to collect *spatial* data in the x direction; such a system is called push-broom. A 1D array requires, *in addition*, a y scanner; such a system is called whisker-broom. The image cube can alternately be collected with an electronically tunable filter (acousto-optic/liquid-crystal) placed in front of a 2D array. While it does away with moving parts, such filters are hard to make, inefficient, and are tunable only from NIR-MW (not in the LW window, where a number of interesting gases have strong spectral signatures). In a QWIP, the detection spectrum itself can be electrically tuned, thereby integrating filter and detector functions. Furthermore, *each* channel (pixel-registered and simultaneously-integrated) of a multichannel array can be made tunable in the MW or the LW. One of the channels can be a p-i-n-diode array for fixed NIR imaging. These inherent properties, together with the

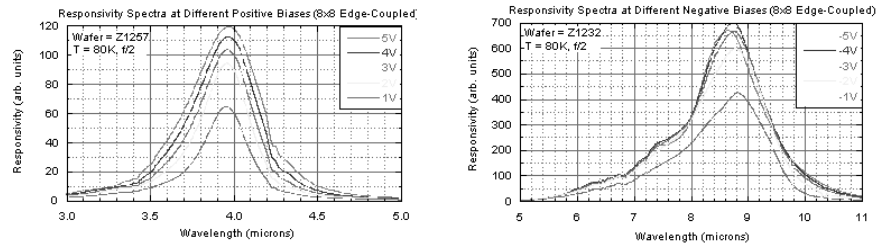


Figure 5.24. *Symmetric* quantum wells do *not* change spectral peak and shape with voltage, both in MW (left) and in LW (right) QWIPs.

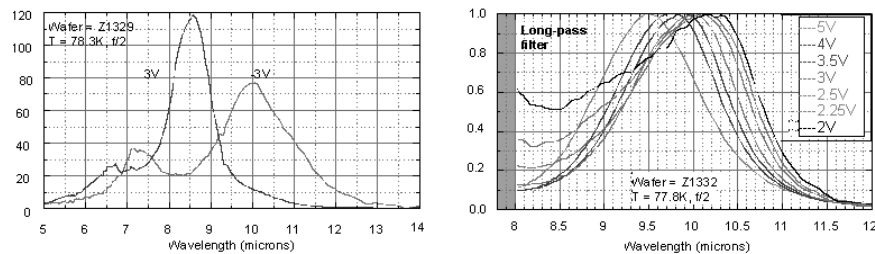


Figure 5.25. Responsivity (a.u.) vs. wavelength showing spectral tunability in *asymmetric* quantum wells: coupled wells (left), stepped well (right).

maturity of GaAs growth and processing technologies, promise an “ λ -imaging-system-on-a-chip” that is powerful, flexible, integrated, reliable, and has no moving parts.

The *quantum-confined Stark effect* in *asymmetric quantum wells* can be used to secure spectral tunability of QWIP response with applied voltage. Normally *symmetric* wells are used, which has the effect of pinning the response to a sharp (FWHM $\sim 0.4\text{--}1\text{ }\mu\text{m}$) band (Fig. 5.24) that does not change spectral peak and shape with voltage (the amplitude of the response does change due to changes in the p.c. gain and the photoelectron escape probability from the well). In *asymmetric* wells, on the other hand, the Stark effect moves the excited state in a controlled way (linearly with voltage, for instance, in a stepped well), that translates to a corresponding movement of the spectral peak. Asymmetric wells are easily constructed as two or more coupled wells of different widths and/or stepped (L-shaped) wells. Preliminary test devices show the promise of both these types of asymmetric wells to achieve tunability (Fig. 5.40).

The spectral bands can be tuned continuously over at least $\sim 2\text{ }\mu\text{m}$ in the LW ($\sim 1\text{ }\mu\text{m}$ in the MW) and can be sharpened to FWHM $\sim 0.4\text{ }\mu\text{m}$ in the LW ($0.2\text{ }\mu\text{m}$ in the MW) by optimizing the recipe and MBE-growth conditions (e.g., modulation doping, interface interrupts, etc.). More complex wells (e.g., coupled asymmetric stepped wells) promise to at least double this tunable range. One challenge for a

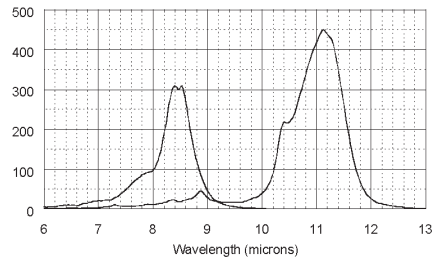


Figure 5.26. Responsivity (mA/W) vs. wavelength. This demonstrates how the broadband coupler couples light efficiently to QWIP over a wide wavelength range.

tunable QWIP FPA is to develop a broadband optical coupler (etched into the pixel top, see Fig. 5.26) that will couple light efficiently to the QWIP over the entire tunable spectral range (normal periodic 2D grating couplers have sharp spectral-coupling efficiency, which is matched to the material spectrum to maximize quantum efficiency). Figure 5.26 shows the coupling efficiency of one such coupler.

Efforts are underway to fabricate a two-channel QWIP array with each channel spectrally agile (Fig. 5.27), using the same fabrication techniques used to realize two-color QWIP FPAs (two channels with *fixed* spectral response), but with asymmetric quantum wells designed to produce the desired spectral tunability (the voltage across each channel can be changed globally by adjusting the bias at the gate of the corresponding channel input switch in the ROIC). If successful, such a two-channel spectroscopic imager will be a significant leap in the continuing evolution of infrared detector arrays.

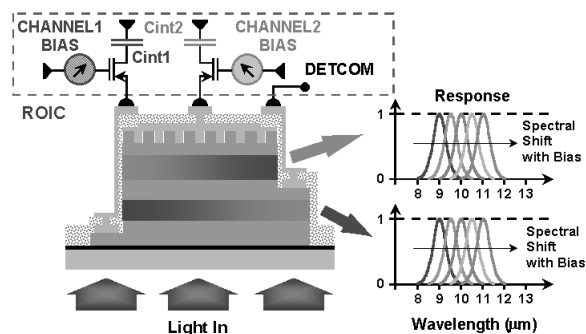


Figure 5.27. Two-channel QWIP architecture showing two vertically-stacked QWIPs connected in parallel electronically by three In bumps/pixel for simultaneous integration by two separate charge wells Cint1 and Cint2 in the ROIC. The voltage across each channel is varied independently. The Stark effect translates this bias sweep to spectral agility. Vertical QWIP stacking ensures perfect pixel registration of the two channel images. The grating etched into the pixel top diffracts and couples normally-incident light to the quantum wells.

5.3.3 SPATIAL-LIGHT MODULATORS

5.3.3.1 SLM operation

Vertical-cavity spatial-light modulators incorporate quantum wells in the cavity region of what forms an optical Fabry-Perot cavity. Electrically, the devices often consist of a p-i-n diode; quantum wells reside in the intrinsic region of this diode. The cavity can employ a top mirror composed of a quarter-wave dielectric stack (though this top mirror is optional) and a bottom mirror which is either purely metallic or a combination of metallic and dielectric mirrors. By reverse biasing the diode, the quantum well exciton's absorptive and dispersive properties are altered across a wavelength range—most importantly at the Fabry-Perot resonance wavelength. This type of behavior allows modulation and detection. As the quantum well changes its absorption/dispersion with bias, the modulators alter their reflected intensity, their reflected transmission, and/or their reflected phase at desired wavelengths.

In devices which use single, independent quantum wells, changes in absorption and dispersion with reverse bias are mediated by the *quantum confined Stark effect* (QCSE). [9] This type of modulation occurs when semiconductor energy bands change as a result of an externally applied electric field. With an electric field, three effects occur. First, the quantum well's lowest order absorptive transition—due to an exciton—shifts to lower energy (longer wavelength). Second, the overlap between electron and hole wave functions decreases. This results in a decrease in the magnitude of the lowest order exciton's absorption peak. Third, the electron and hole wave functions sample the quantum well's walls more directly. This causes a broadening in the lowest energy absorption peak. Figure 5.28 dis-

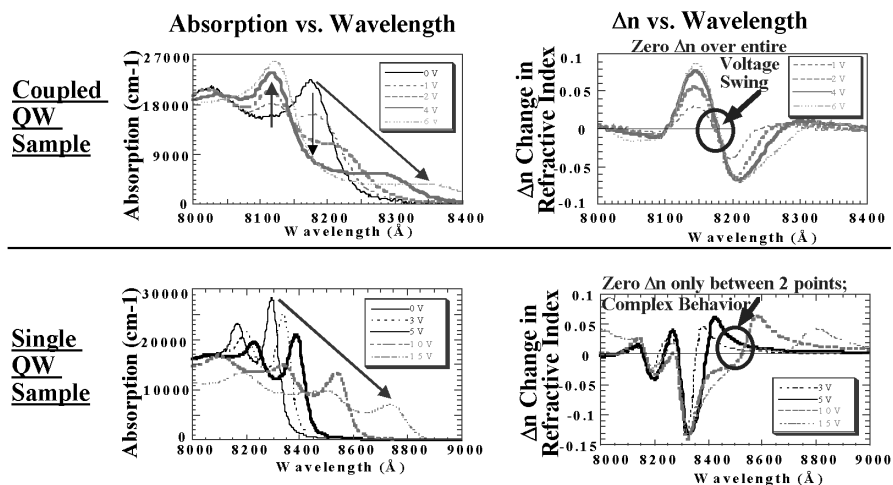


Figure 5.28. Absorption vs. wavelength and refractive index vs. wavelength at various biases for both single and coupled quantum wells.

	Forward Biased Diode	Zero Biased Diode	Reverse Biased Diode
Reflection Modulator (Arrays up to 256 x 256 and 1 x 2048)	Light Emitting 	Reflective 	Absorptive (DETECTOR)
Phase Modulator (Small Arrays 1 x 20)	Light Emitting 	Reflective (0 Phase) 	Reflective (180 Phase)
X Modulator (X-Bar) (individual pixels)	Light Emitting 	Transmissive Bi-directional 	Reflective Bi-directional
Transmission Modulator (individual pixels)	Light Emitting 	Transmissive Bi-directional 	Absorptive (DETECTOR)

Figure 5.29. Vertical cavity modulator configurations.

plays experimental absorption and refractive index change vs. wavelength for both single and coupled quantum wells.

Figure 5.29 illustrates the various configurations of vertical cavity devices we have demonstrated. Vertical cavity modulators that switch only reflected phase [10] are extremely useful for various applications [11, 12]. Such devices are referred to as phase-flip modulators. With zero applied bias, the quantum wells in a phase-flip modulator exhibit very little absorption. Consequently, light, after oscillating in the cavity region, is eventually reflected from the back mirror and exits the device with some given phase. When the phase-flip modulator is “switched,” the cavity’s quantum-well absorption is increased to the point that the device acts like a single large mirror with some residual loss rather than like a cavity device. In this case, the incident light is reflected primarily from the top mirror and the reflected phase is determined by the boundary condition imposed by the quantum wells. If, at the operating wavelength, the quantum wells exhibit the same refractive index with applied bias as with zero applied bias, then the phase

change during operation will be exactly 180° . Zero refractive index change in the quantum wells is the key for this phase modulator.

The simplest and most widely used types of pixels, however, are ones that switch between highly reflective and highly absorptive states. In this device, the cavity contains a back mirror that is much more highly reflective than its front mirror. When the applied bias is zero, the quantum well excitons provide little absorption at the operating wavelength. Incident light enters the cavity region and is eventually reflected out of the cavity by the strong back mirror. With applied bias, the cavity absorption is increased to the point where all of the incident light is absorbed as the photons oscillate in the cavity. Because of this absorption, such pixels can act as intensity modulators or as detectors.

5.3.3.2 SLM operation (binary 2D and grayscale 1D)

Pixelated, GaAs-based spatial-light modulators have been made for several years in both 1D and 2D formats [13]. 2D SLMs with as large as 256×256 close-packed pixels ($> 90\%$ fill factor) have been created. These devices have maximum allowable input data rates of 10 gigabits/sec and maximum frame rates (all pixels updated) of 312,500 frames/sec using 40-MHz-capable silicon drivers. Turn-on and turn-off times for the arrays are approximately 190 ns and 167 ns respectively. The majority of large-format (many pixel) GaAs spatial-light modulators are currently operated at wavelengths of either 855 or 860 nm (or both) with single-pixel-reflectivity changes from 0.5 to < 0.03 at 860 nm and 0.5 to 0.04 at 855 nm. A wide range of wavelengths is achievable using the GaAs/AlGaAs and other material systems. Quantum-well modulators have been made to operate at wavelengths from the visible, blue [14], to $1.5 \mu\text{m}$ [15]. With fill factors of 90%, the 10:1 array contrast ratio implies actual pixel-contrast ratios of about 15 to 21:1 (taking into account the constant interpixel reflectivity of the GaAs/air interface). The optical chips operate optimally with between 3-V and 9-V applied bias. Defect-free 256×256 arrays draw a total current of about 3 mA ($< 0.2 \mu\text{A}/\text{pixel}$). In arrays with a few “bad” pixels these pixels typically draw the majority of device current; this creates arrays with total current draw of only 30 mA. Single pixels have been experimentally demonstrated with reflectivity changes over 90% [16], contrast ratios over 100 (where contrast ratio is defined as $R_{\text{on}}/R_{\text{off}}$), and applied bias as small as 3 volts. In general, GaAs-based modulator design must account for operating voltage, maximum on-state reflectivity, optical contrast ratio, and optical bandwidth (wavelength range over which modulation is achieved). By using simulations, parameter space plots can be made for designs that use our various processing techniques. In this way, tradeoffs in device operation can be made when choosing a modulator design for a particular application.

Electrical data input into the 256×256 array is through a series of 256 input lines, each of which is clocked in parallel into the array. Thus, the array takes 256

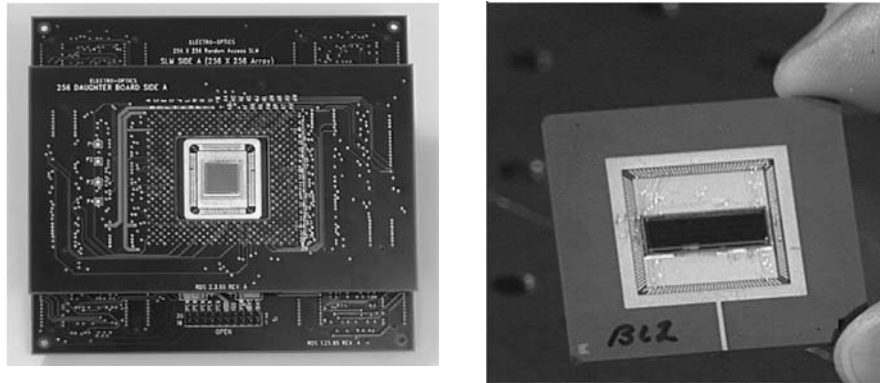


Figure 5.30. 256×256 SLM array (left) and 1×2048 SLM array (right).

clocks to load an image (the 128×128 array handles 128 input lines and takes 128 clocks to load). The CMOS SRAM driver chip is a 2×65 k configuration. This implies that two full frames of binary mode data can be stored at any one time in the driver. Electronic data is loaded into the array row-at-a-time into the “back buffer page” of the SRAM chip while another page of memory is output through the optics. When the back buffer page is fully loaded, the back buffer page can be burst to the front buffer page in just over 160 ns. The SRAM chip can support a maximum clock speed of 50 MHz; this translates into a maximum “frame rate” of 156,000 frames/second for our 256×256 arrays and over 312,000 frames/second for the 128×128 arrays. Electrical input to the chip is supported at over 12 Gb/second, and the page-swapping speed allows burst-data optical outputs of almost 410 Gb/second.

In addition, 1D arrays up to 1×2048 in length have been developed. These arrays have integrated digital to analog converters (D/A) in the CMOS chip. Each pixel has its own 6-bit D/A and allows multilevel outputs to be passed onto each pixel. The intensity of modulated amplitude (phase) or emitted light corresponds to the value of the data passed to that pixel. These devices, known as “grayscale transceivers,” have been targeted toward spectroscopy, beam-steering, image-formation, and image-recording applications.

5.3.3.3 Eight-bit grayscale 2D MQW-SLM development

Grayscale capability in a pixellated optoelectronic array is a critical feature for improving system performance and diversity of applications. A prototype motherboard, two custom 1.2-micron CMOS ASICs, and a 128×128 hybrid device were fabricated and assembled to demonstrate the 2D grayscale SLM concept. In the 2D binary mode devices, all storage and signal conditioning of pixel data is done

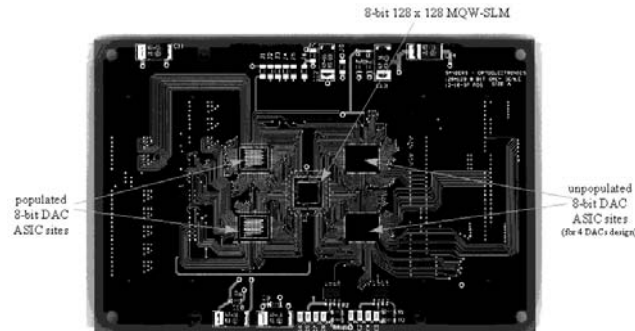


Figure 5.31. 8-bit gray-scale 128×128 MQW-SLM Motherboard.

on the silicon ASIC directly under an optical pixel. Using an existing 128×128 optical chip as a baseline mechanical form factor, the real-estate available at each binary pixel-driver interface was unable to support the increase in circuit size required for the digital-to-analog (DAC) converter necessary for grayscale performance. The real-estate problem was solved by keeping only the analog storage element at each pixel interface on the SLM driver, and moved the DAC function was moved into an external ASIC colocated with the driver on a circuit card assembly.

The center of the 8-bit grayscale 128×128 MQW-SLM motherboard is the location for the SLM hybrid device (analog driver ASIC and optical chip). The components surrounding the hybrid device in the Fig. 5.31 are the 8-bit digital-to-analog (DAC) ASICs. This design allows either one, two (current device), or four DACs to drive the SLM hybrid. This flexibility was implemented to deliver either maximum data to the SLM (256 inputs with 4 DACs) or operation at low frame rates (1 DAC). The chosen option is wire-bond selected and dependent upon the frame speed desired. The four ASICs and SLM hybrid are interconnected through circuit traces within the motherboard. Wire-bond locations at each ASIC location complete the circuit interconnect.

The DAC ASIC contains 64 8-bit DACs to convert data using eight 8-bit ports. Eight consecutive writes of digital data to the DAC fill all of the storage registers for each DAC circuit (64) and initiate a conversion process. When the data is converted, there is enough analog data to fill 64 pixels on the SLM driver. DAC I/O control circuits write the analog data out to the SLM driver in a sequence based upon the number of DAC ASICs employed. This process continues until all pixels on the driver ASIC are updated. The grayscale motherboard is designed to connect directly to the existing 256×256 binary SLM control module. The new device was designed to function with the same host driver module and control signals as the previous 2D binary devices. Since the DAC function is

independent of the specific SLM driver, it can be used in other 2D designs requiring grayscale accuracy.

5.3.4 RESONANT-CAVITY LIGHT-EMITTING DIODES

Similar to the SLM, resonant-cavity light-emitting diodes (RCLEDs) are composed of p-i-n diodes. The p and n regions of the diode either consist of or are adjacent to partially reflective surfaces. The active region, which contains quantum wells, is located in the intrinsic region of the diode. The entire structure forms a Fabry-Perot resonant cavity. RCLEDs have higher efficiency, higher intensity, lower divergence, and higher spectral purity than conventional LEDs. RCLED arrays have been produced using GaAs/AlGaAs as well as GaN/InGaN quantum wells.

The cavity's resonant wavelength determines the wavelength around which the RCLED "operates." The quantum well's gain region is made to coincide with the cavity's resonant wavelength. The 128×128 blue and red MQW-RCLEDs grown by MBE, shown in Fig. 5.32, generate 30–40 microwatts of output power per pixel and operate at a frame rate of 250,000 frames/second. Figure 5.33 depicts the output spectra of RCLEDs that have center wavelengths of 710 nm (red) and 460 nm (blue) and full width at half maximum (FWHM) of about 10 nm (red) and 10 nm (blue). 256×256 RCLEDs at 860 nm have also been demonstrated. This work represents the demonstration of the largest-format RCLED arrays that also operate at the fastest frame rates.

5.3.5 SMART / HIGH-SPEED DETECTORS

While spatial-light modulators take maximum advantage of the switching capabilities of vertical-cavity modulators and the same types of arrays can also take advantage of the emitting properties of the pixels as LEDs, smart detectors take maximum advantage of the detection capabilities of these pixels. In such a configuration, a dense, packed array of vertical-cavity devices is hybridized to a

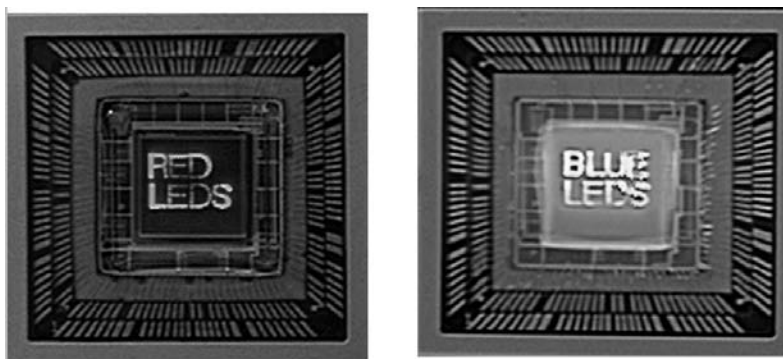


Figure 5.32. 128×128 red and blue RCLED arrays.

specialized CMOS chip. The goal for many applications of smart detectors is data reduction—for example, if you are computing optical correlations at a rate which generates output frames at 10 Gb/second, an electronic postprocessor is not able to work efficiently with this quantity of data. A smart detector can reduce passed data through compression, or by selecting certain frames to pass to a post processor, or by selecting certain regions of the image that merit closer inspection and only transmitting that data. In addition, smart detectors for imagery can contain filter planes stored in the CMOS so that pixel-by-pixel thresholding can be performed on the imagery data. This thresholding is particularly useful for target recognition systems employing an optical correlator. Furthermore, simple signal processing can be performed within the detector circuitry itself; for example, the detector could perform local averaging, edge enhancement, thresholding, contrast enhancement of images, as well as noise reduction.

In addition to the smart processing available with these hybrid detector arrays, high speed is also available. Smart detector arrays may not take full advantage of their speed when being used as a scene-imaging sensor due to the low ambient light levels (low photon density) typically observed. However, in closed systems where a light source is provided, say by a local laser, optical power density can be adjusted. Single-element bulk GaAs-based detectors are commercially available at speeds of over 60 GHz. High-speed detector arrays will likely not achieve such speeds in the near term as the ultimate speed will be limited by the VLSI circuitry that is bump-bonded to the optical pixels. However, given the progress in optical power routing and optical receiver circuitry design as employed in optical transceivers it is believed that speeds of 400 MHz to 1 GHz should be attainable. Thus the 300,000 frame/sec speeds demonstrated in 2D optical correlation should be easily achievable. In 1D detectors, the frame rate can be as high as the VLSI clock speeds. These detectors can also be used for parallel access to optical storage media such as CD and DVD-ROM.

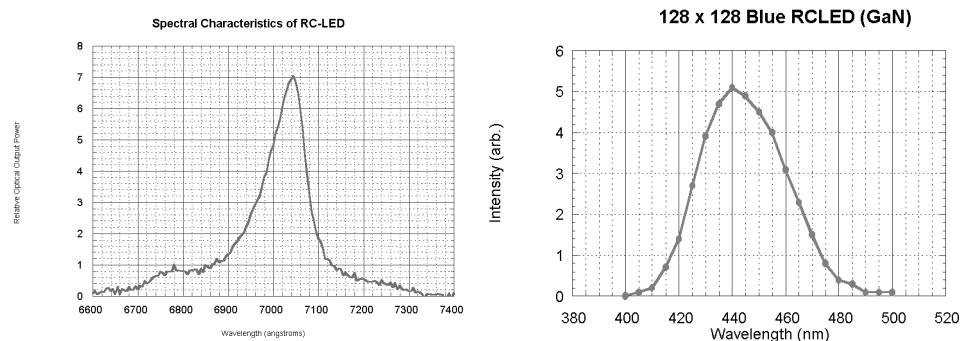


Figure 5.33. Output spectra of red and blue RCLED arrays.

5.3.6 APPLICATION: HIGH-SPEED OPTICAL CORRELATOR

The current MQW-SLM-based optical correlator provides high-speed optical correlation at 250,000 frames/second. An image of a helicopter and filters of a rotating helicopter are used as an input and reference, respectively, for the demonstration of high-speed correlation. The reference filters used for the demonstration are amplitude-encoded binary-phase-only filters (AE-BPOFs). AE-BPOFs provide high discrimination capability in the Vander Lugt correlator. AE-BPOF is obtained by computing the Fourier transform of the reference image and binarizing the phase information. Phase values greater and less than π are encoded to amplitudes of “1” and “0,” respectively. In this demonstration, the reference SLM operates with a clock rate of 32 MHz and cycles through 121 filters of images in different orientations. Since an MQW-SLM requires 128 clocks (128 pixels parallel load) to load a single 128×128 image, the correlator system is expected to detect a correlation peak every $484 \mu\text{s}$. A small aperture on the detector improves the extinction ratio of the correlation signal. The experimental data is compared with a MATLAB simulation throughout the demonstration. Figure 5.34 shows experimental and theoretical output of the correlator for a given input helicopter target and reference filters. In order to simulate the effect of the aperture in front of the detector, a 3×3 -pixel area around the correlation peak is integrated to obtain the theoretical output. It is clear in Fig. 5.34 that the experimental data is well matched with the simulation in extinction ratio and angular dependence. As expected, the correlation has a narrow angular bandwidth ($\pm 3^\circ$ with respect to in-plane image rotations) with a simple AE-BPOF.

The optical correlator is based on the Vander Lugt correlator architecture in which an input target is compared in a Fourier plane using Fourier filter reference images. 860-nm continuous-wave (CW) light from a diode laser is expanded and collimated for uniform illumination over the entire input SLM. Patterned and reflected light from the input SLM is Fourier transformed onto a reference SLM by

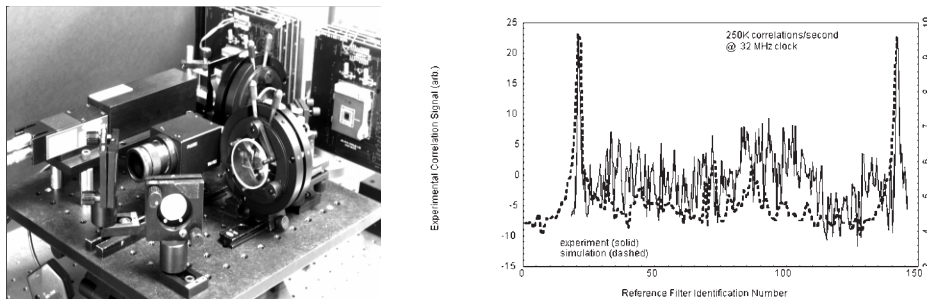


Figure 5.34. Optical correlator and high-speed experimental output.

a Fourier transform lens. The product of the Fourier-transformed input image and a reference filter image is inversely Fourier transformed by another lens. A CCD camera with a frame rate of 30 frames/second captures the spatial information of the output correlation while a high-speed (50 MHz) silicon detector detects the temporal correlation information of a small area. For high-speed target detection, the correlation signal is captured and displayed in real-time with the aid of a 12-bit, 5-MHz analog-to-digital (A/D) converter. The size of the current optical correlator using off-the-shelf optics and subcomponents is $13'' \times 20'' \times 6''$ (also shown in Fig. 5.34). The hardware and software controls of the correlator are integrated in a menu-driven desktop format using a Labview graphical user interface.

5.3.7 TARGET RECOGNITION

The optical correlator has also been examined for versatility using inverse synthetic-aperture radar (ISAR) test imagery provided by U.S. Army Aviation & Missile Command (AMCOM). The test imagery consists of two sets of imagery for each target; imagery with 10° and 12° “look-down” angles. Each data set consists of target imagery in different in-plane orientations. Three similar 8-bit grayscale test targets are used for the evaluation. First, a test target is binarized in such a way that the target and background are set to low and high reflection states, respectively. In a real system, this preprocessing could be performed on the SLM ASIC itself. The matched AE-BPOF for each target is obtained from the binarized target images. This filter is compared with a matched AE-BPOF generated directly from the 8-bit grayscale image and found to be the optimum choice for a binary modulator. No cross-correlation among the three different targets is observed despite the similarity of the targets. Strong correlation between a target and its matching filter in each orientation is observed. The Horner efficiency [9], the highest to secondary peak ratio, is measured to be 2.24:1, similar to the simulated value (2.27:1). We also investigate the use of 10° (look-down angle) AE-BPOF for a 12° (look-down angle) target identification. Experimentally, 10° matching filters in some (but not all) of the target orientations show strong correlation with a 12° target; simulations also show correlation between 10° filters and 12° targets, but not in all in-plane orientations. This discrepancy with the simulation in terms of matching target orientations prompts further investigation in this matter in order to make any conclusion.

5.3.8 APPLICATION: HYPERSPECTRAL DATA PROCESSING

A hyperspectral scene contains information about a scene in which a 2D image of the scene is provided in each of a large number of wavelengths. In order to determine information about the composition of the materials in the scene, each of the “pixels” of the scene must be analyzed across all of the spectral features to determine which materials exist in the various regions of the image. Many of the

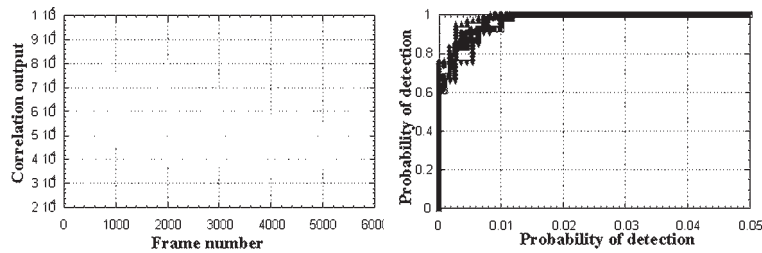


Figure 5.35. Hyperspectral data processing results (a) correlation output vs. time—five peaks correspond to detection of blocks of 50 target pixels, (b) ROC curve for five repetitions of data set.

algorithms required to process this data depend on an inner-product, cross-correlation, or template-matching operation. Optical devices are used to perform this processing *three to five orders of magnitude* faster than the corresponding electronic processing approach. This optical processing is intended to reduce the amount of data that the electronics must process, not replace the electronics.

In Fig. 5.35, we show some sample hyperspectral processing results. What the data in Fig. 5.35 represents is a task in which 50 pixels of target data are discriminated from 900 pixels of clutter data. The left side of the figure shows the correlation output of the system versus time, where the hyperspectral data are loaded onto the input spatial-light modulator in rapid succession. The filter plane of this correlator contains an amplitude-encoded binary-phase-only filter representation of typical target data. The entire data set is processed five times to demonstrate repeatability (thus the five 50-data point peaks in the output). The right side of the figure is an ROC (receiver operating characteristic) curve of Pd vs. Pfa (probability of detection vs. probability of false alarm). This curve demonstrates that all target data can be detected with a false alarm rate of 1%. Thus, if the current optical correlator system is used as the front end to an electronic processing system, we reduce the amount of data that must be processed electronically by a factor of 100 (*and thus increase total processing time by nearly a factor of 100*). The next-generation system will increase this factor to 1,000–10,000.

5.3.9 APPLICATION: OPTICAL BEAM STEERING

Acousto-optic modulators have been widely used as a method of beam steering and beam scanning. It scans by having a periodic change in optical properties inserted into it via the propagation of an acoustic wave along the length of the crystal. This periodic optical change creates a diffraction grating which “steers” an optical beam to a particular angle. Unfortunately, the acoustic wave is a traveling wave. Thus the proper periodic pattern is only available for a brief time (after which the wave is reflected, dissipated, and otherwise distorted). Furthermore, the

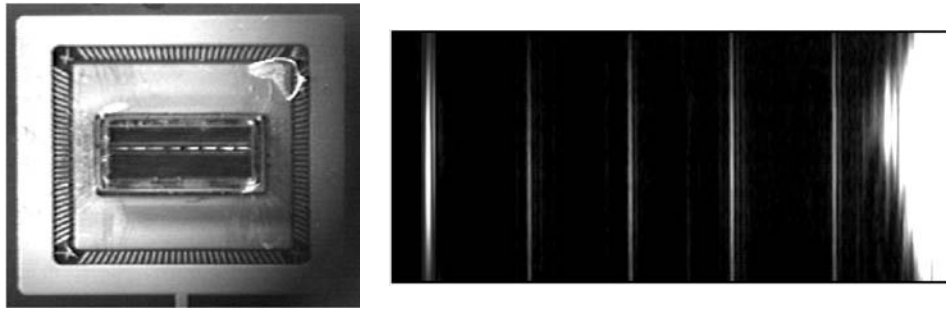


Figure 5.36. (a) 1×512 grayscale SLM, (b) 1D beam steering.

speed of sound in the crystal limits the speed at which the A-O crystal can be updated to a new periodicity (and hence steer optical beams to a new angle).

1D and 2D MQW modulators can be used in a variety of beam-steering applications. MQW SLMs permit random updating of each pixel from frame to frame, enabling a greater diversity of applications—patterns can be rapidly and randomly switched, or remain constant for any time period. Beam-steering devices enable numerous military optical signal- and image-computing systems (such as SAR radar-processing systems, infrared countermeasures, laser communication, laser radar, target designation) in addition to more commercial applications varying from laser printing to permutation and broadcast free-space optical links.

An amplitude-encoded reflective diffraction array can be implemented using 1D or 2D MQW SLM. In the far field of a Fourier transform lens, the grating provides zero-th order (DC) and higher-order diffraction peaks. These peaks are steered by varying the pattern on the SLM. For example, if a sinusoidal amplitude grating pattern $1/2 + m/2 \times \cos(2\pi x f)$ is represented on the modulator array, the Fourier transform lens yields (in far-field amplitude) three delta functions $1/2 \times \delta(s) + m/4 \times \delta(s + f) + m/4 \times \delta(s - f)$ where m represents the peak-to-peak amplitude change for the grating, f represents the grating spatial frequency ($30 \text{ } \mu\text{m}^{-1}$ in this example). By changing the grating periodicity, the steered angle changes. Vertical-cavity phase modulators can also be used to implement the diffraction grating, where a phase modulator has large reflectivity with bias-controlled phase. [10] Phase modulators can eliminate the DC peak in proportion to the degree of phase precision of the modulator.

Figure 5.36(a) shows our 1D grayscale amplitude SLM (1×512 pixels) displaying a sinusoidal steering pattern. Figure 5.36(b) shows the resulting steering of first order diffraction peaks (temporally integrated output of five steering patterns). Each pixel was capable of displaying 6 bits (64 levels) of grayscale. 256 linearly spaced angular points were resolvable. The maximum deflection angle is 1.64 degrees (28.6 milliradians) with an angular resolution of 0.0064 degrees (0.011 milliradians).

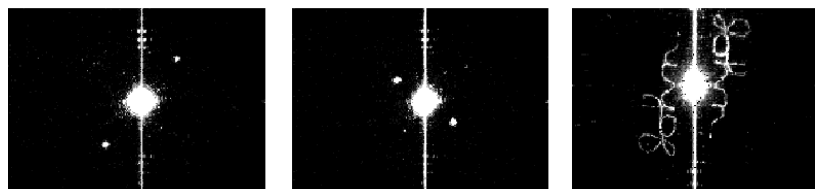


Figure 5.37. 2D beam steering. (a) Beam steering to one location, (b) beam steering to second location, (c) temporally integrated 125-kHz beam steering.

Figure 5.37 shows 2D beam steering via a 128×128 binary amplitude SLM. The maximum diffraction angle is ± 0.6 degrees (± 10.5 milliradians); the angular resolution is 0.0095 degrees (0.16 milliradians). Approximately 30% of the light is steered to the two first order diffraction peaks. The use of phase-modulating SLMs would increase this value. Figure 5.37(a) shows an example frame pattern placed on the SLM; Fig. 5.37(b) shows the experimental diffraction pattern generated from this pattern. Figure 5.37(c) shows the temporally integrated output as a series of 400 frames steer the first order diffraction peaks in 2D. These 400 frames are cycled through at the rate of 125,000 per second.

5.3.10 APPLICATION: HIGH-SPEED OPTICAL PRINTING

Computer-to-plate printing is now undergoing a rapid technology-driven revolution. Print-plate media that can be written onto and used in printing without photographic processing steps has provided a new market niche: direct-to-plate printing. A critical performance criterion for computer-to-plate printing is the rapid, high-contrast modulation of high-power printing-laser light onto the printer plate. Multiple-quantum-well SLMs currently afford the highest-performance SLMs for this application. The direct-to-plate application is a direct-plate printer in which areas of a thin titanium plate are vaporized by exposure with a highly focused light beam. The titanium plate is then used in direct printing of large-format (poster-size) print jobs. Many companies currently use acousto-optic modulators (AOMs) to shutter the ablation light onto the titanium plate. Single-channel AOMs are inexpensive and afford good power densities and contrast ratios (50:1), but provide only modest modulation speeds and resolutions (hundreds of spots/mm). On the other hand, multiple-channel AOMs that provide the larger numbers of resolvable spots required for high-resolution printing are expensive.

The direct-printing system requires that substantial optical power be delivered to the print plate in a pattern that can be modulated or controlled. Reflective or transmissive modulators can be used to deliver patterned optical power to the plate. Reflective-mode modulators provide several key advantages over transmissive modulators. Reflective-mode devices can be tuned after growth to match the

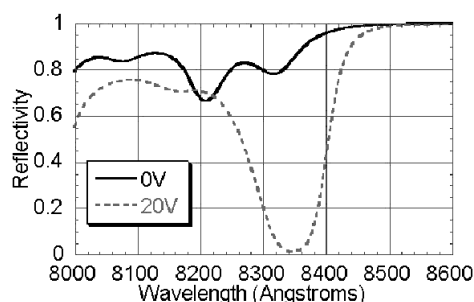


Figure 5.38. Wavelength vs. reflectivity for optical printing SLM.

specific operating optical wavelength; transmission-mode devices require that the MBE growth of the material be exact to the design dimensions to better than 1% for proper device operation. Reflective-mode modulators permit additional design flexibility that can be used to provide higher on-state reflectivities, larger optical bandwidths, and higher contrast ratios; transmission-mode devices require additional design iterations to achieve similar characteristics. In Fig. 5.38 the on-state (0 V) and off-state (20 V) reflectivities are plotted for a reflective-mode modulator design similar to that required for this application. The reflective-mode design provides a large optical bandwidth (permitting laser tuning or reduced temperature sensitivity), in addition to reduced incidence-angle dependence. These devices have been characterized to measure their optical power-handling capability with and without active (thermoelectric) cooling devices. With active cooling, reflective-mode modulators have operated with over 8 kW/cm² optical power density (Fig. 5.40).

5.3.11 APPLICATION: OPTICAL STORAGE

High-speed and high-volume data access are enabled by smart-pixel technology. The use of smart-pixel arrays for holographic storage, whether thin-film or vol-

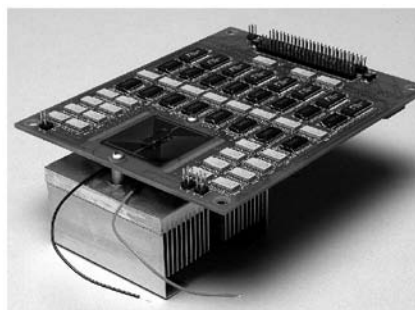


Figure 5.39. Optical printing module (thermoelectrically cooled).

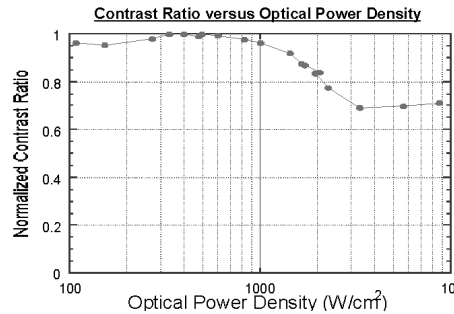


Figure 5.40. Optical power handling of the array.

ume, will enable ultra-high-speed data access. Such memories take advantage of the large-format pixel arrays and the ability for individual pixels to act as both SLM and RCLED. Large-format pixel arrays offer the potential for improvement in data retrieval from memory at speeds four to nine orders of magnitude faster than the fastest competing electronic approaches. For example, a $1 \times 32,000$ 1D detector array has enough pixels to read the entire radius of a CD-ROM. Assuming a modest frame rate of 40 MHz, the data throughput potential is 1.3 Tb/s—an entire CD-ROM could be read out in a single rotation of the disk, taking less than a millisecond. Such a system could be developed today using existing technologies—no breakthroughs would be required. In the longer term, two-photon materials should allow storage densities on the order of 100 Mb/cm^3 (compared with 6.5 Mb/cm^2 for a CD-ROM). Arrays of modulators and detectors would allow for page-at-a-time access to such memories with microsecond access times. Again, the advantage would be speed of data access if not overall storage density; the advantage of optical material over magnetic material is the ability to bring data to and from CMOS chips thousands or millions of bits at a time. To store and retrieve large amounts of data rapidly, and to update stored look-up tables, fixed masks are not acceptable. Instead, high-speed electrically-addressed SLMs and RCLEDs coupled with detectors are required to provide programmable readout and update of the optical memory. Figure 5.41, currently in process of integration for bench-top demonstration, shows the modified optical-memory look-up-table (LUT) implementation using MQW-SLMs and a two-color photorefractive storage method. Here, the input data (x) and of the desired output of the look-up table [$f(x)$] are input to the signal arm and reference arm of the optical memory respectively using reflective MQW-SLMs. The two-color photorefractive storage method provides the additional benefit of nondestructive readout, since both the energizing light and recording light wavelengths are necessary to modify the stored data. The data loading to the MQW-SLMs is accomplished using a high-speed digital IO board custom designed to provide data to the MQW-SLMs.

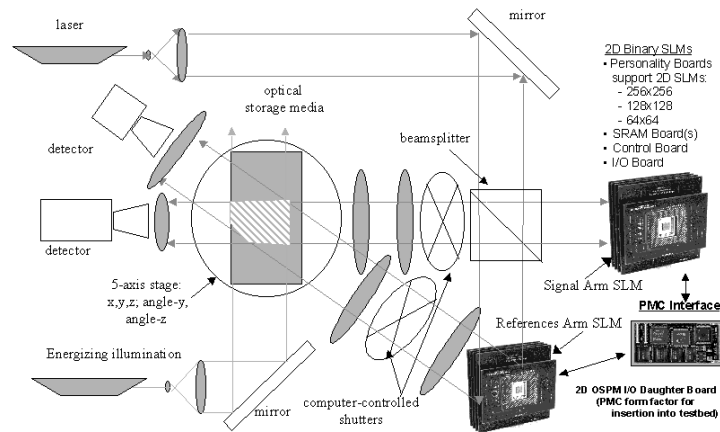


Figure 5.41. Volume storage LUT implementation using multiple-quantum-well spatial-light modulators (work in progress).

5.4 SUMMARY

Very-large-scale heterogeneous integration, while not widespread in the commercial world yet, is still a technology ready for deployment. From a technological viewpoint, the main attraction of optoelectronic heterogeneous integration is the ability to move data between electronic chips quickly. At the most basic level, this means

- *Sensing* data from the world and bringing it into chips through arrays of devices (bring millions of bits per clock cycle into CMOS chips as in focal plan arrays)
- *Processing* data (by optically transferring it between electronic chips in parallel chunks and performing operations on it either in the electronics or in the optical domain. (Optical-electronic processing systems run customer algorithms hundreds of times faster than the fastest pure-electronic computers can run them; 10,000 times improvements will be demonstrated in the near term))
- *Routing (moving)* data (as in pure transceiver modules where no processing is performed or as in beam-steering modules where data can be directed for free-space data communications)
- *Rapid storage and access* of data (e.g., as in the ability to use existing technology to read a CD-ROM in about the same time as today's CD-ROMS can access 1 bit of data or in the ability to perform printing at ultra-high speed.)

These four areas (sensing, processing, routing, and storing) are the four pillars of modern computing and communication systems; very-large-scale heterogeneous integration has a place in each of them.

The entry of these technologies into volume applications is slowly beginning. Today, the commercial world employs small-scale heterogeneous integration techniques in some areas (as in the use of small-scale flip-chip bonding in Pentium III processors to get a few hundred bump bonds). As the comfort with these technologies grows and systems are designed from scratch to include these technologies, there will be a natural evolution to include these technologies. Unlike the electronic integrated circuit that took the better part of three decades to gain widespread acceptance, very-large-scale heterogeneously integrated modules seem poised for acceptance in a far shorter time.

REFERENCES

- [1] John Baliga, associate editor, "Ball grid Arrays: the high-pincount workhorse," *Semiconductor International*, Vol. 22, No. 10, p. 86 (1999).
- [2] J. F. Lancaster, *The Metallurgy of Welding, Brazing and Soldering*, J. W. Arrowsmith LTD., Bristol, U.K., pp. 106–117 (1965).
- [3] M. Sickmiller, A. Pike, S. Jacobs, L. Vo, P. Zampardi, G. Sullivan, B. McDermott, International Conference on GaAs Manufacturing Technology, Digest of Papers, 187 (1999).
- [4] R. E. Williams, "Gallium arsenide processing techniques," Chap. 16.2, Artech House, Dedham, MA (1984).
- [5] Bock, A. Grub, and H. L. Hartnagel, *J. Electrochem. Soc.*, Vol. 137, 3301 (1990).
- [6] R. J. Olson, Jr., M. F. Taylor, R. J. Williams, T. S. Faska, and M. Sundaram, International Conference on GaAs Manufacturing Technology, Digest of Papers, 179 (1999).
- [7] For a comprehensive review of the physics and technology of QWIPs and one-color QWIP FPAs, see B. F. Levine, *J. Appl. Phys.*, 74, R1 (1993); and references therein.
- [8] M. Sundaram *et al.*, in Proc. IRIS Spec. Grp. Mtg. Dets., NIST, Boulder, CO (August 12–13, 1998).
- [9] D. A. B. Miller, "Electric field dependence of optical absorption in quantum well structures: physics and applications," *Proceedings of the SPIE—The International Society for Optical Engineering*, Vol. 792, pp. 176–177 (1987).
- [10] J. A. Trezza, J. S. Harris, Jr., "Creation and Optimization of Vertical Cavity Phase Flip Modulators," *Journal of Applied Physics*, Vol. 75, pp. 4878–4884 (1994).
- [11] B. Pezeshki, R. B. Apte, S. M. Lord, J. S. Harris, Jr., "Quantum Well

- Modulators for Optical Beam Steering Applications,” *IEEE Photonics Technology Letters*, Vol. 3, pp.790–792 (1991).
- [12] J. A. Trezza and J. S. Harris, Jr., “Two-state electrically controllable phase diffraction grating using arrays of vertical cavity phase flip modulators,” *IEEE Photonics Technology Letters*, Vol. 8, pp. 1211–1213 (1996).
 - [13] T.L. Worchesky, K.J. Ritter, R. Martin, B. Lane, “Large arrays of spatial light modulators hybridized to silicon integrated circuits,” *Applied Optics*, Vol. 35, pp. 1180–1186 (1996).
 - [14] S. Y. Wang, P. Thompson, G. Horsburgh, J. T. Mullins, I. Hauksso, K. A. Prior, B. C. Cavenett, “A ZnSe/ZnCdSe quantum well symmetric self-electro-optic effect device operating in the blue-green region,” *Journal of Crystal Growth*, Vol. 138, pp. 647–651 (1994).
 - [15] Sam-Dong Kim, J. A. Trezza, J. S. Harris, Jr. “Observation of 1.5 μm quantum confined Stark effect in InGaAs/AlGaAs multiple quantum wells on GaAs substrates,” *Journal of Vacuum Science & Technology B*, 13, pp. 1526–1528 (1995).
 - [16] R. H. Kingston. “Detection of Optical and Infrared Radiation,” Springer Series in Optical Sciences, Vol. 10, Edited by David L. MacAdam, pp. 10–23 (Springer-Verlag, New York, 1978).

Molded Integration of Optical, Optoelectronic and Electronic Components

Volkan H. Ozguz and John C. Carson
Irvine Sensors Corporation

6.1 INTRODUCTION

The demand for more functionality in smaller sizes fuels the development of integration technologies. The classical example is the rapid growth in silicon transistor integration, identified by the semiempirical Moore's Law, which states that "the number of transistors on a given amount of silicon doubles every two years." However, silicon can not fulfill all functionality needed for a stand-alone system that consists typically of a sensor (electrical, mechanical, optical, chemical, etc.), a processing and decision circuit (analog, digital, mixed mode) and an actuator (electrical, mechanical, optical, chemical, etc.). The integration of a high level of functionality is not trivial due to non-silicon-based devices such as GaAs and nonplanar devices such as micromechanical structures. Even simpler cases of integration of memory and processors or analog and digital circuits are not straightforward due to fine-tuned processing techniques that are not directly compatible. The efforts toward integration of dissimilar materials and devices lead to the concept of "heterogeneous integration." A planar version of heterogeneous integration will yield "systems-on-a-chip," where more complex functionalities can be obtained in a single piece of silicon.

Optical (e.g., refractive or diffractive components, waveguides) optoelectronic devices (e.g., lasers, detectors, etc.) are based on nonsilicon materials and, mainly, compound semiconductors of type III-V (e.g., GaAs). Optical devices have advantages especially in communication due to large bandwidth. The integration of optoelectronic devices with silicon circuits is highly desirable for developing optically interconnected systems.

There were several attempts to join optoelectronic devices such as vertical-cavity surface-emitting lasers (VCSELs), light-emitting diodes (LEDs) and detectors with silicon circuits. These attempts can be grouped into three categories:

1. *Bump or pad bonding*: joining metal pads of the plane comprising optoelectronic elements with the pads of silicon circuits with or without reflowing the

metal. This technique is the extension of conventional flip-chip bonding and is usually restricted to joining single pairs of components. In addition, the resulting structure is thicker, and mechanically fragile unless supported by underfill material. However, due to its simplicity and maturity of the supporting techniques, it is widely used. Bump pitches on the order of 25 microns can be achieved if dense interconnects are needed. A variant of this approach has been developed by Lucent labs in which the bumped layer is further thinned for optical emission through the substrate [1–12].

2. *Thin-film bonding*: thinning either silicon or optoelectronic devices to less than 10–15 μm and joining with the thick substrate containing silicon or other type circuits using adhesive bonding, metal bonding or weak-force (Van Der Walls) bonding. There are several variations of this technique differing in how thin film is obtained (epitaxial lift-off, cleaving, grinding) and carried onto the host substrate (direct, via rigid transfer substrates, via polyimide membranes) [13–20]. In terms of final dimensions, this technique results in structures closest to monolithic integration. However, obtaining and bonding the thin film is not trivial especially if the receiving surface is not flat. Most of the silicon circuits have to be specially prepared since typical integrated circuits have surface topologies on the order of few microns resulting from underlying structures.
3. *Wafer fusion*: fusing two blank wafers of dissimilar materials (e.g., silicon and GasAs) together followed by respective fabrication of devices in each layer. In some cases, the devices in one of the wafers may be ready processed. The major difficulty in this approach is incompatibilities in the subsequent processing steps with the devices already processed. For example, heating above 450°C will affect the first layer metal (aluminum), and may cause spiking in junctions or changes in the interface state density at the gate oxide [21–26].

For applications where more than two types of material need to be joined together in a small form factor, a new approach is needed. Joining electrical and optical components in a transparent-molding compound can circumvent the difficulties in material incompatibilities and related process development in other approaches.

6.2 DESCRIPTION OF THE DEVELOPED MOLDING TECHNIQUE

The quest for dense and multifunctional packaging can be approached by combining components in 2D (planar packaging) and 3D (stacking). Multichip modules are an answer to 2D packaging. Stacked packages and stacked-die tech-

nologies address the need of 3D packaging. However, historically, all technologies addressing more than one component either in 2D or 3D packaging formats were confined to niche markets due to the complexity and cost of the existing approaches. Furthermore, these technologies lacked the versatility of accepting components of various origins such as MEMS or optoelectronic devices. The drawbacks of existing approaches are mainly due to the use of nonstandard processing steps and tools. One existing approach uses packaged chips that suffer from pin-density limitations and constrains the physical size. Another existing approach is limited to chips of the same size. In both cases, process steps are complicated and the flexibility of selecting components of various form factors and technology origin is lost. Recently, ISC developed a versatile, highly manufacturable packaging approach to allow the integration of silicon integrated circuits of different size and origins referred to as *neostacking* [26–28]. In this approach, individual components are joined in a molded frame that will generate a wafer-like structure (*neowafer*). Molded components can be modified and diced (*neochip*, *neolayers*) for next-level packaging or stacking. Components of different sizes, functionality, and technology can be joined together. This method provides several benefits over conventional stacking technologies. It allows the use of ICs which are only available as bare dies. Conventional ISC processes involve building a wiring board onto the wafer (referred to as modification), and subsequent thinning, requiring the starting material to be at wafer form. Furthermore, the laminated layers must all be of the same size. This means either being limited to homogenous stacks with the same die in all layers, or dicing wafers with smaller ICs to the same size as the largest IC, thereby greatly reducing wafer yields. Neochips surmount these difficulties by encapsulating dies in such a way as to effectively convert them back to wafer form. As such, they may be modified, thinned, and diced to a common dimension even if the original dies were of different size.

Our approach is illustrated in Fig. 6.1. Figure 6.2 shows the cross sections of a typical layer with multiple dies and two-layer interconnection. Figure 6.3 is a photograph of an actual multichip layer with two-layer metalization. The process starts with the tested chips. Gold thermosonic ball bonds are added on the existing pads, except the tail of the wire is cut off by the automated bonder, leaving a gold bump on the pad. The use of gold bumps that are bonded on the I/O pads allows a very efficient, negligible resistance, and highly reliable interface which is common in any electronic wirebonding application. Concurrently, polyimide films are spun on Al shims. Polyimide is exposed and developed using a mask made to match the pad locations and dimensions of the dies to be used. The developed assembly is bonded on a glass-epoxy-based carrier with openings that match the size of the chip (or chips) using a heat-assisted bonding process. The Al layer is etched from inside the cavity to expose polyimide with holes. The chips with gold bumps are inserted in the matching hole pattern in the polyimide inside the cavity. The cavity is filled with potting epoxy and

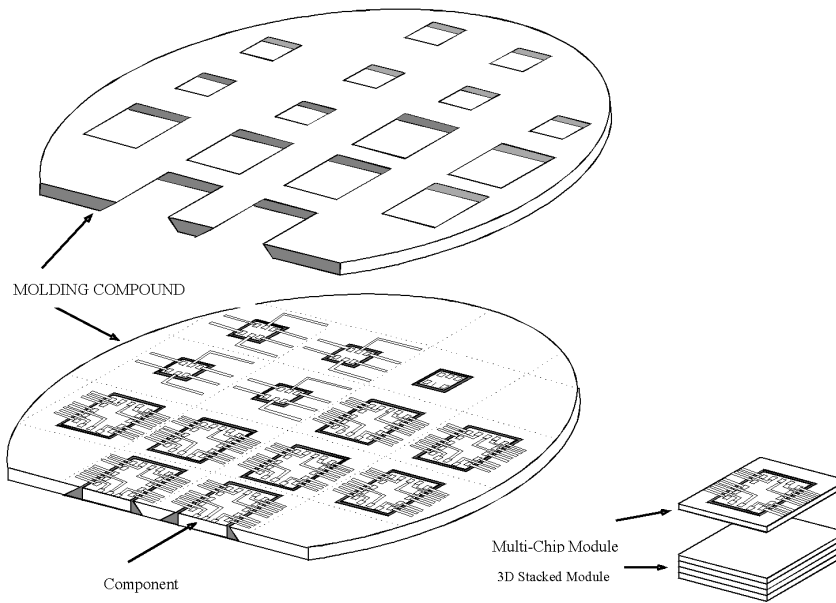


Figure 6.1. The concept of molding various components in a wafer-like structure. The known-good die and other components are embedded in a molding frame which is filled with a potting material. The multichip assemblies are obtained by dicing. These assemblies can be stacked further if needed.

the assembly is cured. The portion of the bumps protruding through polyimide at the front face is lapped to expose a planar surface of polyimide with planar gold pads flush with the surface. Metal traces are added using metal deposition and photolithography at the front surface, extending the leads of the die (or to connect dies). The metal-interconnect traces are implemented using a physical-deposition technique (e.g., evaporation or sputtering) to capture the bumps. Therefore, that metallic interface also is highly reliable with negligible electrical resistance (similar to multilayer interconnects in silicon technology). A protection layer of polyimide is added and vias are opened using a mask corresponding to the location of the new pads. The last steps can be repeated to create a multilayer connection structure. The overall assembly is ground from the backside to the final desired thickness and diced to the final dimensions using industry standard equipment. The unused portions of the frame are diced out during the operation. The advantages of this new technique are:

1. Step-coverage problem is eliminated by allowing a planar interconnection surface.
2. A good electrical contact is insured by exposed gold flush with polyimide surface.

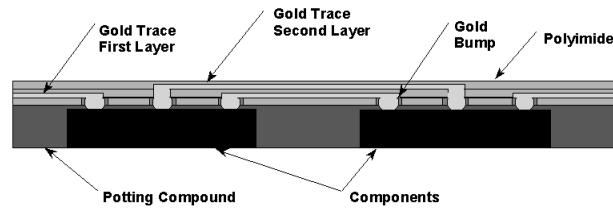


Figure 6.2. Cross section of a layer with multiple dies embedded in potting compound.

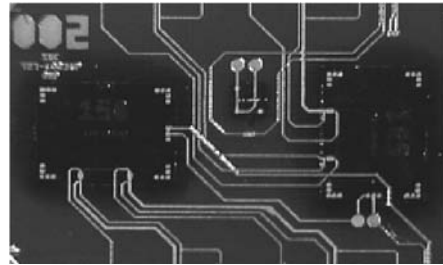


Figure 6.3. A proof-of-concept implementation with three components in one layer embedded in potting compound and two-level metalization.

3. Die-to-die alignment is achieved for multiple chips for future system-in-a-stack applications.
4. Several dies of the same type or different dies can be handled at a given time, allowing a large wafer-like structure that increases the manufacturability.
5. The disposable frame acts as a cost-effective molding fixture.

These steps are based on standard pick-and-place and molding techniques allowing a low-cost, highly manufacturable approach to multichip packaging. Therefore components of different sizes and multiple chips can be accommodated in the same layer. The materials used for molding are selected specifically to minimize CTE mismatches, component movements and to allow excellent surface planarity with mechanical stability. The material selection is not limited to conventional encapsulation materials such as phenolic or anhydride epoxies commonly used for plastic encapsulation in the electronics industry. The molding materials considered in this approach are derived from optical-grade epoxies and liquid-crystal polymers (LCP). Materials used in the construction are also low-outgassing, high-reliability materials. This combination allows an extremely versatile and manufacturable high-volume, low-cost approach and eliminates cost and manufacturability concerns of existing multichip approaches. The fixture and chip placement is

designed to allow die-to-die alignment for subsequent photolithography steps. Each submodule in one layer of this technology can be considered a low-cost multichip module in a chip-scale package by itself. Due to the excellent surface planarity, multiple metal layers can be formed between the components allowing dense interconnects. Since metalization steps are performed in a wafer-like structure in batch mode, the cost added per component is minimal. The concept works well with any combination of components. The resulting subassemblies can be stacked if needed to increase the level of integration. The size of the subassemblies can be standardized (e.g., small, large, etc.) allowing a major manufacturing advantage by eliminating new tools for each new stack configuration, thereby eliminating a major cost concern in stacking resulting from high nonrecurring charges. Known-good dies can be used for molding. The use of known-good dies further increases the manufacturability and the yield, thus reducing the cost. Layers fabricated using the above approach can then be laminated into a single stack if needed. A ceramic substrate connected through to the stack can be placed on the top of the stack to provide a conventional interface to the next packaging level. Metalization is added to the two sides of the stack to complete the interconnection between dies, bringing all input/output signals to the cap chip.

A powerful feature of this approach is the capability of routing interconnects through the molding compound to realize area interconnects between the layers. Therefore, interconnection densities can be increased by orders of magnitude.

The thermal dissipation properties of the molded structure are not worse than a comparable semiconductor-based layer or stack. Note that the molding material is only on the periphery of the active components. Any potential degradation results from the epoxy surrounding the semiconductor chips eliminating dissipation from the edges. However, since during lapping the bottom of the semiconductor is exposed, there exists a direct thermal dissipation path (same as the original chip) to the next level of packaging that may consist of a heat sink. The expected thermal resistance will therefore be limited only by the thermal conductivity of the die-attach material used and can be as low as 0.02–0.2 W/mK. Therefore, the power dissipation of a single molded layer can be high, up to several tens of watts. In stacking applications, the thermal path may pass through layers of molding epoxy due to layout choices. In most cases, average thermal resistance of the assembly will be in the range of 10–20 C/W therefore limiting the total power dissipation to a few watts which in turn limits the number of layers that can be stacked. For cases where lower thermal resistance is needed for higher-power-dissipation in the stacks, blank silicon pieces can be added in the layers, under the high power dissipating components, to improve heat conduction. The technology is also versatile enough to accept an active-cooling (i.e., thermoelectric) device needed if the temperature of one of the components (such as a laser) requires stabilization.

The molded-integration approach is novel in many aspects. There are only few approaches that have some similarities to this molding technique in the literature. R. Johnson of Auburn University described a multichip approach in which photolithographic cavities are etched first in blank silicon wafers using the anisotropic wet-etching properties of silicon [29]. Electronic chips are placed in the cavities and using a faceplate to planarize, an adhesive is used to fill the cavities. The front face of the silicon-in-silicon structure is later processed for metalization and interconnects. This approach has the advantage of good thermal matching since the majority of the structure is silicon. However, silicon being a semiconductor also creates electrical isolation issues. Furthermore, the processing of the front side, including access to pads, is not trivial. Another approach was the high-density interconnect (HDI) technology developed by GE [30]. In this chip-first approach, laser-defined cavities are etched in ceramic substrates and chips are placed in the cavities. The metalization and dielectric layers are added using a special laser-based stitching operation to compensate for the chip-to-chip alignment errors. HDI multichip modules were inserted into some aerospace and military applications. Other variations of GE's HDI included the use of flexible nonceramic substrates. The drawback of the GE's approach is the customization needed for errors at each module level.

The first demonstration vehicle using molded-integration molding technology was an imaging sensor that can store up to 32 image frames into nonvolatile memory directly on the sensor, and is also capable of TV-rate real-time video imaging. The fabricated short stacks each have four layers of chips, including FPGA, flash memory, analog digital converter and an I/O chip (cap chip). Another design is a memory stack consisting of 16 layers of 64-Mb flash dies for a total capacity of 1024 Mb (1 Gb or 128 MB) and a layer of built-in line drivers (Fig. 6.4). The use of line drivers enables the realization of a stand-alone nonmoving part hard-disk replacement in a small form factor for use in portable-processing applications. The same technology allows the realization of very-small-form-factor portable computers with built-in data storage and processing units, as shown in Fig. 6.5.

In summary, this approach allows the realization of "system-in-stack" in chip-scale packages without expensive technology development, using existing known-good dies of various sources while maintaining aggressive integration goals. It is possible to encapsulate multiple dies in one neowafer. This allows the addition of receiver and transmitter components along with off-the-shelf ICs, resulting in a optoelectronic neochip. In addition, encapsulants identified so far are transparent, making this technique suitable for optoelectronic implementations, allowing optical receiver and transmitter elements to communicate between the layers of a laminated stack or as a single layer in a simple package ready to interface with optical components.



Figure 6.4. A molded integration example: the stack contains 16 layers of 64-Mb flash memory (layer shown on the right) and one layer of line drivers (layer shown on the left). Uppermost layer of the stack is a glass-epoxy based interface chip to the next packaging layer. The overall dimensions are $20 \times 12 \times 6$ mm for a total capacity of 1Gb.

6.3 APPLICATION TO OPTOELECTRONIC INTEGRATION

The neolayer approach described allows also the realization of complex optoelectronic integrated circuits and multichip modules. Various components can be joined together, including vertical-cavity surface-emitting lasers, detectors (e.g., Si, III-V, II-VI), waveguides, diffractive optical elements, integrated circuits (Si, GaAs) and even micro-electro-mechanical components. The surfaces of molded modules are planar and suitable for photolithographic formation of high-density interconnect waveguides or DOEs. Component-to-component placement accuracy

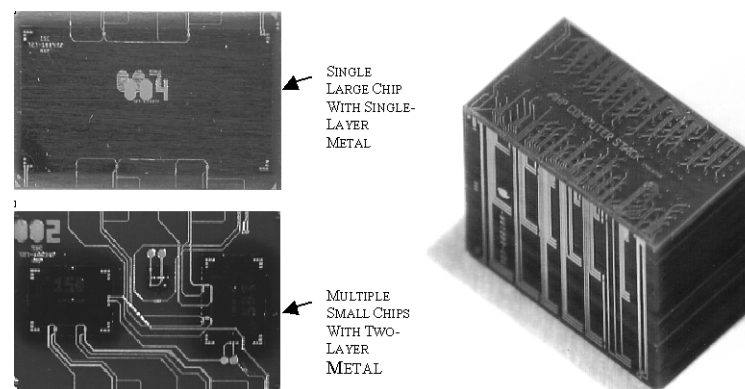


Figure 6.5. Typical complex neostack design consisting of 32 flash memory, 4 DRAM memory, 1 FPGA, 1 low-power CPU, 1 line driver, 1 I/O interface and 4 built-in decoupling capacitor layers with a total chip count of 50.

is $\pm 5 \mu\text{m}$. Since there are no moving or loose parts, misalignments do not present a problem. Molded optoelectronic integrated circuits can be fitted directly with molded optical components, thereby drastically reducing the assembly costs. Resulting integrated modules are fully compatible with existing electronic packaging technology with an inherent path for the implementation of power and ground lines needed to operate the system, a detail omitted in most of the existing optoelectronic integration techniques. Layers can be stacked together to form more complex assemblies as shown in Fig. 6.6.

6.3.1 MATERIAL SELECTION

One constraint for optoelectronic implementations is the transparency of the substrates of the optoelectronic receiver transmitter components (e.g., GaAs). If at least either the transmitter or the receiver substrate is not transparent at the selected wavelength, these components have to be placed in the stack format face-to-face to allow optical communications. On the other hand, even if their substrates are transparent, the electrical-connection requirement of the component and power supplies may require laying out metalization layers on both sides of the assembly. These constraints are illustrated in Fig. 6.7.

Another constraint is the transparency of the embedding material. Several materials have been evaluated, selected based on their mechanical and optical properties and relatively low curing temperatures. In order to establish the optical transparency, potted and fully-cured cross sections of various candidate compounds have been prepared. The thickness of the epoxy was approximately $480 \mu\text{m}$. The cured epoxy layer was diced into samples of $1 \times 1 \text{ cm}$ each. The absorption spectra of the samples in the wavelength range of 350–1100 nm are

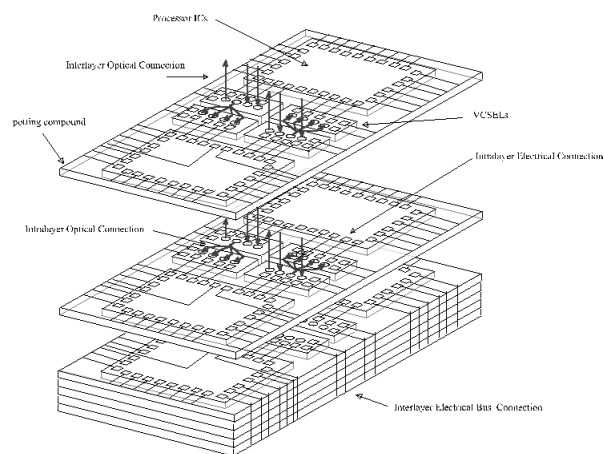


Figure 6.6. The concept of embedded optical links.

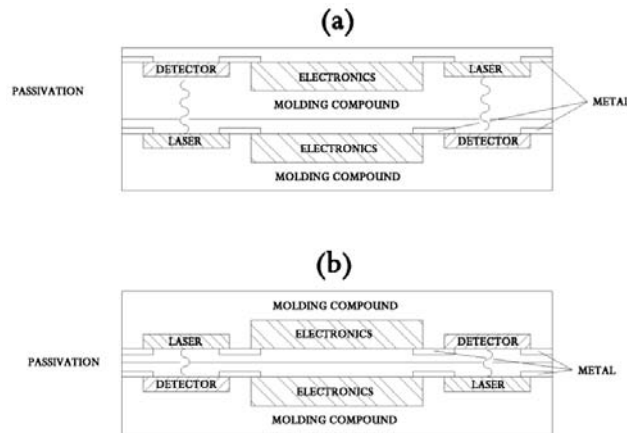


Figure 6.7. (a) The case of detector substrate being transparent at the sensitivity range. (b) The case of the detector substrate being opaque at the sensitivity range.

shown in Fig. 6.8. The initial operational wavelength range selection for demonstrations was in the 850-nm range. A magnified portion of the absorption spectra from 820–880 nm is shown in Fig. 6.9. Note that the transmission results included losses due to index mismatch (the refractive index is in the range of 1.6 for typical materials) and surface scattering. An epoxy-based material over 85% optical transmission has been selected for further experiments. The mechanical, optical, and electrical properties of a typical embedding material are listed in Table 6.1.

6.3.2 ALIGNMENT ISSUES

In this process, components are inserted into a frame using bumps and encapsulated. The manufacturing process has tolerances that will result in misalignments. During the original process development for solely electronic components, the misalignment requirements are based on the capturing bumped I/O pads. For optoelectronic application and especially for multilayers and stacks, transmitter position needs to be adjusted with respect to the adjacent layer's detector position. The final alignment tolerance required from the assembly process for stacked optoelectronic layer case is a strong function of the respective sizes of the beam waist at the detector interface and the detector active area. As shown in Fig. 6.10, the expected final placement tolerance can be calculated by geometric relationships in the stack. If the layers are not stacked, the alignment requirements are similar to a single multichip layer.

In Fig. 6.10, DV is the aperture of the transmitter (i.e., VCSEL), DD is the aperture of the detector, PD is the detector pitch (assumed to be equal to the VCSEL pitch), α is the divergence angle of the VCSEL and T is the distance

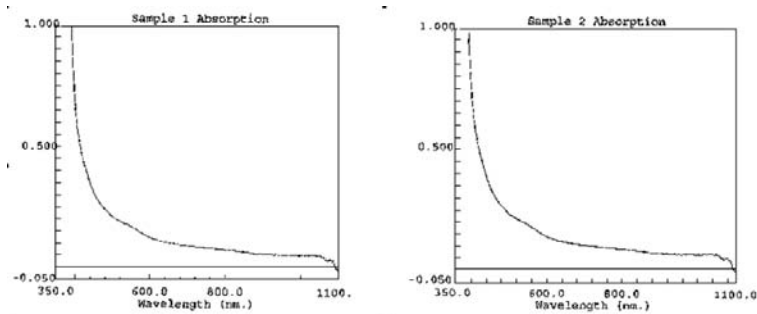


Figure 6.8. Absorption spectra of the molding compounds samples in the 350–1100-nm range.

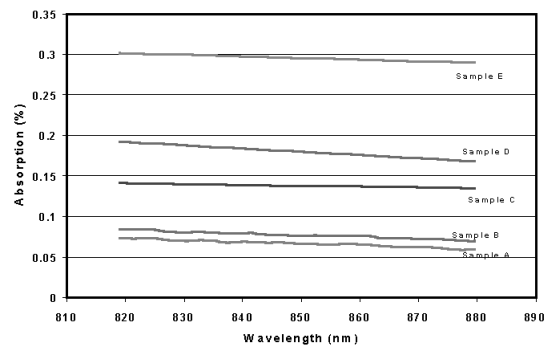


Figure 6.9. Absorption properties in the selected wavelength region of various molding compounds used in neochip formation.

Table 6.1.

Typical Properties of Molding Material

Cure schedule	1.5 hr at 60°C; 1 min at 150°C
Index of refraction	1.56 at 0.5 μ m
Spectral transmission	> .95% at 0.7–0.9 μ m
Flexural strength	10600 psi
Compressive strength	20200 psi
Coefficient of thermal expansion	12 ppm/C
Service temperature	–50 to 180°C
Tg	124°C
Moisture resistance (% weight increase)	0.03% at 96% RH, 1 wk
Dielectric strength	570 V/mil
Dielectric constant	5.1
Volume resistivity	1.0 E +16 ohm

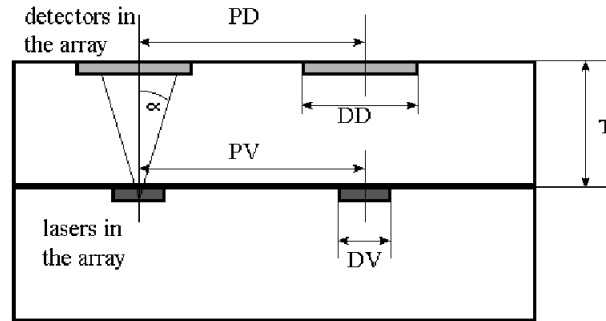


Figure 6.10. The relationships that determine the final assembly tolerances for the case of stacked optoelectronic layers.

between detector and VCSEL planes after assembly. Assuming that there is no misalignment, the projection of the divergent VCSEL beam onto the detector plane is:

$$B = DV + 2T \tan \alpha \quad (6.1)$$

The final alignment tolerances can be found assuming that 50% of the beam power should be falling onto the detector without illuminating the adjacent detector. Therefore, the final assembly tolerance will be defined by the smaller of the following quantities:

$$DD/2 - T \tan \alpha \text{ or } PD - DD/2 - DV/2 - T \tan \alpha \quad (6.2)$$

This relation indicates that for arrays where $PD \gg DD$ and $DD > DV$, the misalignment will be dictated by the detector aperture. Note that there is another trade-off on the detector side. From the assembly viewpoint it is advantageous to have larger detectors that will relax alignment tolerances. However, when the aperture of the detector is larger, the capacitance will increase and therefore the maximum speed of operation of the detector will decrease. As a numerical example, one can use a 15-micron VCSEL aperture with a 9° divergence angle, 50-micron detector aperture, 500-micron pitch, and 20-micron distance between layers that results in a final misalignment tolerance requirement of about ± 6 microns. This analysis is conducted for translation errors only. However, rotational errors can be transformed into translation errors by simple geometrical manipulations. Based on a typical square array size of 2 mm on the side, a numerical analysis using typical values used above results in a tolerable rotational error about 0.5° . There are several sources of error that may contribute to the final misalignment:

1. Photolithographic errors in the fabrication of the components and the holding frames: these errors are in the range of submicrons for modern devices and can be omitted.
2. Dicing errors in the fabrication of components: these errors are due to the tolerances of the dicing equipment and they are in the range of a few microns. However, they can be eliminated if some alignment feature on the die is used for alignment to the molding layer instead of the edge of the die as outlined previously in the approach based on the use of bump-and-plug technique. In this case, the photolithographically registered I/O pads are used. This is especially important if the supplier, without any precautions, dices the incoming wafer.
3. Placement errors during molding: these errors can be as large as several microns depending on the accuracy of approaches, tools, and fixtures used. The placement errors also include the location of bumps in corresponding pads. The use of the bump-and-plug technique, as outlined previously, minimizes the registration errors. The bump is placed automatically by the pick-and-place equipment. Therefore, the bump-to-pad registration and bump-to-plug tolerances will dictate the final alignment. For low cost and manufacturability, a fast-placement approach and relaxed tolerances are used which contribute to alignment errors that can be as large as several microns.
4. Dicing errors during dicing of the neodie layers: these errors are due to the inaccuracy of the dicing operation and can be as large as few microns. They can be minimized if the dicing references are selected based on the registration marks on the optoelectronic die and the edge-connect leads are large enough to absorb residual errors.
5. Lamination errors: these errors occur due to the inaccuracies of the lamination fixture used in the process of 3D stacking and can be on the order of several microns. If a single layer is used these errors are not applicable.

Therefore, the goal is to minimize errors in all other areas before lamination to minimize alignment errors between layers. Consequently the final alignment error will be dictated by the error of the lamination tool. The total alignment error resulting from these approaches for alignment will be discussed in experimental sections.

6.3.3 DEMONSTRATION OF MOLDED OPTOELECTRONIC LAYERS

The first demonstration of molded optoelectronic neolayers was performed using components that were available immediately. A VCSEL wafer segment, designed originally for individual lasers, was received from Honeywell for this

purpose and was diced to form 6×6 arrays as shown in Fig. 6.11. An IC designed and manufactured by ISC in the previous year was also selected to emulate the electronic component in the feasibility study. The selected IC contains several LED driver circuits and receiver amplifiers that are directly accessible. The IC chip operates on a single 5-V supply. The LED drivers have a measured output of 10 mA. The IC was designed to operate at 40 MHz, however both the receiver and transmitter sections functioned normally at 50 MHz. The LED drivers were designed to control the LED output by controlling the ground path of the LED supply voltage. The laser diodes, on the other hand, were of a common ground configuration as built by Honeywell. As a result, the demonstration circuit was fabricated by connecting the laser common to the output of the LED driver, thus limiting the number of utilized drivers to one. Power was individually applied to VCSELs in the matrix (a 6×6 array of VCSELs was deemed the smallest size appropriate for handling) in order not to exceed the capacity of the driver. Due to the common ground of the laser-diode array on the bottom of the die, an interplanar connection was also added. This connection is made through the thickness of the epoxy. The detector was eliminated from the final layout because no compatible detector array could be found. The layout is shown in Fig. 6.12. The 6×6 -VCSEL array was positioned adjacent to the long edge of the IC die. This configuration allows more room for interconnection metal, and a final neochip dimension slightly closer to square. The neochip output pads were arranged to match an existing probe card. Two edges of the neochip were utilized, with the IC signals and 22 VCSEL power lines on one edge, and 14 VCSEL power lines on the other. The interplanar connection is achieved with a through-hole via filled with a conductive polymer. The subsequent modification metal overcoats the top of the via. After modification, the back of the VCSEL array is cleared of any epoxy, and field metal is applied to the back of the neochip, connecting the array to the bottom of the via. The first completed optoelectronic neolayer is shown in Fig. 6.13. The lasers were all operational as shown in Fig. 6.14.

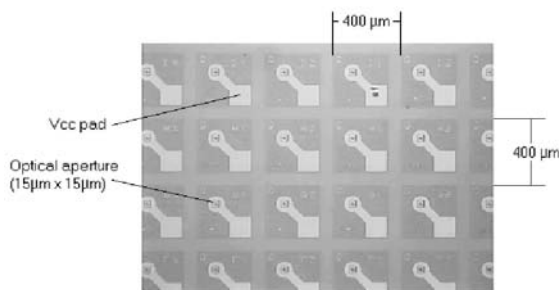


Figure 6.11. Photomicrograph and dimensioned layout of the Honeywell VCSEL diced into an array.

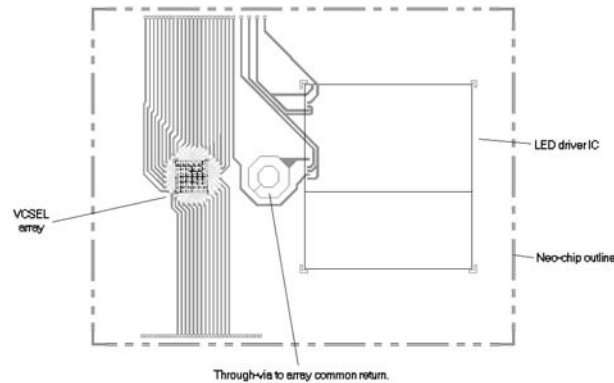


Figure 6.12. The layout of the first test circuit consisting of a 6×6 VCSEL array and the electronic driver/receiver IC.

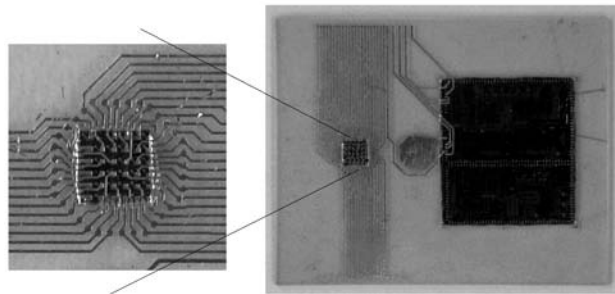


Figure 6.13. Photomicrograph of the completed optoelectronic neolayer with close-up showing VCSEL array area.

The second demonstration vehicle used to validate optoelectronic neolayer technology for optoelectronic applications was of an optical-link module illustrated in Fig. 6.15. Each layer has a transmitter array, its driver, a receiver array and its amplifier. In addition, an electronic chip has been selected to simulate electronic functionality. The following components were used in the prototype link demonstrator:

VCSEL Array

The laser array is the 1×12 VA-12 array at $250\text{-}\mu\text{m}$ pitch with an aperture of $15\text{-}\mu\text{m}$ diameter produced by Honeywell. These lasers have common cathode and enable speeds up to 1 GHz. The lasers have about 4–5-mA threshold currents and 1.7–1.8-V threshold voltages. They emit at 850 nm with a differential efficiency of about 50%. The chip size is $3600 \times 1800\text{ }\mu\text{m}$. A photomicrograph of a VCSEL array is shown in Fig. 6.16.

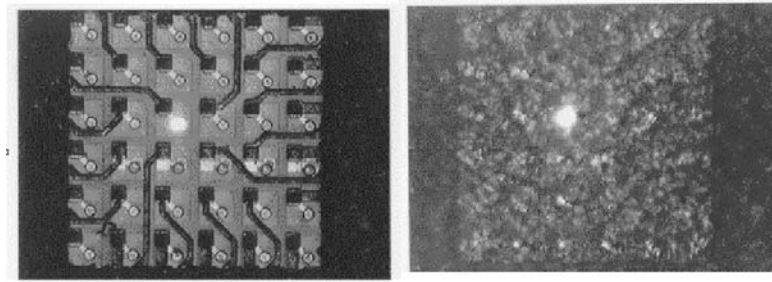


Figure 6.14. The operation of VCSELs in the fabricated optoelectronic neolayer. The photomicrograph on the right is the lasing through 500-micron additional epoxy covering the laser array demonstration the transparency of the material used.

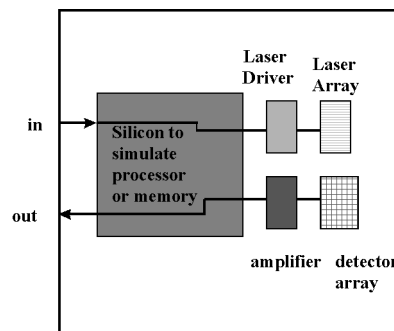


Figure 6.15. A typical layer used in the prototype link demonstration.

Detector Array

The detector array is the 1×22 DA-20 array at 140- μm pitch with an aperture of 80×80 - μm metal-semiconductor-metal (MSM) type produced by Honeywell. These detectors have common ground and enable speed up to 2 GHz. Since they are of MSM type, the capacitive loading is minimal and is not a function of the aperture. Therefore, larger apertures can be used for improved alignment tolerances without sacrificing speed. The responsivity of these detectors is about 0.25–0.27 A/W at 135 μW input power at 850 nm. The chip size is 3600×1800 μm . A photomicrograph of a detector array is shown in Fig. 6.17.

Laser Driver Array

The laser driver is the 1×12 channel HXT2000 driver-array chip from Helix. Although HXT2000 is optimized for edge-emitting laser, it has more current sourcing capabilities compared to HXT2001 designed for VCSELs. The drive cur-

rent can be modulated up to 8 mA from a 3.3-V power supply. Therefore, it is more suitable for large-threshold-current lasers such as Honeywell's. HXT2000 can support up to 1.25 Gb/s data rates. The chip size is $2000 \times 4100 \mu\text{m}$. A photomicrograph of a driver array is shown in Fig. 6.18.

Receiver/Preamplifier Array

The receiver/preamplifier is the 1×12 -channel ARX12 receiver array from Helix. ARX12 can support up to 2.5-Gb/s data rates therefore providing enough tolerance on the operation speed. It operates from a 3.3-V supply and has a versatile differential output stage. The chip size is $3800 \times 3800 \mu\text{m}$. Ten receiver-preamplifier arrays were received from Helix. A photomicrograph of a driver array is shown in Fig. 6.19.

Electronic Processing Chip

The CLDA2 laser-diode-array driver from OIDA-UNC-Charlotte will be used as an alternative to a dummy electronic chip to simulate the future processor chip. The built-in self-test and clock-generator circuit capabilities of the CLDA2 can be useful in performance measurements. This self-test mode allows 500-MHz operation and provides pseudorandom number generation. Furthermore, CLDA2 includes

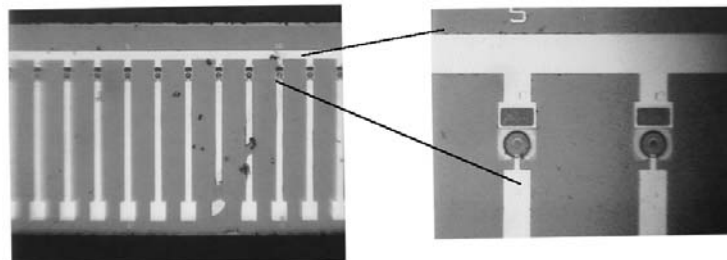


Figure 6.16. The 1×16 laser array and the detail of the VCSELs.

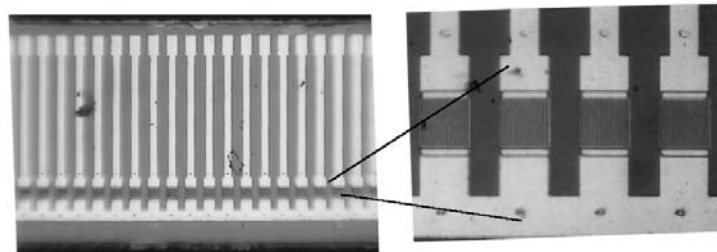


Figure 6.17. The 1×16 detector array and the detail of the MSM detector.

an additional driver that allows independent testing of lasers. CLDA2 has the capability of driving 32 laser channels at 150 MHz and its layout is based on four identical quadrants. The driver has the capability of 20 mA of adjustable bias and up to 40 mA of adjustable modulation at 3.3–5-V supplies. One quadrant (eight drivers) were used in these experiments. The CLDA2 chip is $4400 \times 4400 \mu\text{m}$. A photomicrograph of the CLDA2 chip is shown in Fig. 6.20.

The design of the optical link experiment was based on two layers with mirror images providing optical communication in the stack as shown in Fig. 6.21. Each layer contained a laser array, a driver array, a detector array and a receiver array as shown in Fig. 6.22.

Emulators for technology characterization

Due to the limited availability, only six actual parts were received for each component. Mechanical replicas were designed for process troubleshooting and characterization of mechanical and reliability properties. The neolayering process was applied to these emulator components. The result is presented in Fig. 6.23.

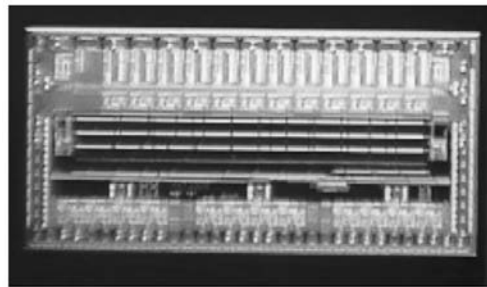


Figure 6.18. The laser driver chip.

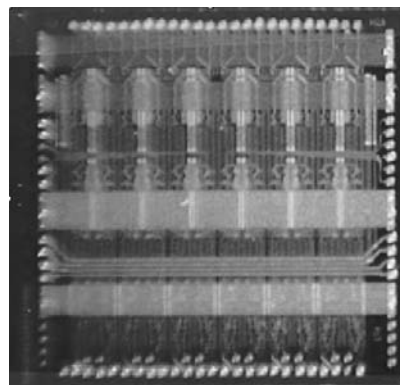


Figure 6.19. The receiver chip.

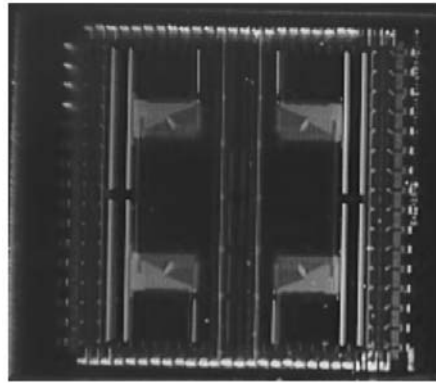


Figure 6.20. The CLDA laser driver signal generator chip.

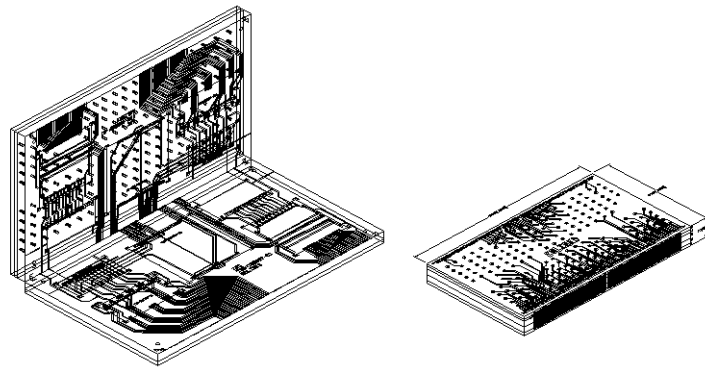


Figure 6.21. The optical-link design for the second demonstration.

Component-to-component misalignment, crucial for the optical interconnects, was less than $\pm 7 \mu\text{m}$, well within the expected range.

After the successful results of the emulator experiments, actual components were assembled in the molded optoelectronic neolayer as shown in Fig. 6.24. A total of four layers were completed due to the limited number of components in hand. The embedded detectors could not be used due to the metal-peeling problem in the detectors. A fast external detector from Thorlabs (SV2-FC, rated at 2 GHz) was used for optical detection. The lasers driven through the Helix driver chip could be operated up to 1 GHz as shown in Fig. 6.25. An external signal generator provided the driver input. No electrical cross talk was observed between the channels. The lasers (#1, #5, and #10) connected to CLDA2 were also found to be functional when driven from the CLDA2 chip as shown in Fig. 6.25. The pseudo-random bit generator of the CLDA2 chip was used to drive data through the lasers.

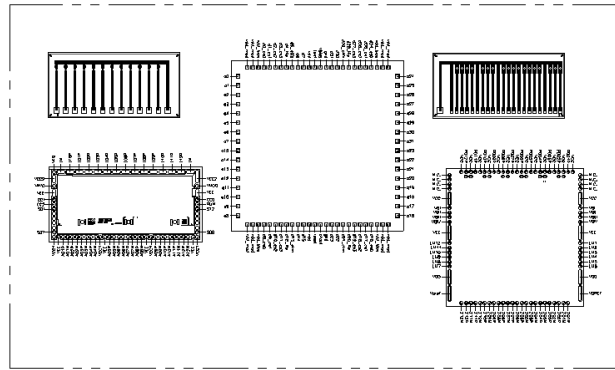


Figure 6.22. The details of the layout for the second demonstration of the optoelectronic neo-layer.

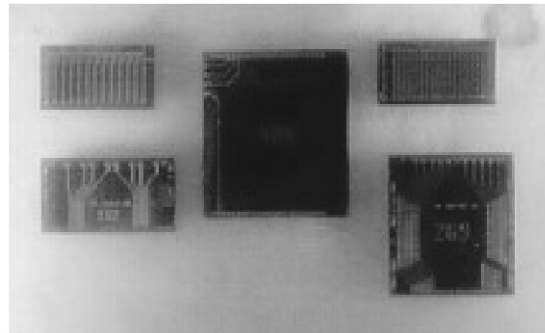


Figure 6.23. Mechanical emulator for optoelectronic integrated neo-multi-chip layers embedded in epoxy used for alignment studies. The bottom surface is visible through the transparency of the embedding epoxy. The upper-left chip is a laser replica and upper-right chip is a detector replica. The center chip represents the electronic processor. The lower-left chip represents the laser driver and the lower-right chip represents the receiver.

The speed of the data was as high as 500 Mb/s, limited only by the capability of the CLDA2 chip. An attempt was made to wire bond the detectors with defective metalization with the hope that they may survive a delicate wedge bonding cycle. However, this attempt was also unsuccessful. With the existing component selection, this module should allow 12 optical transmission channels at 25 GHz of usable bandwidth. This is the first time a molded over multichip module designed without any specific high-frequency precautions has ever operated in the 500 MHz–1-GHz regime. This demonstrates the capability of the technology developed.

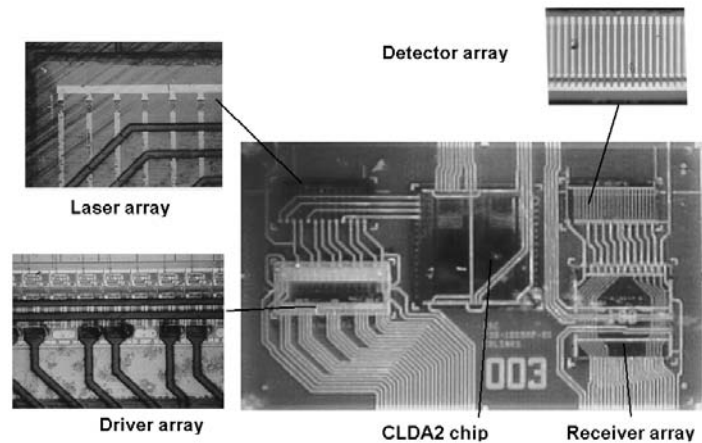


Figure 6.24. Second optoelectronic multichip in neolayer. A VCSEL array and its driver are on the left and a detector array and its receiver are on the right. The CLDA2 electronic test chip with self-testing capabilities is at the center.

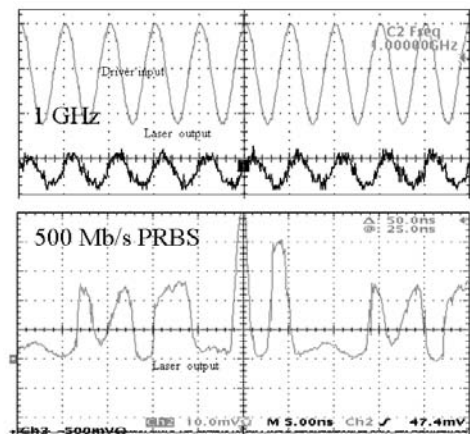


Figure 6.25. The operation of second optoelectronic neolayer multichip. Upper: operation of a laser driven by the Helix driver and external signal generator at 1 GHz. Lower: operation of a laser driven by the CLDA2 chip with 500-Mb/s pseudorandom bit sequence.

6.4 FUTURE DEVELOPMENT, SYSTEM APPLICATIONS AND CONCLUSIONS

The molding-based optoelectronic integration technology described here has been demonstrated in its basic form. For optoelectronic applications, optical and mechanical interfaces are as important as the optoelectronic assembly. One way to introduce more complex optical functions would be the use of waveguiding

effects. The materials used for the implementation of molded optoelectronic layers have an index of refraction in the range of 1.5–16. Therefore, optical components such as diffractive and refractive lenses can be added on the surface of the layer. In this scheme, waveguiding and beam-splitting effects can be created using index-of-refraction differences of different epoxies and polymers offering several advantages:

- optical interconnects can be realized between arbitrary layers of the stack
- layout of the layers becomes simple since the requirement of overlapping the detectors and transmitters is no longer present
- optical interconnects can be taken out from the side of the module either into free space or to fibers, eliminating the need of placing an additional array

The development of such powerful optical interconnect modules will require the development of 45° mirrors on the top of the transmitters and receivers, formation of the waveguides along with metal traces for optical interconnects and formation of a 3D waveguide structure. Some of these developments have already taken place, for example within the POINT program by AMP and Allied Chemical. The basic waveguide fabrication has been also demonstrated as shown in Fig. 6.26. More technology development is underway to produce an optoelectronic assembly that can be inserted easily into many application platforms.

Military applications using such modules include onboard processing of radar signals, airborne reconnaissance, target recognition, naval/airborne/land multiple-sensor fusion and associated real-time data processing. Commercial applications

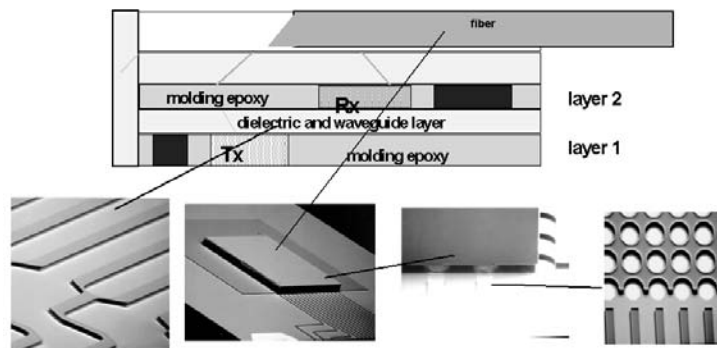


Figure 6.26. Conceptual rendering for various techniques that can be added to optoelectronic neochips for communicating optically out-of-the-stack.

include high-capacity public switches, real-time data and image compression (or decompression) such as in medical images, and high-speed interface to large databases in optical memories. Typical examples in the telecommunication area are ATM switch modules or gigabit ethernet interface circuits. Examples in the processing area are optically interconnected image-recognition modules or LIDAR interface modules. One important commercial application will be integrated parallel read/write heads for optical memories.

The real revolution, however, derives from the fact that the neolayer technology eliminates the need for printed circuit boards and their interconnection structure, overcoming one of the major obstacles for ultra-high-speed systems.

ACKNOWLEDGMENTS

The optoelectronic-molding technology development and module demonstration was supported by the U.S. Army, Aviation and Missile Command (monitored by Pete Black) under an SBIR contract. The authors would like to thank Andrew Camien, Jim Yamaguchi and Angel Pepe of ISC for their contributions to the development of the neochip technology. Honeywell Technology Center (Drs. Mary Hibbs-Brenner and Yue Liu) provided the lasers and detector arrays used for the demonstrations. The OIDA-JOP program provided the CLDA2 chip designed by Dr. Kiamilev. Their support was instrumental in the realization of the demonstration vehicles.

REFERENCES

- [1] T. S. McLaren, S. Y. Kang, W. G. Zhang, T. H. Ju, Y. C. Lee, "Thermosonic bonding of an optoelectronic transceiver based on an 8×8 vertical cavity surface emitting laser array," *IEEE Trans. Components, Packaging and Manufacturing Technology*, Vol. CHMT-20, No. 2, pp. 152–160 (1997).
- [1] W. Imler, K. Sholz, Maria Cobarruviaz, V. K. Nagesh, Clinton Chao, R. Haitz, "Precision Flip-chip Solder Bump Interconnects for optical packaging," *IEEE Trans. Components, Packaging and Manufacturing Technology*, Vol. CHMT-15, No. 6, pp. 977–981 (1992).
- [2] C. Armiento, A. Negru, M. Tabasky, R. Boudreau, M. Rothman, P. Haugsjaa, "Gigabit transmitter array modules on silicon waferboard," *IEEE Trans. Components, Packaging and Manufacturing Technology*, Vol. CHMT-15, No. 6, pp. 1072–1080 (1992).
- [3] M. Goodwin, A. Moseley, M. Kearly, R. Morris, R. Goodfellow, "Optoelectronic component arrays for optical interconnection of circuits and systems," *J. Lightwave Technol.*, Vol. 9, No. 12, pp. 1639–1645 (1991).
- [4] T. Hayashi, "An innovative bonding technique for optical chips using solder bumps that eliminates chip positioning adjustments," *IEEE Trans. Com-*

- ponents, *Packaging and Manufacturing Technology*, Vol. CHMT-15, No. 2, pp. 225–230 (1992).
- [5] K. Kjackson, E. Flint, M Fina, D. Lacey, Y. Kwark, J. Trehwella, T. Caulfield, P. Buchmann, C. Harder, P. Vettiger, “A high density, four channel, OEIC tranceiver module utilizing planar processed optical waveguides and flip-chip solder bump technology,” *J. Lightwave Technol.* Vol. 12, No. 7, pp. 1185–1191 (1994).
 - [6] J. Jahns, R. Morgan, H. Ngyuen, J. Walker, S. Walker, Y. Wong, “Hybrid integration of surface emitting microlaser chips and planar optics substrate for interconnection applications,” *IEEE, Photonics Technol. Lett.* Vol. 4, No. 12, pp. 1369–1372 (1992).
 - [7] T. Worchevsky, K. Ritter, R. Martin, B. Lane, “Large arrays of spatial light modulators hybridized to silicon integrated circuits,” *Appl. Optics*, Vol. 35, No. 3, pp. 1180–1186 (1996).
 - [8] Y. Ohiso, K. Tateno, Y. Kohama, A. Wakatsuki, H. Tsunetsugu, T. Kurokawa, “Flip-chip bonded 0.85 μm bottom-emitting vertical cavity laser array on an AlGaAs substrate,” *IEEE, Photonics Technol. Lett.* Vol. 8, No. 9, pp. 1115–1117 (1992).
 - [9] T. Woodward, A. Krishnamoorthy, A. Lentine, K. Goossen, J. Walker, J. Cunningham, W. Yan, L. Asaro, L. Chirowsky, “1 Gb/s two-beam transimpedance smart-pixel optical receivers made from hybrid GaAs MQW modulators bonded to 0.8 μm silicon CMOS,” *IEEE, Photonics Technol. Lett.* Vol. 8, No. 3, pp. 422–424 (1996).
 - [10] E. Yablonovitch, T. Gmitter, P. Harbison, R. Bhat, “Extreme selectivity in the lift-off of Epitaxial GaAs films,” *Appl. Phys. Lett.* Vol. 51, No. 26, pp. 2222–2224 (1987).
 - [11] A. Ersen, I. Schnitzer, E. Yablonovitch, T. Gmitter, “Direct bonding of GaAs films on silicon circuits by epitaxial lift-off,” *Solid State Electronics*, Vol. 36, No. 12, pp. 1731–1739 (1993).
 - [12] E. Yablonovitch, T. Sands, D. Wang, I. Schnitzer, T. Gmitter, S. Shastry, D. Hill J. Fan, “VanDerWalls bonding of GaAs on Pd leads to a permanent solid-phase topotaxial metalurgical bond,” *Appl. Phys. Lett.*, Vol. 59, No. 24, pp. 3519–3161 (1991).
 - [13] D. Mathine, R. Droopad, G. Maracas, “A vertical surface emitting laser applique to a 0.8 μm NMOS driver,” *IEEE Photonics Technol. Lett.* Vol. 9, No. 5, pp. 869–871 (1997).
 - [14] H. Fathollahnejad, D. Mathine, R. Droopad, G. Maracs, S. Daryanani, “Vertical cavity surface emitting lasers integrated onto silicon substrates by PdGe contacts,” *Electron Lett.*, Vol. 30, No. 7, pp 1235–1236 (1994).
 - [15] D. Pllentier, Y. Zhu, D. DeMaulmester, P. Van Daele, P. Demeester, “Low stress Pd/Ge based ohmic contact for GaAs epitaxial lift-off,” *Microelectron.*

- Eng.* Vol. 15, No. 1, pp. 153–156 (1991).
- [16] M. Jin, J. Wang, V. Ozguz, S. H. Lee, “Monolithic Integration of Si Driver Circuits onto PLZT Substrates for Smart Spatial Light Modulator Applications,” *Appl. Optics*, Vol. 33, No. 14, pp. 2842–2849 (1994).
- [17] M. Jin, V. Ozguz, J. H. Wang, S. Lee, “Direct bonding and Flip-chip Bonding Technologies Applied to Spatial Light Modulator Fabrication,” *Proc. 45th ECTC*, pp. 194–201 (1995).
- [18] S. Matsuo, T. Nakahara, K. Tateno, T. Kurokawa, “Novel technology for hybrid integration of photonic and electronic circuits,” *IEEE Photonics Technol. Lett*, Vol. 8, No. 11, pp. 1507–1509 (1997).
- [19] C. Camperi-Ginstet, Y. Kim, N. Jokerst, M. Allen, M. Brooke, “Vertical electrical interconnection of compound semiconductor thin-film devices to underlying silicon circuits,” *IEEE Photonics Technol. Lett*, Vol. 4, No. 9, pp. 1003–1006 (1997).
- [20] F. Ejeckam, Z. Zhu, Y. Quian, J. Zhang, Z. Zhang, G. Chistenson, Y. Lo, “Wafer bonding technology and its applications in optoelectronics devices and materials,” *IEEE J. Sel Topics in Quantum El.*, Vol. 3, No. 3, pp. 927–939 (1997).
- [21] A. Black, A. Hawkins, N. Margalit, D. Babic, A. Holmes, Y. Chang, P. Abraham, J. Bowers, E. Lu, “Wafer fusion: Materials issues and Device results,” *IEEE J. Sel Topics in Quantum El.*, Vol. 3, No. 3, pp. 943–951 (1997).
- [22] B. Levine A. Jwkins, S. Hiu, B. Tseng, C. King, R. Johnson, D. Zolnovski, J. Bowers, “Characterization of wafer bonded photodetectors fabricated using various annealing temperatures and ambients,” *Appl. Phys. Lett.*, Vol. 71, No. 11, pp. 1507–1509 (1997)
- [23] K. Mori, K. Tokutome, S. Suogu, “Low threshold pulsed operation of long wavelength lasers on Si fabricated by direct bonding,” *Electron. Lett.*, Vol. 31, No 4. pp. 284–185 (1995).
- [24] Y. H. Lo, R. Bhat, D. Hwang, C. Chua, C. Lin, “Semiconductor lasers on Su substrates using the technology of bonding by atomic rearrangement,” *Appl. Phys. Lett.*, Vol. 62, No. 10, pp. 1038–1040 (1993).
- [25] R. Ram, J. Dudley, J. Bowers, L. Yang, K. Carey, S. Rosner, K. Nakua, “GaAs to InP wafer fusion,” *J. Appl. Phys.* Vol. 78, No. 6, pp. 4227–4237 (1995).
- [26] R. Lea, I. Jalowiecki, D. Boughton, J. Yamaguchi, A. Pepe, V. Ozguz, J. Carson, “A 3-D stacked chip packaging solution for miniaturized massively parallel processing,” *IEEE Tran. Adv. Packaging*, Vol. 22, No. 3, pp. 424–432 (1999).
- [27] U. S. Patent No. 5,953,588, A. Camien, J. Yamaguchi, Stackable layers containing encapsulated IC chips (1999).

- [28] K. Gann, "Neo-wafer 3D Packaging," *Proc. IMAPS 3D ATW '98*, pp. 35–39 (1998).
- [29] R. W. Johnson, T. Philips, R. Jaeger, S. Hahn and D. Burdeaux, "Multi-chip thin film technology in silicon," *Trans. IEEE CHMT*, Vol. 12, No. 2, pp. 185–194 (1989).
- [30] W. Dann, W. Burdick and R. Fillion, "A chip-first multi-chip module technology," *IEE Computer*, Vol. 23, No. 4, pp. 23–29 (1993).

Scalable Micro-Optical Modules for Short-Distance VLSI-Photonic Interconnections

Hugo Thienpont, Valerie Baukens, Christof Debaes, Heidi Ottevaere,
Pedro Vynck, Patrik Tuteleers, Guy Verschaffelt, Bart Volckaerts,
Alex Hermanne, Irina Veretennicoff
VUB, Vrije Universiteit Brussel

Marnik Brunfaut, Jan Van Campenhout
Department of Electronics and Information Systems, University of Gent

Mike Hanney
CTP-Combined Optical Industries Limited

7.1 INTRODUCTION AND RATIONALE

7.1.1 THE SIA ROADMAP

Fiber optics and free-space optoelectronic technologies have been widely investigated as means to alleviate data communication bottlenecks at the machine-to-machine, back-plane and board-to-board interconnection level [1–3]. Recent breakthroughs in the fabrication of spatial arrays of optoelectronic emitters and detectors [4, 5] and their heterogeneous integration with Si-CMOS electronic chips [6–9] now also encourage the use of photonics as a wire replacement technology at the inter- and intra-multichip-module (MCM) interconnection level [10]. With this approach one aims to increase the communication functions between the latter electronic processing modules, relaxing their bandwidth limitations. These limitations are primarily imposed by fundamental electrical signal propagation issues and the limited number of electrical chip pin-outs [11, 12].

The requirements on these short-range interconnect fabrics in high-performance systems are spelled out in the SIA Semiconductor road map [11]. The main thrust in this evolution is that, over the next seven years, high-performance ASIC or processor chip sizes are expected to grow to 520 mm² (implying lateral dimensions of 22.8 × 22.8 mm), and that they will need up to 4000 connections to their environment (chip package or high-density substrate), of which 3000 will carry high-speed signals (see Table 7.1).

Locally on the chip, frequencies will grow to as high as 3.5 GHz, and across-chip frequencies are expected to hit values ranging from 0.9 to 2 GHz. Chip-to-board frequencies will reach roughly half that frequency, driving buses with widths of 256 bits and more.

In order to interconnect chips to their interconnection substrates, area flip-chip bonding techniques will have to be developed that will reach down to a 100- μm pitch on the chip. The substrates themselves will require track widths and spacings as low as 20 μm in order to access two rows of bumps per wiring layer.

Reaching these goals will require significant research and development, as existing galvanic interconnection technology is currently not capable of providing the required performance. The road map document indeed marks several key requirements as “Research Required” or “Development Underway,” “Solutions Being Pursued,” or even as having “No Known Solution.”

7.1.2 SHORT-RANGE HIGH-DENSITY OPTICAL INTERCONNECTS: A POSSIBLE SOLUTION

While some of the problems identified relate to the fabrication of the VLSI chips themselves, a significant part has to do with providing the external chip interconnect hierarchy that will offer the required combination of interconnect density, escape boundary, average interconnect length, aggregate bandwidth, signal integrity, manufacturability and, of course, cost. Significantly better inter-

<i>Year of First Shipment</i>	1999	2001	2003	2006
<i>Sizes and dimensions</i>				
Chip size (mm^2)	340	385	430	520
Flip-chip pitch (mm)	180	150	130	100
PWB line width (mm)	34	30	26	20
PWB line spacing (mm)	38	30	26	20
<i>Frequencies (MHz)</i>				
Local on-chip	1250	1500	2100	3500
Across-chip	958	1570	1768	2075
Off-chip	479	785	884	1037
<i>Counts and densities</i>				
# pins (chip)	1991	2432	2970	4010
# pins (package)	1493	1824	2228	3008
I/O bus width	128	256	256	256
Substrate wire density (cm/cm^2)	278	334	385	500

Table 7.1. The projected evolution of some relevant chip and chip-to-board parameters, according to the 1997 SIA Road Map [11].

connect substrates are needed than what is available today, and short-range high-density optical interconnects are believed to provide an effective solution.

The trade-offs involved have been discussed by several authors. George et al. [12] investigate the possibility of implementing complex systems in the form of multiple, relatively small VLSI chips interconnected using flip-chip technology and high-density substrates as opposed to one, much larger, silicon chip. The trade-off involves comparing the larger on-chip interconnect delays and lower process yield of the one-chip solution with the need for extra on-chip drive capability to bring the signals off-chip, and the required interconnect capabilities of the substrate. One of their conclusions is that large and complex designs (high off-chip Rent exponents) impose very high demands on the interconnect substrate for a multichip solution to be profitable. They find that even the current advanced MCM technology is not capable of offering the required wiring capacity to tile complex designs. On the other hand, if appropriate substrates would be available, the gains could be significant, as borne out by their simulations.

One of the underlying mechanisms is that the relative strength of MCM or PCB based interconnect—the fact that traces can be considered almost lossless transmission lines with propagation delays proportional to their length—is gradually lost in the modern, extreme fine-line MCMs. Even at interconnection lengths of only a few cm, the 20- μm -wide traces of, e.g., an MCM-D substrate start behaving like lossy transmission lines, with a faster-than-linear increase of signal travel time and a similar reduction of operating frequency. It is this effect that was identified by Miller and Ozaktas as the main limitation of galvanic interconnects [13, 14].

Haney and Christensen [15] perform a similar analysis, and make a comparison between high-density but short-range interchip interconnects using galvanic means and optical means; in particular they consider various forms of free-space optical interconnect as alternatives for the traditional galvanic interconnect. They conclude that the interconnect density offered by free-space optical interconnects, and the absence of an interconnect hierarchy with successively lower densities, as is the case in galvanic interconnects, are the key strengths of optical approaches. These properties lead to a far better scaling performance than exhibited by galvanic interconnects, in particular in systems that have extremely large global interconnect requirements (which is e.g. the case in systems with high off-chip Rent exponents, a similar conclusion as the one by George et al.). Their comparison also indicates that optical interconnects hold a far greater promise than galvanic ones in terms of power efficiency. Just switching the 5-pF capacitance of a 4-cm MCM-D interconnect to 3.3 V and back dissipates more than 50 pJ, not counting the dissipation in the multistage pad driver. Optical links with a much better total speed-power product are currently available [16] (better than 10 pJ), even when using active light sources such as VCSELs or even MCLEDs.

Among the optical alternatives, Haney and Christensen [15] consider the macro-optical approach, using a small number of “large” lenses, as well as the micro-optical one, using an individual lens per optical channel. They argue that the macro-optical approach is the superior one if the interconnect pattern is indeed global, i.e., if a large number of interconnections run between any pair of chips in the system. At this point, however, we could raise a question concerning the expected nature of the system-level interconnect pattern of future high-performance systems. As argued before [10], it is believed that systems that intend to take full advantage of the increased interchip interconnect density will indeed exhibit other Rent properties and wire length distributions than the current ones, which often have an artificially low incremental Rent exponent in the higher levels of their interconnect hierarchy. Increasing the high-level (off-chip) Rent exponent can be accomplished by, e.g., gradually replacing buses and other time-multiplexed interconnect structures by dedicated point-to-point interconnects. Such a transformation indeed leads to vastly increased pin numbers on individual chips, and allows for higher performance since the contention for the shared interconnect structure has been removed.

However, it is expected that systems that indeed require a fully global interconnect fabric will be the exception rather than the rule. Instead, a large fraction of the interconnections, although far more numerous than in current-day systems, will still be local. This feeling is supported by several emerging systems that aim at exploiting the interconnect density advantages of optical interconnect: high-bandwidth ring networks [17] or optical backplanes [18], optically interconnected 3D systems [19, 20] and many others. In systems with such a less global or even local interconnect pattern, the micro-optical approach gains in importance; in systems where most of the interconnect is local (only connections to the nearest neighbors), the main drawback of the micro-optical approach—its limited throw distance—vanishes all together. One can then fully benefit from the particular advantages of the micro-optical approach advocated in this chapter, namely its compact, solid nature, and its great potential for providing easy-to-produce low-cost building blocks of tomorrow’s VLSI interconnect substrates.

7.1.3 STRUCTURE OF THE CHAPTER

This chapter therefore reports on the design, fabrication and demonstration of such a prototype multichannel interconnection module. Fabricated with deep proton lithography [21] in polymethylmethacrylate (PMMA) the component can be mass-produced at very low cost by injection-molding replication techniques in highly advanced optical plastics [22] agreeable to the wishes of the semiconductor industry.

The chapter starts by introducing the concept of this free-space MCM interconnection component and the parameters involved in its design. It then describes the basic principles of deep proton irradiation of PMMA and shows how this deep

etch lithographic technique is applied to fabricate the micro-optical structure for short distance optical interconnects, integrating refractive microlenses, micromirrors and passive alignment features. Typical optical characteristics of the micro-optical structures and surfaces thus obtained are highlighted.

The next part of the chapter is devoted to the experimental set-up and the performances of this multichannel module. The last two sections are devoted to a sensitivity analysis for misalignments and fabrication tolerances and to fabrication issues related to the mass-replication of these components through injection-molding techniques with high-optical-quality plastics.

7.2 THE CONCEPT

7.2.1 FREE-SPACE AND GUIDED-WAVE SCHEMES

Research groups all over the world are presently exploring various optical schemes to interconnect densely-packed photonic pin-outs regularly distributed over entire CMOS chip areas. They try to accommodate the actual requirements imposed by the performance specifications and geometrical limitations of their feasibility demonstrators. Key considerations here are the relative positions of the CMOS chips to be interconnected, the available overhead space, the distances to be bridged, the interconnection configuration, the photonic pin-out densities, the overall power budget, the high signal integrity to be achieved at high aggregate bit rates, and the power dissipation and packaging issues. Different schemes are possible to realize these interconnection systems in practice. Most of these implementations are based on a beam-guiding approach. They either combine or embed rib wave guides, individual fibers or imaging fiber bundles into flexible [23, 24] or rigid [25] modules. A second route is the use of free-space beam-shaping and beam-delivering micro-optical structures combined in compact hybrid or monolithic modules [26, 27].

7.2.2 THE FREE-SPACE APPROACH

The optical pathway block described in this chapter is an optical bridge consisting of two components. The first one is a base plate featuring two arrays of spherical microlenses and a pair of alignment holes. The second component integrates the counterpart mounting features and two micromirrors. Figure 7.1 shows the individual components as well as the assembled optical bridge aligned above a VCSEL and a detector array heterogeneously integrated with a CMOS chip.

The concept is based on the idea that the microlens array positioned above the emitters would act as a beam collimator, while the second microlens array should refocus the beams onto the detectors. Each micromirror surface should make an angle of 45° with respect to the optical axes of the lenses in order to deflect the collimated beams by 90° . For this interconnection module only VCSELs will be

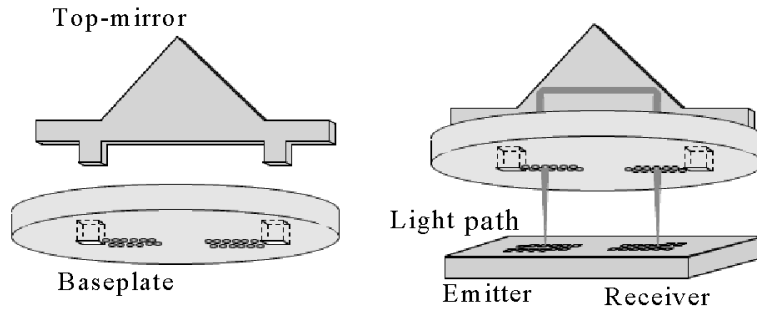


Figure 7.1. Concept of a 16-channel micro-optical bridge and its components.

considered as sources because for this free-space microlens relay concept, low-divergence-angle emitters are imperative to avoid cross talk between adjacent channels.

Contrarily to “individual fiber”-based schemes, where the diameters of the fiber claddings limit the maximum channel density of the optical interconnects, the free-space bridges proposed here have the advantage that there are no major technological fabrication limitations for small lenslet diameters at different focal lengths. One only has to take care that the diameters of the lenslets are smaller than the channel pitch. This provides the free-space approach with the potential advantage of being scalable, because lower lens diameters imply higher channel densities and consequently higher total aggregate bit rates.

7.2.3 THE GAUSSIAN BEAM MODEL

The minimum lens diameter needed to bridge a geometrical interconnection length L in this component is determined by the diffraction of the VCSEL beam, which is characterized by a Gaussian angular intensity distribution, a beam waist w_0 , a beam divergence θ and a wavelength λ_0 . From the minimum lens diameter one can calculate the maximum channel density as a function of the distance traveled in the optical pathway block (OPB), assuming that the pitch of the channels equals the lens diameter. Assuming that in the middle of the OPB (at $z = 0$) the beam waist is $w'(0) = w'_0$ and that the material index of refraction of PMMA is $n = 1.48$ @ $\lambda_0 = 850$ nm, then the beam radius at the position of the lens ($z = L_{\max}/2 = L$) is given by

$$w'(L) = w'_0 \sqrt{1 + \frac{\lambda_0^2 L^2}{n^2 \pi^2 w_0'^4}} .$$

A schematic drawing of a single channel of this confocal system is given in Fig. 7.2.

If one now applies the rule of thumb that the laser beam must always be smaller

than $2/3$ of the lens diameter so that more than 99% optical energy throughput through the lenses is achieved and that cross talk is absent in a perfectly fabricated and aligned system, then

$$w'(L) = \frac{2}{3} \cdot \frac{\phi_{\text{lens}}}{2}$$

and

$$L = \frac{n\pi}{\lambda_0} \sqrt{\frac{w_0'^2 \phi_{\text{lens}}^2}{9} - w_0'^4}.$$

Calculating the beam waist w_0 such that the interconnection distance L is maximum, it is found for

$$\frac{\partial L}{\partial w_0'} = 0$$

that

$$w_0' = \frac{\phi_{\text{lens}}}{3\sqrt{2}}$$

and that the maximum geometrical pathway length that can be traveled in the bridge is

$$L_{\text{max}} = 2L = \frac{n\pi}{\lambda_0} \frac{\phi_{\text{lens}}^2}{9}.$$

Due to the diffraction of the laser beam the lens diameter ϕ_{lens} for an interconnection length L therefore has to be equal to, but preferably larger than

$$\phi_{\text{lens}} = 3 \cdot \sqrt{\frac{\lambda_0}{n\pi} \cdot L}.$$

This requirement on the minimum size of the lens diameter will restrict the device density N to

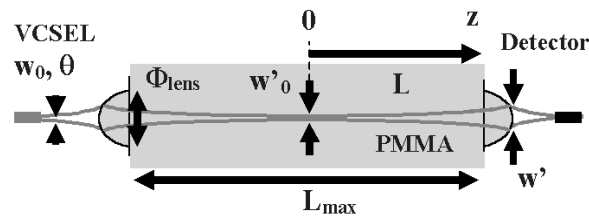


Figure 7.2. Propagation of the laser-beam profile in a confocal microlens relay system.

$$N = \frac{1}{(\phi_{\text{lens}})^2}$$

in case power dissipation considerations do not play an even more severe limiting role.

In this calculation it is assumed that for every laser diode with beam waist w_0 and divergence θ , there exists a lens with a focal length such that

$$w'_0 = \frac{\phi_{\text{lens}}}{3\sqrt{2}} .$$

This is true only if

$$f \geq \frac{n\pi}{\lambda_0} w_0 w'_0 = \frac{n\pi}{\lambda_0} w_0 \frac{\phi_{\text{lens}}}{3\sqrt{2}}$$

or

$$f/\# = \frac{f}{\phi_{\text{lens}}} \geq \frac{n\pi}{3\sqrt{2}} \cdot \frac{w_0}{\lambda_0} = \frac{1}{3\sqrt{2}} \cdot \frac{1}{\theta} .$$

7.2.4 CHANNEL DENSITY VERSUS INTERCONNECTION LENGTH TRADE-OFFS

For VCSELs with a FWHM angle of 12° or a beam divergence $\theta = 6^\circ$ and a wavelength $\lambda_0 = 850$ nm, this would mean that the focal number of the lenses will have to have a value $f/\# \geq 2.24$. With this simple analytical model the allowed channel density N as a function of the geometrical pathway length L_{max} can be calculated for different lens diameter sizes. The results are displayed in Fig. 7.3.

It can be concluded from this graph that one can trade off pathway length and therefore interconnection length for channel density. Indeed, one can increase the interconnection length by using larger diameter microlenses, but this comes at the cost of lower channel densities. For example a microlens with a diameter of $100 \mu\text{m}$ could yield a channel density of $10^4/\text{cm}^2$. However, this density only allows a pathway length L_{max} of 6 mm. With a microlens diameter of $200 \mu\text{m}$, on the other hand, one can bridge lengths up to 24 mm, but at the cost of a lower channel density of $2.5 \cdot 10^3/\text{cm}^2$. A very important characteristic of these optical data transmission modules is their total data throughput capacity. The data throughput per chip area of a data link is defined as the bit rate per channel multiplied by the channel density in the interconnection module. In general, the channel densities and the bit rates that can be achieved are primarily limited by the maximum allowed power dissipation of the optoelectronic emitters and receivers and their CMOS or bipolar drivers [28]. Usually power dissipation densities are limited to $10\text{W}/\text{cm}^2$ for air-cooled silicon or to $100\text{W}/\text{cm}^2$ when the chip is actively-cooled. As stated above,

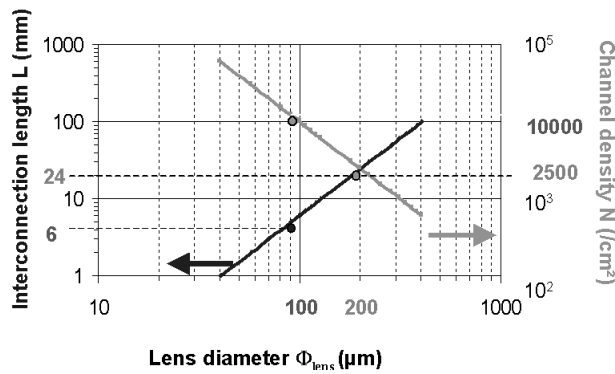


Figure 7.3. Channel density versus interconnection length.

an additional limiting factor in the particular case of free-space OPBs is the diffraction of the laser beams rather than their attenuation in the optical medium. This causes the channel density to drop with increasing interconnection length and limits the total throughput per chip area for an interconnection length beyond a couple of centimeters [30]. Notwithstanding these and other types of limiting factors, it is important to notice that for this type of free-space interconnection module all the different performance analyses converge to the same conclusions: interconnect densities of $10^3/\text{cm}^2$ at around 1 Gb/s bit rate can be achieved for passively cooled chips, which corresponds to an aggregate throughput capacity of 1Tb/s.cm^2 , while actively cooled systems could even perform aggregate bandwidths of several Tb/s.cm^2 over future inter- and intra-MCM interconnection distances [28–31]. Furthermore the performances of both plastic optical fiber wave-guide and free-space optical pathway blocks were compared [30]. As a general conclusion one can say that while both approaches have the potential for Tbits/s.cm^2 fire-hose data-handling capacity the rigid free-space module and the more flexible guided-wave approach are complementary in that the first is better suited to bridge distances smaller than a few centimeters while the latter is more appropriate for interconnection lengths of several centimeters and more [30].

7.3 DESIGN OF THE OPTICAL BRIDGE

In this case study, the design of the prototype free-space OPB has been based on the layout and the characteristic dimensions of the photonic chip proposed and fabricated within the European Community-funded MEL-ARI project “Optically Interconnected Integrated Circuits” [16]. In this specific case the optical bridge will be used to interconnect an 8×2 array of $250\text{-}\mu\text{m}$ -pitched 850-nm VCSELs with an 8×2 array of $150\text{-}\mu\text{m}$ -diameter InP receivers. The geometrical distance L_{max} to be traveled inside the OPB for this configuration is

8 mm. Using the analytical model describing the confocal microlens relay, it is estimated that for this distance a minimum lens diameter of 115 μm is needed to achieve 99% transmission efficiency while avoiding cross talk. A lens diameter of 200 μm is used, which is the maximum possible lens diameter obtainable with the noncontact deep proton lithographic LIGA mask while respecting the device pitch of 250 μm (see Section 7.4). To calculate the focal number of the microlenses and to optimize the optical throughput of this system, ray-tracing simulations and radiometric calculations using the photonics design software SOLSTIS have been used. In doing so, the simple Gaussian beam confocal model has been generalized. Here, optical throughput has been defined as the ratio of the optical power received by the detector to that emitted by the VCSEL. The interchannel cross talk is calculated as the ratio of the power that falls on the adjacent detector to that impinging on the designated one. The fixed reflection and absorption losses of the OPB are not included in the calculations. For the VCSEL, a circular geometry of 15- μm diameter with a Gaussian emittance and a Gaussian angular intensity distribution with a FWHM angle of 12° is assumed. The VCSEL is simulated using a Monte Carlo method. This way, rays are emitted from the source in a quasi-random manner in accordance with their emission probabilities. These rays are propagated through the system and their intersection with the detection plane is calculated. Figure 7.4 shows the optical throughput of the free-space optical module for different focal numbers of the lens and for different working distances between the chip and the bridge. Optical throughputs above 80% can be achieved for working distances between 450 and 580 μm . In principle, a maximum efficiency of 83% can be reached for a working distance of 510 μm with a focal number for the lenses of 2.6. For these optimized design parameters an optical cross talk of smaller than -40 dB was calculated.

Based on these design characteristics, two basic micro-optical bridge components were fabricated with dimensions as displayed in Fig. 7.5 using deep proton lithography.

7.4 FABRICATION WITH DEEP PROTON LITHOGRAPHY

The technology of deep proton lithography [21] in PMMA is a high-precision rapid prototyping technology for the fabrication of micro-optical elements and micromechanical structures. With this technology different optical functions can be realized in one block to form monolithic micro-optical systems. In addition mechanical support structures and alignment features can be integrated. The fabrication process consists of three basic procedures: a selective irradiation of a PMMA substrate with ion particles followed by either etching the irradiated regions with a specific developer or swelling them with an organic monomer

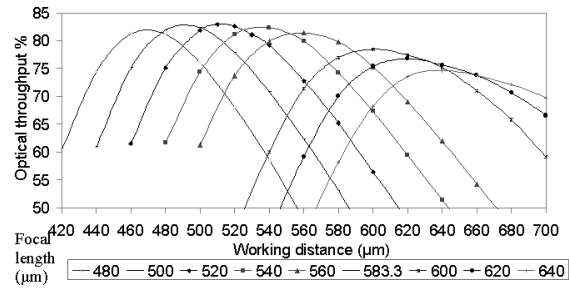


Figure 7.4. Optical throughput of the free-space approach using VCSELs: working distances increase from left to right.

vapor. If required, both processes can be applied to the same sample. During the irradiation procedure the PMMA sample is covered with a patterned high-precision metal LIGA mask [32] which is only transparent for the proton beams at the apertures. The shape of this mask is directly projected onto the PMMA sample where impinging high-energy protons create well-defined domains with reduced molecular weight (Fig. 7.6).

The process is based on the fact that protons transfer energy to the PMMA molecules while propagating in the substrate. This interaction causes molecular chain scissions, reducing the molecular weight of the polymer and changing the chemical properties of the material. Once the irradiated zones have received a large enough dose they will feature a low enough molecular weight, making them susceptible to a binary chemical solvent. The irradiated domains can therefore be selectively dissolved with a special developer, because they show higher solubility than the nonirradiated domains.

In this way, complex structures can be microdrilled or precision cut by moving the sample during irradiation behind a mask, with, for example, a single hole as aperture. The surfaces thus obtained have a high optical quality and can be used as micromirrors.

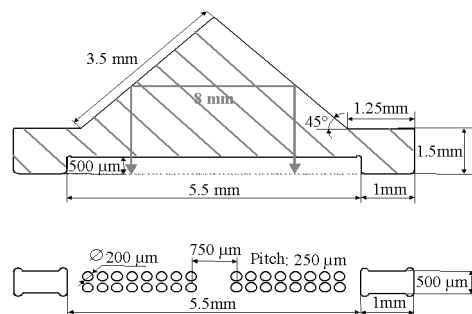


Figure 7.5. Dimensions of the free-space OPB.

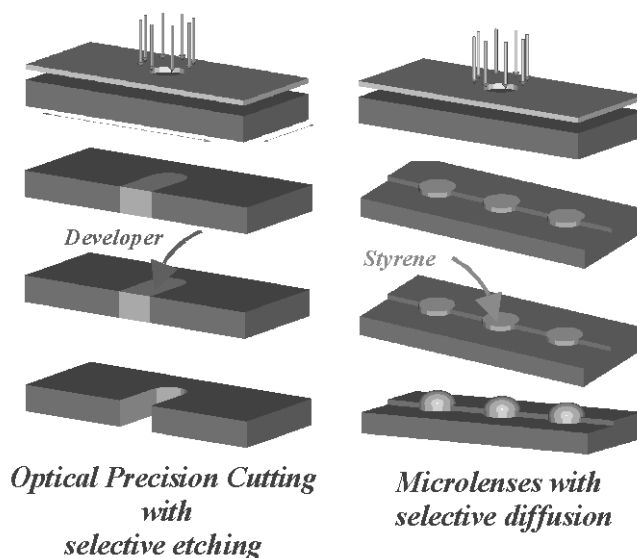


Figure 7.6. Basic fabrication processes of deep proton lithography.

The width of the mirrors is determined by the penetration depth of the ions in PMMA, which in its turn depends on the kinetic energy of the incident particles. To achieve structural depths of several hundred micrometers up to 1 mm, the ions have to be accelerated to relatively high energies. In our experiments we have used protons with a specific energy of 8.3 MeV making it possible to cut through 500- μm -thick PMMA samples. The flatness of the micromirrors in the direction of the ion beam is limited by the lateral straggling of the ions in the material. The flatness of the surface along the length of the mirrors is determined by the rectilinearity of the movement of the translation stage holding the sample and the homogeneity of the dose deposited during this movement. The surfaces of both micromirrors constituting the top prism of the optical bridge have been characterized with a vertical scanning noncontact optical profiler. In Fig. 7.7 one can see that the flatness of the micromirrors along their length can be better than 1 μm and that the roughness is in the range of about 20 nm, while the flatness in the direction of the ion beam is less than 0.5 μm . Figure 7.8 shows a photo of the top prism with its mounting features.

As an alternative to dissolving the irradiated zones, they can be swelled by exposing them to an organic monomer vapor such as styrene. Indeed, only those zones featuring a low enough molecular weight will be receptive to organic monomer material through an in-diffusion process (Fig. 7.6). During this process there is almost no chemical reaction between the broken polymer chains and the in-diffused molecules. However, the in-diffusion of the monomer brings about a volume expansion resulting in a hemispherical surface. Thus refractive microlenses can be fabricated with low focal numbers over a wide range of diameters D ,

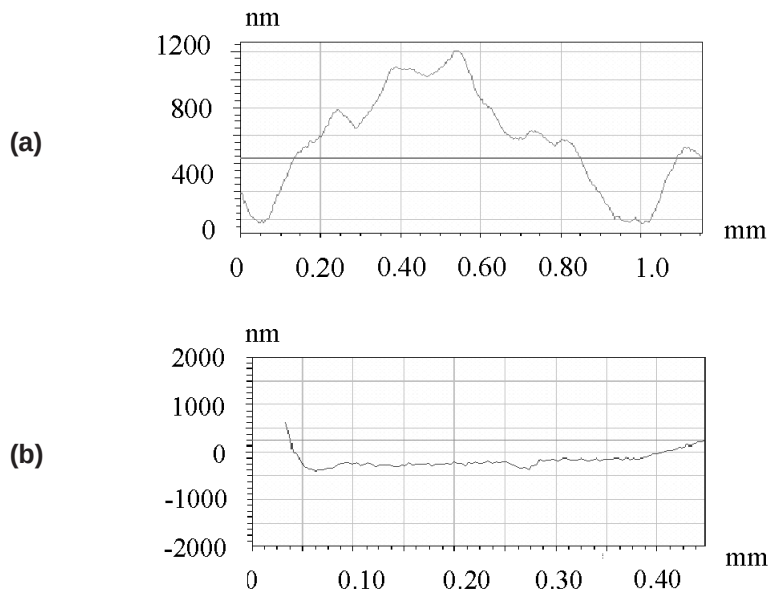


Figure 7.7. (a) Flatness of the micromirrors along their length. (b) Flatness of the micro-mirror in the direction of the ion beam.

corresponding to the circular apertures available on the lithographic LIGA mask. A thermal polymerisation procedure finally prevents the out-diffusion of the styrene monomers and fixes the shape of the microlenses. Two factors are governing the above mentioned irradiation process: the average molecular weight in the irradiated region and the volume of this region.

- The molecular weight can be controlled by the dose deposition. It has been observed that when the dose is too low the molecular weight of the irradiated region is too high to allow a sufficient in-diffusion of the styrene monomer to

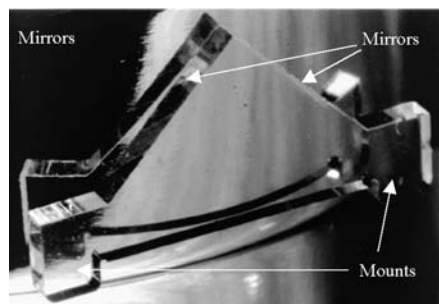


Figure 7.8. Top prism with its mounting features.

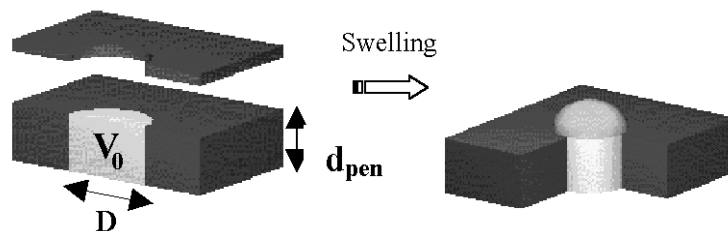


Figure 7.9. Definition of the swelling parameters.

cause a volume expansion. Too high a dose, on the other, hand will result in a local temperature increase which in its turn will destroy the sample.

- The shape of the irradiated volume is a cylinder with a diameter D equal to the circular aperture in the mask and a height determined by the penetration depth d_{pen} of the protons and therefore by their energy (Fig. 7.9).

The volume expansion strongly depends on the amount of monomer diffused into the irradiated zones. This amount can be controlled by the diffusion time and the temperature that influences the monomer concentration in the reactor and the speed of the diffusion. Furthermore, both diffusion time and temperature limit the range in which lens-like shapes are practically achievable. Too short a diffusion time leads to an insufficient volume expansion while too long a diffusion time will destroy the sample. On the other hand, the temperature has to be sufficiently high to create the monomer vapor phase while too high a temperature will start the substrate material to flow.

An accurate calibration of the swelling process is thus necessary to perfectly predict the height and the radius of curvature R of the lenses, and hence their focal lengths f and focal numbers $f/\#$. All these parameters are related by the formulas

$$R = \frac{h^2 + \left(\frac{D}{2}\right)^2}{2h},$$

$$f = \frac{R}{n - 1}$$

and

$$f/\# = \frac{f}{D}$$

in the paraxial approximation.

For calibration purposes, a sample was irradiated with 8.3-MeV protons with doses ranging from 0.1 nC to 2 nC in steps of 0.1 nC. After irradiation this sam-

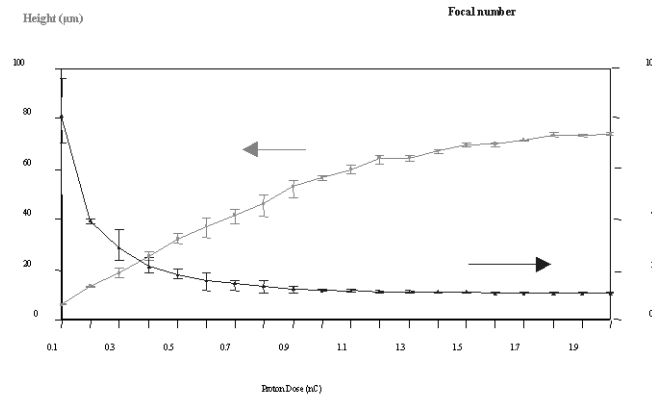


Figure 7.10. Height and focal numbers of 200- μm -diameter lenses as a function of the deposited dose.

ple was placed in a temperature-controlled reactor at 90°C; next, styrene was injected with a syringe while a pressure probe was used to detect a possible leakage of the reactor. Diffusion then took place for 45 minutes. Finally the microlenses were polymerised by reducing the temperature to 65°C and sustaining it for five hours. The dimensions and quality of the microlenses were measured in transmission mode with a Mach-Zehnder interferometer and in reflection mode with a vertical-scanning noncontact optical profiler. All lenses in this 4×20 array had a diameter of 200 μm . Their heights ranged from 6.4 to 74 μm corresponding to a range of focal numbers between 8.12 and 1.08 with standard deviations as plotted in Fig. 7.10.

Figure 7.11 shows the experimental profile of a single 200- μm -diameter lenslet of this array. The height of this microlens is 15.4 μm , which corresponds to

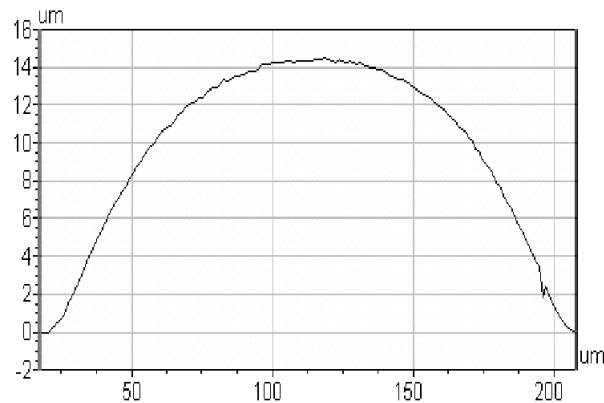


Figure 7.11. 2D profile of a 200- μm lens.

a focal length of 678.3 μm and a focal number of 3.5. The optical surface has a roughness of 20 nm. From these results we can derive that a dose of 0.4 nC is appropriate to obtain microlenses with the desired focal number of 2.6.

Then two arrays of 8×2 identical microlenses were irradiated on the base plate of the optical bridge.

The characteristics of the different individual input and output lenses are displayed in Table 7.2, resulting in an array of spherical microlenses with an average height of $17.6 \pm 2.6 \mu\text{m}$, an average focal number $f/\#$ of 3.3 ± 0.2 and a pitch of $250 \pm 4 \mu\text{m}$.

Figure 7.12(a) shows a photo of the base plate with the alignment holes and the two microlens arrays while Fig. 7.12(b) displays a detail of the microlens array.

7.5 PROOF-OF-PRINCIPLE DEMONSTRATION

Figure 7.13 shows the side view of the assembled free-space 2D multichannel interconnection module. All 16 channels have been used individually to establish an optical link between a single-mode pigtailed 1-mW CW 680-nm laser and a multimode pigtailed silicon detector. Here the single mode input fiber with an NA = 0.11 was used to mimic the emission characteristics of a VCSEL source whereas the multimode fiber with an NA = 0.2 and a core diameter of 50 μm simulated the photodetector. The distance traveled within the bridge is 8 mm and corresponds to a maximum on-chip interconnection distance of 4.25 mm.

h		Input lenses						
top	13.4	17.1	16	14	18.3	15.8	15.3	16.9
bottom	11.9	12.8	19.4	20.4	19.1	19.7	16.8	16.2
f/#		Input lenses						
top	4.0	3.1	3.3	3.8	2.9	3.4	3.5	3.2
bottom	4.4	4.1	2.8	2.7	2.8	2.7	3.2	3.3
h		Output lenses						
top	16.8	17.2	15.4	15.3	16.5	16.6	19.6	14.8
bottom	23.6	17.2	19.1	19.1	14.2	16.7	14	13.1
f/#		Output lenses						
top	3.2	3.1	3.5	3.5	3.2	3.2	2.8	3.6
bottom	2.3	3.1	2.8	2.8	3.7	3.2	3.8	4.0

Table 7.2. Height and focal numbers of an array of input and output lenses.

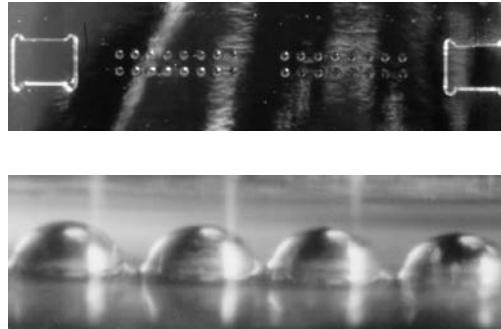


Figure 7.12. (a) Base plate with alignment holes and microlenses; (b) detail of the microlens array.

The optical transfer efficiencies for the 16 channels were measured to be in the range of 40% to 46% with a cross talk between neighboring channels of -22 dB to -27 dB. To demonstrate the potentialities for high-bandwidth intra-MCM data communication, the single-mode input fiber was connected to a standard telecommunication card operating at 1300 nm on the OC-12 SONET carrier level, providing a quasi-random 622Mb/s bit stream. Figure 7.14 shows the open eye-diagram at this bit rate and demonstrates a very reliable data transmission with a bit error ratio less than 10^{-13} .

Because of the short interconnection distances neither dispersion nor absorption will play a limiting factor on the maximum achievable bit rate. Indeed, the demonstration of a 2.48-Gb/s one-dimensional multichannel module based on cylindrical microlenses fabricated with deep proton lithography [23] has already been reported on. In both cases, bit rate was only limited by testing equipment. Although the experimentally obtained optical transfer efficiencies of 40–46% are sufficiently high and the cross talk acceptably low to obtain reliable short-distance

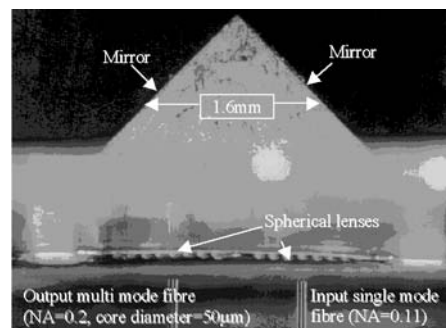


Fig 13. Assembled free-space 2D multichannel interconnection module.

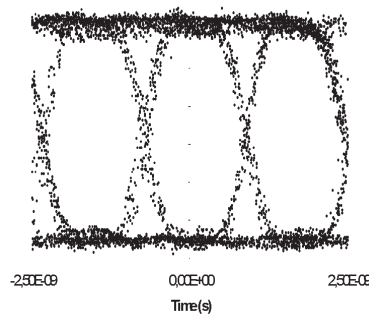


Figure 7.14. Eye-diagram for data communication through a single micro-optical channel at 622 Mb/s and 1300 nm.

parallel data transmission, it is clear that with this first prototype component we do not reach the calculated maximum transmission efficiency of 83% and the -40 dB cross talk. This can be imputed partly to a combination of geometrical and optical fabrication imperfections, partly to misalignment errors and partly to the fact that the input and output fibers are only approximate imitations of the VCSEL source and the photodetector.

7.6 MASS-FABRICATION AND TOLERANCES

Although deep proton lithography is an attractive rapid prototyping technique for micro-optical components, it is far too expensive to be used for mass-production. It is however possible to use the individually fabricated components as masters and to electroplate them by a galvanofforming process similar to that used in the LIGA technique [21, 32]. With this process an inverse of the master is obtained that subsequently can be used as a mold for mass-replication. This section investigates the possible use of injection molding techniques for the replication of free-space OPB with high-optical-quality plastics [33]. To that aim, the tolerance requirements on the injection molded product are derived through a sensitivity analysis based on the same ray-tracing model and similar radiometric calculations as mentioned in Section 7.3. The sensitivity analysis is based on the calculation of the tolerance to manufacturing errors and mechanical misalignments in terms of the degradation of the optical throughput and the increase of interchannel cross talk. The nominal system parameters used are detailed in Table 7.3.

Figures 7.15 to 7.17 show some of the longitudinal, lateral and rotational misalignments that can occur due to manufacturing and assembly errors. To determine the tolerance for each single parameter it is necessary to first define minimum system performance thresholds. In this case, cross talk should stay below -25 dB and the optical throughput is not allowed to drop below 70%.

VCSEL λ	850 nm
VCSEL FWHM	12°
Detector diameter	150 μm
Pitch	250 μm
Microlens diameter	200 μm
Microlens f	520 μm
Working distance	510 μm
Path length bridge	8 mm

Table 7.3. Nominal system parameters.

Table 7.4 lists the perturbations on the different system parameters considered in this study, such as the variation of the top angle of the prism $d\alpha$ and the variation of the focal length of a microlens df .

Figure 7.18 shows the simulation results for the optical throughput and the cross talk for the most sensitive parameters: changes in the working distance, dZ_a , a lateral misalignment of the detector/emitter array, dY_{ed} , and deviations in the top-angle prism, $d\alpha$.

It can be observed for example that for the top angle of the prism a deviation of 22 Arc min is allowed to respect the optical throughput performance threshold of 70%, but that the optical cross talk threshold of -25 dB is what sets the tolerance limit to a deviation of only 16 Arc min, or 0.27° .

The results of this sensitivity analysis are summarized in Table 7.4 and the perturbations ranked in order of importance from 1 to 7. The most critical parameters are those caused by angular fabrication errors or rotational misalignments in the system. Also the lateral alignment of the emitter/detector array is very stringent. The reason for this is that changes of the parameters $d\alpha$, $d\beta$, dY_{ed} , dY_{pb}

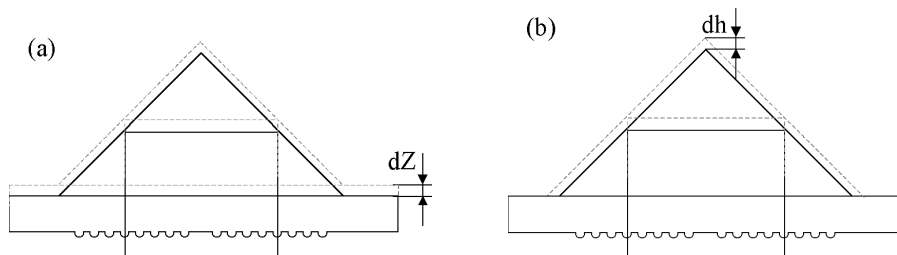


Figure 7.15. Longitudinal misalignment: tolerance on the optical path length in the bridge (a) because of a change of thickness dZ of the base plate, and (b) because of a change in prism height dh .

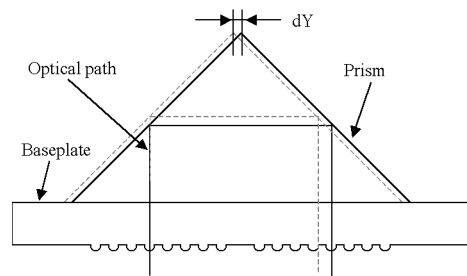


Figure 7.16. Lateral misalignment: tolerance on the lateral shift of the prism with respect to the base plate.

induce coupling of optical energy into adjacent microlenses, causing the cross talk to increase.

Longitudinal misalignments, on the other hand, give rise to a loss in optical throughput due to vignetting at the lens apertures. In the latter case the tolerance is set by the optical throughput threshold rather than by cross talk. The mold proposed here consists of two parts, as can be seen in Fig. 7.19.

The first part combines the tools to replicate the top prism. The flat surfaces of the prism mold, required to shape the mirror surfaces at 45° , will be constructed with separate polished inserts to achieve a high surface flatness combined with a low surface roughness. The second part of the mold integrates the mold-insert for the microlens array. This mold insert will be fabricated by electroplating the master of the microlens array obtained by deep proton lithography.

This mold design makes it now possible to compare the demands on the tolerances with previous results achieved on similar part geometries fabricated by CTP-COIL, a British company that specializes in reproducing on a large scale ultra-high-quality clear optical components. This information has therefore been added to Table 7.4 together with the possible tolerances achievable using further improved tool constructions. From these numbers it is clear that using the ade-

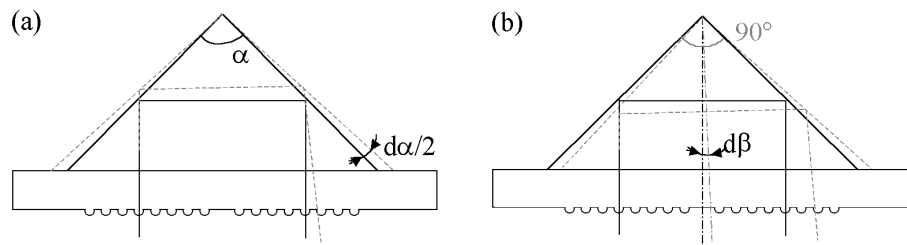


Figure 7.17. Rotational misalignment: (a) tolerance on the top angle and (b) tolerance on the tilt of the prism.

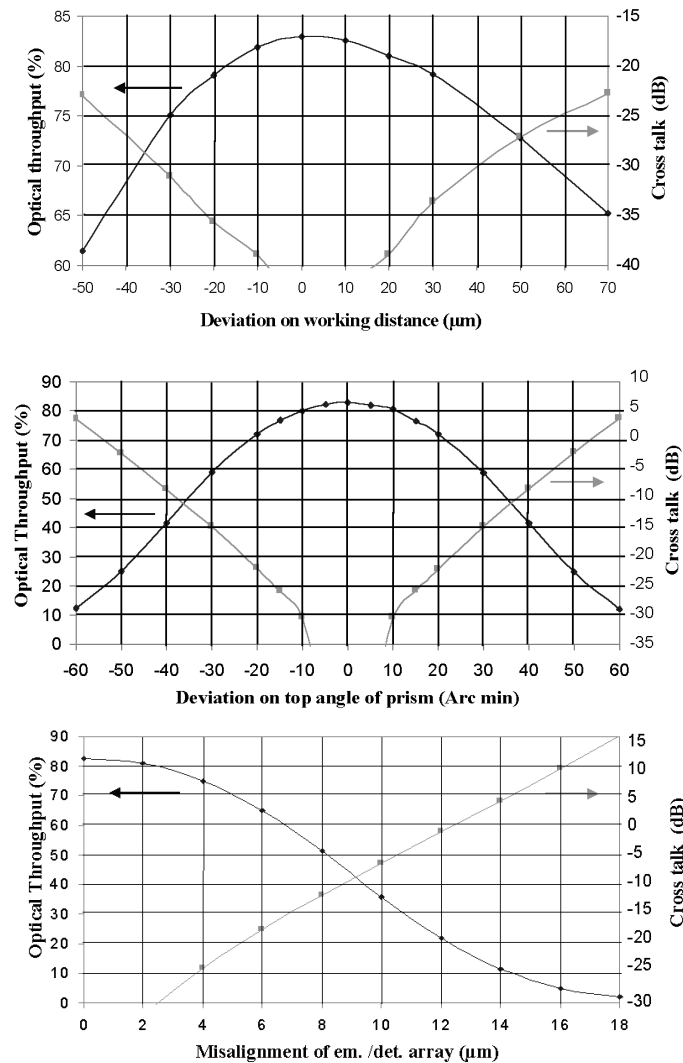


Figure 7.18. Sensitivity of optical throughput and cross talk to (a) dZ_a , (b) dY_{ed} and (c) $d\alpha$.

quate tooling techniques it must be feasible to fabricate a high-quality injection mold and to technologically achieve the production of sufficiently precise photonic-VLSI compatible optical interconnection modules at a very low cost.

7.7 CONCLUSION

This chapter reports on the concept of a scalable multichannel optical data link with Tb/s.cm^2 aggregate bit rate capacity for inter- and intrachip module interconnection distances. A simple prototype module was designed. The individual

Tolerance Parameter		Symb.	Required Tolerance	Rank	Tolerances Previously Achieved on Similar Part Geometries	Possible Tolerances Achievable Using Improved Tool Constructions
Longitudinal Misalignments	Working Distance	dZa	±35 μm	5	Function of Assembly Method	—
	Optical Path Length	dZb	±500 μm	7	±200 μm	±50 μm (Adjustable)
	Focal Length of Lenses	df	±40 μm	6	±25 μm	±10 μm
Lateral Misalignments	Position emitter/detector Array	dYed	±4 μm	3	Function of Assembly Method	—
	Position prism/base	dYpb	±19 μm	4	±10 μm	±5 μm (Adjustable)
Rotational Misalignments	Top Angle of Prism	dα	±0.27°	1	±0.083°	±0.02 (Adjustable)
	Tilt Angle of Prism	dβ	±0.33°	2	±0.083°	±0.04° (Adjustable)
Prism Quality Parameters	Prism Surface Flatness	—	—	—	2λ @ 632.8 nm	λ/4 @ 632.8 nm
	Prism Surface Roughness	—	—	—	300 Å RMS	100Å RMS

Table 7.4. Sensitivity analysis: a summary of the results.

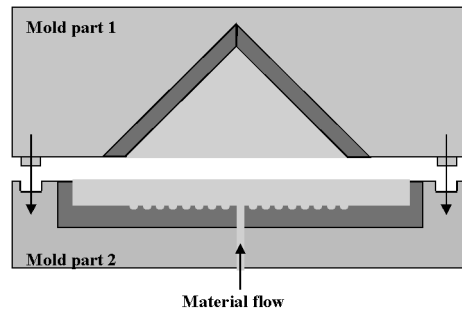


Figure 7.19. Schematic view of the mold for replicating the micro-optical bridge.

components were fabricated with deep proton lithography in PMMA and their optical characteristics were measured. The potentialities for intra-MCM data communication were demonstrated with the assembled module at high bit rates for 16 channels @ 622 Mbit/s per channel with a BER smaller than 10^{-13} while the cross talk between neighboring channels was lower than -22 dB. A sensitivity analysis for misalignments and fabrication errors was performed, which found the resulting requirements on tolerances less stringent than tolerances previous achieved by industry on similar part geometries fabricated with injection molding.

Based on these first results one can conclude that the use of free-space micro-optical modules for fire hose aggregate bit rate capacities over short interconnection distances could be turned into a practical and commercially attractive route to alleviate the interconnection bottlenecks at the inter- and intrachip interconnection level. The perspective of fabricating low-cost, chip-compatible, alignment-tolerant, high-precision, snap-on components with well-established injection molding techniques could bring us one step closer to a fully heterogeneous integrated three-dimensional photonic processing module. Moreover the approach described in this chapter paves the way for the use of novel high-tech plastics, such as long-lived dye-doped polymers and optically transparent polymers with very high glass temperatures, and for taking full advantage of all the peculiar and unique optical, optoelectrical and optomechanical properties that they will have in store.

ACKNOWLEDGMENTS

This work was funded by the European Commission ESPRIT-MELARI project 22641 “OIIC,” by DWTC IUAP 13, by FWO, GOA, IWT ITA-II GBO, and the OZR of the Vrije Universiteit Brussel.

REFERENCES

- [1] H. Scott Hinton, *An Introduction to Photonic Switching Fabrics*. New York, USA: Plenum Press, 1993.
- [2] R. T. Chen and P. S. Guilfoyle, *Optoelectronic Interconnects and Packaging*. Washington, USA: SPIE Critical Review of Optical Science and Technology CR62, Optical Engineering Press, 1996.
- [3] H. Thienpont and P. Chavel, *Journal of Applied Optics A, Pure and Applied Optics, Special issue on Optics in Computing*, vol. 1, no. 2, 1999.
- [4] C. Jung, R. King, R. Jäger, M. Grabherr, F. Eberhard, R. Michalzik and K. J. Ebeling, "Highly efficient oxide-confined VCSEL arrays for parallel optical interconnects," *J. Opt. A: Pure Appl. Opt.*, vol. 1, no. 2, pp. 272–275, 1999.
- [5] T. K. Woodward, A. Krishnamoorthy, "1-Gb/s Integrated Optical Detectors and Receivers in Commercial CMOS Technologies," *IEEE Journal of Selected Topics in Quantum Electronics.*, vol. 5, pp. 146–156, 1999.
- [6] D. A. Louderback, O. Sjölund, E. R. Hegblom, J. Ko and L. A. Coldren, "Flip-Chip Bonded Arrays of Monolithically Integrated, Microlensed Vertical-Cavity Lasers and Resonant Photodetectors," *IEEE Photon. Technol. Lett.*, vol. 11, no. 3, pp. 304–306, 1999.
- [7] R. Pu, C. Duan, C. M. Wilmsen, "Hybrid Integration of VCSELs to CMOS ICs," *IEEE Journal of Selected Topics in Quantum Electronics.*, vol. 5, pp. 201–208, 1999.
- [8] A. V. Krishnamoorthy, L. M. F. Chirovsky, W. S. Hobson, R. E. Leibenguth, S. P. Hui, G. J. Zyzdik, K. W. Goossen, J. D. Wynn, B. J. Tseng, J. Lopata, J. A. Walker, J. E. Cunningham and L. A. D'Asaro, "Vertical-Cavity Surface-Emitting Lasers Flip-Chip Bonded to Gigabits-per-Second CMOS Circuits," *IEEE Photon. Technol. Lett.*, vol. 11, no. 1, pp. 128–130, 1999.
- [9] E. M. Hayes, R. Jurrat, R. Pu, R. D. Snyder, S. A. Feld, P. Stanko, C. W. Wilmsen, K. D. Choquette, K. M. Geibb and H. Q. Hou, "Foundry-fabricated array of smart pixels integrating MESFETS/MSMs and VCSELs," *Int. J. Optoelectron.*, vol. 11, pp. 229–237, 1997.
- [10] J. Van Campenhout, "Optics and the CMOS interconnection problem: a systems and circuits perspective," *Int. J. Optoelectron.*, vol. 12, no. 4, pp. 145–154, 1998.
- [11] The National Technology Roadmap for Semiconductors, San Jose, CA, Semiconductor Industry Association, 1997.
- [12] A. George, J. Krusius, and R. Granitz, 1996, "Packaging Alternatives to Large Silicon Chips: Tiled Silicon on MCM and PWB Substrates," *IEEE Transactions on Components, Packaging, and Manufacturing Technology—Part B*, 19(4), pp. 699–708.

- [13] D. A. B. Miller, "Physical reasons for optical interconnection," *Int. J. Optoelectron.*, vol. 11, no. 3, pp. 155–1668, 1997.
- [14] D. A. B. Miller and H. M. Ozaktas, "Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of the system architecture," *Journal of Parallel and Distributed Computing*, vol. 41, pp. 42–52, 1997.
- [15] M. Haney, and M. Christensen, 1998, "Performance scaling comparison for free-space optical and electrical interconnection approaches," *Applied Optics*, **37**(14), pp. 2886–2894.
- [16] Technology Roadmap: Optoelectronic Interconnects for Integrated Circuits, Microelectronics Advanced Research Initiative (MELARI-OPTO) European Communities, 1998.
- [17] C. Chen, B. Hoanca, C. Kuznia, A. Sawchuck, J. Wu, 1999, "TRANSLucent Smart Pixel ARray (TRANSPAR) Chips for High Throughput Networks and SIMD Processing," *IEEE Journal of Selected Topics in Quantum Electronics*, **5**(2), pp. 316–329.
- [18] T. Szymanski and V. Tyan, 1999, "Error and Flow Control in Terabit Intelligent Optical Backplanes," *IEEE Journal of Selected Topics in Quantum Electronics*, **5**(2), pp. 339–352.
- [19] S. Bond, O. Vendier, M. Lee, S. Jung, M. Vrazel, A. Lopez-Lagunas, S. Chai, G. Dagnall, M. Brooke, N. Jokerst, D. Wills, and A. Brown, 1999, "A Three-Layer 3-D Silicon System using Through-Si Vertical Optical Interconnections and Si CMOS Hybrid Building Blocks," *IEEE Journal of Selected Topics in Quantum Electronics*, **5**(2), pp. 276–286.
- [20] J. Van Campenhout, H. Van Marck, J. Depreitere, J. Dambre, 1999, "Optoelectronic FPGAs," *IEEE Journal of Selected Topics in Quantum Electronics*, **5**(2), pp. 306–315.
- [21] M. Kufner and S. Kufner, *Handbook Micro-optics and Lithography*. Brussels, Belgium: VUB Press, 1997.
- [22] J. P. Menendez, F. Erismann and M. A. Gauvin, "The advantages of plastic optical components," *Opt. Phot. News*, vol. 10, no. 8, pp. 28–30, 1999.
- [23] M. Jöhnck, B. Wittmann and A. Neyer, "64 channel 2D POF-based optical array interchip interconnect," *J. Opt. A: Pure Appl. Opt.*, vol. 1, no. 2, pp. 313–316, 1999.
- [24] R. Bockstaele, A. Van Hove, T. Coosemans, C. Sys, I. Moerman, B. Dhoedt, R. Baets, P. Van Daele, J. Koetsem and L. Van der Torren, "Microcavity LED-based parallel data link using small-diameter (125 μm) plastic optical fibres," *J. Opt. A: Pure Appl. Opt.*, vol. 1, no. 2, pp. 233–236, 1999.
- [25] Y. Li, J. Popelek, L. Wang, Y. Takiguchi, T. Wang and K. Shum, "Clock delivery using laminated polymer fibre circuits," *J. Opt. A: Pure Appl. Opt.*, vol. 1, no. 2, pp. 239–243, 1999.
- [26] M. Kufner and S. Kufner, "Micro-optics for optical interconnects," *Int. J.*

- Optoelectron.*, vol. 12, no. 4, pp. 119–126, 1998.
- [27] G. Verschaffelt, R. Buczynski, P. Tuteleers, P. Vynck, V. Baukens, H. Ottevaere, C. Debaes, S. Kufner, M. Kufner, A. Hermanne, J. Genoe, D. Coppée, R. Vounckx, S. Borghs, I. Veretennicoff and H. Thienpont, “Demonstration of a monolithic multi-channel module for multi-Gb/s intra-MCM optical interconnects,” *Photon. Technol. Lett.*, vol. 10, pp. 1629–1631, 1998.
 - [28] O. Kibar, D. A. Van Blerkom, C. Fan and S. C. Esener, “Power Minimization and Technology Comparisons for Digital Free-Space Optoelectronic Interconnections,” *J. Lightwave Technol.*, vol. 17, no. 4, pp. 546–555, 1999.
 - [29] A. V. Krishnamoorthy and D. A. B. Miller, “Firehose architectures for free-space optically interconnected VLSI circuits,” *J. of Parall. and Distribut. Comp.*, vol. 41, pp. 109–114, 1997.
 - [30] V. Baukens, G. Verschaffelt, P. Tuteleers, P. Vynck, H. Ottevaere, M. Kufner, S. Kufner, I. Veretennicoff, R. Bockstaele, A. Van Hove, B. Dhoedt, R. Baets and H. Thienpont, “Performances of optical multi-chip-module interconnects: comparing guided-wave and free-space pathways,” *J. Opt. A: Pure Appl. Opt.*, vol. 1, no. 2, pp. 255–261, 1999.
 - [31] A. V. Krishnamoorthy and D. A. B. Miller, “Scaling optoelectronic-VLSI circuits into the 21st century: a technology roadmap,” *IEEE of Selected Topics in Quantum Electron.*, vol. 2, no. 1, pp. 55–76, 1996.
 - [32] E. W. Becker, W. Ehrfeld, R. Hagmann, A. Maner, D. Münchmeyer, “Fabrication of microstructures with high aspect ratios and great structural heights by synchrotron radiation lithography, galvanofforming and plastic moulding (LIGA process),” *Micro-electronic Eng.*, vol. 4, pp. 35–56, 1986.
 - [33] J. F. Monk, *Handbook Thermosetting Plastics: Moulding Materials and Processes*. Dorchester, Great Britain: Addison Wesley Longman, 1997.

INDEX

- acousto-optic modulators (AOMs), 218
- air post, 78, 98
- AlGaAs substrate, 76-77
- AlGaAs as wide bandgap material, 20
- aligned wafer bonding, 35-36
- aligned pillar bonding, 35-36
- Allied Chemical, 246
- AMP, 246
- amplitude-encoded binary-phase-only filters, 214
- Antoniadis, Dimitri, 28
- Arizona State University, 79 109
- Asymmetric wells, 205
- asynchronous-transfer-mode (ATM), 112
- AT&T Bell Labs, 15
- ATM Switch, 112-115
- Auburn University, 231

- back-emitting VCSEL arrays, 77, 91
- ball-grid arrays (BGA), 182, 185-187
- bidirectionally emitting I²-VCSELs, 60-61
- bit-error-rate (BER), 165-168, 273
- bottom-emitting VCSEL arrays, 89-90
- bottom-emitting VCSELs, 125
- bottom-emitting I²-VCSELs, 59
- Bump or pad bonding, 225-226
- buried heterostructure, 78

- carrier diffusion, 124
- CD-ROM, 220, 221
- chemical mechanical polishing, 191
- Chirovsky et al., 79
- Choquette, 109
- CLDA2 chip, 242

- CMOS chips, 61, 64, 75-116, 196, 210, 213, 220, 221, 255, 258
- CMOS/VCSEL chip, 63, 75-116
- CMOS integrated circuits (ICs), 75
- coefficients of thermal expansion (CTEs), 8-10
- conductive-attachment top-bottom contacts, 80-110
- coplanar flip-chip bonding, 80-110
- current spreading, 124

- Dapkus, 109
- Deep proton lithography, 260-266
- deposited dielectric Distributed Bragg Reflector (dd-DBR), 50, 51, 55, 57, 59
- Detcom layer, 194, 199, 201
- Detector Array, 240
- Developed molding technique, 226-231
- Dicing errors during dicing of the neodie layers, 237
- Dicing errors in the fabrication of components, 237
- dielectric growth windows (DGWs), 19-20
- dielectric aperture, 78
- differential efficiency, 149-150
- digital-to-analog (DAC), 211
- dual in-line package (DIP), 182

- Eight-bit grayscale 2D MQW-SLM development, 210-212
- electronic integration, 4-5
- Electronic Processing Chip, 241
- emitter-based transceiver array, 196-197
- Emulators for technology characterization, 242-244
- epitaxy on SonG, 33
- epitaxy-on-electronics, 14-27

- fine-pitch wire bonding, 182, 185
- first null device, 125-136
- first peak device, 125-136
- flip-chip bonded, 49, 61-64, 68, 77, 123, 142, 153, 181, 226
- flip-chip (FC), 182, 185
- flip-chip details, 187-188
- focal plan array (FPA), 186, 191-192, 96, 199
- focal plane array fabrication, 193-195
- free space optical link, 165-168
- free space optical interconnects (FSOIs), 122, 123, 136, 144
- free-space 2D multichannel interconnection module, 266-267 free-space OPB, 259, 268
- Free-space approach, 255-256
- full width at half maximum (FWHM), 143, 205, 212, 258, 260
- GaAs MESFET integrated circuits, 75, 76, 79, 115
- GaAs substrate, 93, 101, 104, 111
- gallium-arsenide Ics, 16-18
- gallium-arsenide-on-silicon, 10-14
- Gaussian Beam Model, 256-258, 260
- GE, 231
- George, et al., 253
- Georgia Tech, 79
- Haney and Christensen, 254
- Helix driver, 240-244
- high-density interconnect (HDI), 231
- high-speed P²I²-VCSELS, 66-70
- high-speed optical correlator, 214-215
- High-speed Optical Printing, 218-219
- Honeywell, 79, 109, 237, 238, 239
- hybrid integration, 5, 104, 108
- hybrid integrated I²-VCSELS, 64
- Hyperspectral Data Processing, 215-216
- I²-VCSEL, 51-72
- I²-VCSEL design, 52, 54-55
- IBM, 23
- III-V electronics, 34-35
- III-V materials systems, 6
- Implant Apertured and Index-Guided VCSELS, 51, 70
- implanted apertures, 52-58
- independent optical beam guide, 51, 70
- InGaP as wide bandgap material, 20
- inter- and intra-multichip-module (MCM), 251, 259
- internet-protocol (IP), 112
- intra-MCM data communication, 267, 273
- intracavity contacted current guide, 51, 70
- inverse synthetic-aperture radar (ISAR), 215
- ion implantation, 78
- ISC, 238
- Johnson, R., 231
- Jokerst, 79
- Krishnamoorthy, 79, 109
- Lamination errors, 237
- large-signal modulation response, 164
- Lascola, 99
- Laser Driver Array, 240
- lateral misalignments, 268-272
- layer oxidation, 149
- Lee et al., 80
- LIGA Mask, 260-261, 263, 268
- light-emitting diodes (LEDs) in OEIC applications, 23-24
- light-emitting diodes with microlenses, 136
- limit of bonding density, 102
- liquid phase epitaxy (LPE), 10
- liquid-crystal polymers (LCP), 229
- Liu, 79
- logical devices, 196-197

- long-wavelength infrared (LWIR)
 - imaging, 198
- longitudinal misalignments, 268-272
- Lucent, 15, 36, 79, 109, 226
- LW/LW QWIP FPA, 199, 203
- LWIR focal-plane arrays, 198
- Maracas, 109
- materials issues in integration, 7-10
- Mathine et al., 79
- Matuso et al., 80, 109
- McLaren, 109
- MCT (mercury-cadmium-telluride), 183
- medium-scale integration (MSI), 6
- MEL-ARI project "Optically Interconnected Integrated Circuits, 259-260
- MEMS, lessons from, 27
- MESFET Chip 81, 83, 94, 97, 111, 115
- metalorganic chemical vapor deposition (MOCVD), 8-10
- microelectromechanical systems (MEMS), 6
- microlens fabrication, 137-140
- microlens integration and flip-chip bonding, 153-155
- microlens integration, 140-141
- Miller and Ozaktas, 253
- MIT Lincoln Lab, 28
- MIT, 23
- modulation conversion efficiency factor (MCEF), 68, 161
- modulator-based transceiver array, 196-197
- molecular beam epitaxy (MBE), 8-10, 20, 189-191, 212
- monolithic integration, 3-41, 76, 123, 142-168
- Monolithic Refractive Microlenses, 136-137
- Monte Carlo method, 260
- Moore's law, 5, 225
- Morgan, 109
- MOSIS Foundry Service, 84, 94, 95
- MQW-RCLEDs, 212
- MQW-SLM-based optical correlator, 214
- MSM detector, 143
- MSM-based detectors, 197
- multimode fiber (MMF) links, 168
- multiple quantum well electro-absorption (MQWEA) diode, 49
- Multiple quantum well (MQW) spatial light modulators (SLMs), 214, 217, 218, 220
- MW/LW dualband combination, 199, 203
- MW/MW dualband combination, 199, 203
- nonconductive-attachment top contacts, 80-111
- NTT Optoelectronic Laboratories, 80, 98, 109
- OIDA-UNC-Charlotte, 241-242
- On-wafer Testing, 90-92
- one-color FPA processing, 193-195,
- one-color LW QWIP FPAs, 199
- one-color FPA architecture, 199
- One-color LW QWIP FPA performance, 201-202
- optical scattering losses, 123-125, 130
- optical confinement, 126-132
- optical cross talk, 168-169
- Optical Beam Steering, 216-218
- Optical Storage, 219-220
- OPTOCHIP Project, 25-27
- optoelectronic integrated circuits (OEICs), 3-41
- optoelectronic devices, 5-6
- optoelectronic-VLSI (OE-VLSI), 49, 72
- oxide shrinkage, 148-149
- oxide apertures, 52-58, 129
- oxide confinement, 78, 81, 109
- oxide-aperture layer, 146
- oxide-to-oxide bonding, 37

- p-i-n diode, 207
- p-i-n-based detectors, 197
- P²I²-VCSEL design strategy, 52, 55, 70
- P²I²-VCSEL, 49-72
- parasitic capacitance, 66-67, 106, 125
- patterned Si wafers, 9-10
- patterned GaAs wafers, 9-10
- photodetectors, 22, 122, 142
- photodetectors (PD), 76
- photolithographic precision, 51, 56, 70
- Photolithographic errors, 237
- photolithographically defined ion-implanted apertures, 53
- physical devices, 196-197
- Placement errors during molding, 237
- plasma-assisted bonding technique, 32
- PMMA substrate, 260
- polydimethylglutерimide (PMGI), 137-140
- polyimide bonded, 98
- polymethylmethacrylate (PMMA), 254, 260-262, 273
- postfabrication bonding, 80
- postfabrication processing, 79
- preform wafers, 38-39
- printed wiring board (PWB), 182
- proton implant, 98, 109
- Pseudorandom bitstream (PRBS), 65
- Pu et al., 79
- quantum wells (QW), 141, 151-152, 156, 208-209
- quantum confined Stark effect (QCSE), 207
- quantum-efficiency spectrum, 147
- quantum-efficiency, 145-152
- quantum-well infrared photodetector (QWIP), 186, 189-193, 198, 205
- QWIP architecture, 199-201
- radius of curvature (ROC), 137-140
- reactive-ion etching (RIE), 137-140, 193
- Reactive-ion Etching and Molecular Beam Epitaxy (RIE-MBE), 189-193
- Receiver/Preamplifier Array, 240
- Research Devices M-8A flip-chip aligner/bonder, 155
- resonant cavity enhanced photodetectors, 143
- resonant-cavity light-emitting diodes (RCLEDs), 212, 220
- responsivity spectrum, 159
- RF-cross-talk, 168-169
- rotational misalignments, 268-272
- Sanders, 79
- Sandia, 98, 109
- scanning-electron micrograph (SEM), 82-83, 87, 89, 140, 146, 154, 155, 184, 191, 192, 238, 239
- second null device, 125-136
- secondary ion mass spectroscopy (SIMS), 134
- self-assembly wafers, 38-39
- short-range high-density optical interconnects, 252-254
- Si substrate, 80
- Si substrate, 184
- Si Substrate, 99
- Si readout integrated circuit (ROIC), 186, 190, 192, 199
- Si-on-GaAs, 27-35
- SIA Semiconductor road map, 251-252
- silicon beam lead interconnect (SBLI), 183-185, 186
- silicon-on-gallium arsenide (SonG) wafers, 27-35
- silicon-on-insulator (SOI) technologies, 27
- SLM operation (binary 2D and grayscale ID), 209-210
- SLM Operation, 207

- small-signal modulation response, 160-164
- smart/high-speed detector arrays, 212-213
- smart-photodetector array (SPA), 113-115
- Smith, 109
- SOLSTIS, 260
- spatial light modulator (SLM), 189, 195-196, 220
- spatial-light modulator (SLM), 113-115
- standing wave null device, 129-131
- Stark effect, 152, 205, 206
- static thermal transfer, 100-101
- substrate removal, 76-77, 86
- symmetric wells, 205

- tapered apertures, 126-136, 164
- The Horner efficiency, 215
- thermal cross talk, 168-169
- thermal resistance, 95-104 111, 115, 124, 230
- thermal expansion in GaAs-on-Si
- thermal resistance reduction, 102
- thermal stability of electronics, 18
- thermal cross talk, 95-104, 115
- thermal conductivities, 100-101
- thermal impedance of flip-chip bonded VCSELs, 158
- thermal transfer simulation, 115
- thermal expansion, 8-10
- thermal properties, 110, 116
- thermal dissipation, 122, 230
- thermal-compression bonding, 81-82
- thin apertures, 124, 131, 164
- Thin-film bonding, 226
- Thorlabs, 243
- tilt, 93, 105-106
- top-emitting I²-VCSELs, 57-58
- top-emitting VSCSEL arrays, 89-92, 142
- Trezza, 79
- Tunable QWIP FPAs, 204-205
- two-color FPA architecture, 200
- two-color FPA processing, 194-195
- Two-color QWIP FPAs, 202-204

- Typical Properties of Molding Material, 235

- U.S. Army Aviation & Missile Command (AMCOM), 215
- UC Berkeley, 80, 98, 99, 109
- UC San Diego, 80
- UCSB, 98
- University of Colorado, 109
- USC, 98, 109

- Vander Lugt correlator architecture, 214
- VCSELs on CMOS, 104
- vertical cavity surface emitting laser (VCSEL), 5, 9
- very large-scale integration (VLSI), 6
- Vitesse Semiconductor, 28, 83-84, 94, 95
- VSCSEL Array, 239
- VSCSEL-CMOS smart pixel, 94-95
- VSCSEL-MESFET smart pixel, 94-95

- Wafer fusion, 226
- wall-plug efficiency, 134-136
- WDM arrays, 168-169

- Yeh and Smith, 80