High-Speed Modulation of Resonant CMOS Photonic Modulators in Deep-Submicron CMOS

by

Benjamin Moss

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2014

Signature redacted

Author .....
Department of Electrical Engineering and Computer Science
August 29, 2014

Signature redacted

Certified by .....
Vladimir Stojanović
Thesis Supervisor

Signature redacted

Accepted by .....
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Theses

High-Speed Modulation of Resonant CMOS Photonic Modulators in Deep-Submicron CMOS

by

Benjamin Moss

Submitted to the Department of Electrical Engineering and Computer Science
on August 29, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Processor manufacturers have turned to parallelism to continue to improve processor performance, and the bandwidth demands of manycore systems are rising. Silicon photonics can lower the energy-per-bit of core-to-core and core-to-memory interconnects while simultaneously alleviating bandwidth bottlenecks. In this work, methods of controlling the amount of charge entering the diode structure of a photonic modulator are investigated to achieve high energy efficiency in a constrained monolithic process. Two digital modulator topologies are simulated, fabricated and tested. One circuit topology, intended to drive a carrier-injection-based ring modulator, uses a digital push-pull topology with preemphasis to reduce the energy-per-bit and to prevent the ring's optical passband from shifting to the next optical channel. The second circuit topology drives a depletion-mode modulator device for high energy efficiency and speed. High-level system modeling is addressed, as well as practical considerations such as packaging. This work marks the first monolithic transceiver in a zero-change CMOS process, and the most energy-efficient monolithically-integrated modulator in a sub-100 nm CMOS process.

Thesis Supervisor: Vladimir Stojanović

Acknowledgments

I owe thanks to a great number of people for their guidance, inspiration, friendship, and support.

First I would like to acknowledge my family, especially my parents and two brothers, who have always been there with encouragement to keep me on the right track.

I could not have hoped for better colleagues at the Integrated Systems Group. Thank you for your incredible teamwork skills through all of our challenges.

I also owe a heartfelt thanks to Susan Bates, Daryl Beetner, Van Stoecker, and Scott Blomquist for being truly inspiring individuals to me.

Finally, I wish to acknowledge my illustrious research advisor, Vladimir Stojanović, for his continued support and dedication to my graduate education at MIT, and for setting a tremendous example as a scientist and researcher.

Contents

1 Motivation **19**

1.1 Introduction 19

1.2 A New Technology: Silicon Photonics 23

1.2.1 A Conceptual Photonic Integrated Link 24

1.2.2 Different Approaches to Silicon Photonics 26

1.3 Monolithic Integration Technology Development Platforms in Commercial Processes 33

1.3.1 Processor-Side 35

1.3.2 Memory-Side 36

1.3.3 Silicon Photonics Area and Energy Advantage 38

1.4 Thesis Focus 42

2 The Modulator Device **45**

2.1 Carrier-Injection Modulator Device Analysis 48

2.1.1 Carrier Injection Ring Modulator Dimensions 49

2.1.2 Free Spectral Range 50

2.1.3 Free Carrier Refractive Index Change 52

2.1.4 Phase Shift Due to Refractive Index Change 55

2.1.5 Resonance Shift 56

2.1.6 Uniform Carrier Injection 59

2.1.7 A Ring Modulator at $\lambda = 1220$ nm 63

2.1.8 Finger Design 64

2.1.9 The P-I-N Diode 66

2.1.10 P-I-N Diode Simulation Using Sentaurus 66
2.1.11 The Physical Carrier-Injection Design 70
2.2 Charge-Depletion Modulator Device 72
2.2.1 Charge-Depletion Ring Modulator Dimensions 73
2.2.2 Determining the Depletion Region Width 74
2.2.3 Determining the Index Shift, Absorption Change, and Resonance Shift 76
2.2.4 Determining Transmissivity 78
2.2.5 The Physical Carrier-Depletion Design 81
2.3 Conclusion 82

3 The Modulator Driver **83**
3.1 Introduction 83
3.2 A General Modulator Driver Model 84
3.3 Carrier Injection Modulator Driver 86
3.3.1 A Case for Pre-Emphasis 86
3.3.2 Carrier-Injection Driver Circuit 90
3.3.3 Advantages of Signal Conditioning 99
3.3.4 Experimental Test Setup 101
3.3.5 Experimental Results 103
3.4 Depletion-Mode Modulator Drivers 106
3.4.1 Design Constraints 106
3.4.2 Processor (EOS) Platform: Charge-Pump Carrier-Depletion Driver Circuit 107
3.4.3 Processor (EOS) Platform: DCVS-NAND Thick-Oxide Carrier-Depletion Driver Circuit 110
3.4.4 Memory (Micron) Platform: Carrier-Depletion Driver Circuit . 112
3.5 EPHI 114
3.6 Conclusion 118

4 Post-Processing and Bonding **121**

4.0.1 Frontside Release 123

4.0.2 Full Release on Handle 124

4.0.3 Full Substrate Release and Transfer 127

4.0.4 Flip-Chip Full Release 130

4.0.5 Flip-Chip Partial Release 133

4.1 Processing Conclusion 133

5 System Modeling **135**

5.1 Introduction 135

5.2 Photonic Link Components 136

5.2.1 Optical Data Receiver Modeling 137

5.2.2 Single Channel Link Tradeoffs 141

5.3 Towards a Full WDM Link 142

5.4 WDM Photonic Link Evaluation 142

5.5 Conclusion 144

A Tables **147**

A.1 Depletion-Mode Modulator Verilog-A Model 151

List of Figures

1-1 Gordon Moore's 1965 Observation[6] 19

1-2 20

1-3 Interconnect bottlenecks in a manycore system 21

1-4 Interconnect trends in high-performance computing.[2] 22

1-5 A conceptual integrated link [4] 24

1-6 A 256-Core integrated photonic system [4] 25

1-7 Discrete integration 28

1-8 Hybrid integration examples. 28

1-9 Hybrid integration examples: Die-to-die bonding. 29

1-10 Hybrid integration examples: Oracle's die-to-die solution [?]. 30

1-11 Monolithic integration with thick-BOX SOI. 31

1-12 Monolithic integration (this work). 32

1-13 Technology Development Platforms in both Processor and Memory Processes 34

1-14 EOS: A monolithically-integrated photonic technology development platform in CMOS. A cell in a multi-λ WDM link is shown. This figure is a 3-D ray-traced representation of the actual taped-out GDS electronic-photonic circuit layout. 35

1-15 Substrate undercut using XeF_2 in a bulk-CMOS process [12]. 37

1-16 An example of XeF_2 over-etch causing catastrophic circuit damage. . 37

1-17 Electrical backend of a photonic link 41

2-1 Ring modulator Lorentzian frequency response, both passive and under carrier-injection (under carrier-depletion, T_n would shift to the right instead of the left). 46
2-2 Carrier-injection ring modulator diagram top view (not to scale) . . . 48
2-3 Carrier-injection ring modulator dimensions (heater ommitted) 49
2-4 Measured EOS18 ring response; note the Lorenzian response repeating, for an FSR of 4 nm for this particular ring. 51
2-5 Silicon free carrier index change at $\lambda = 1220\,\text{nm}$[33] 54
2-6 Free electron and hole concentration 60
2-7 P-I-N finger structure. 65
2-8 Simulation design flow. 66
2-9 P-I-N cross-section showing Sentaurus simulation mesh. 67
2-10 Sentaurus simulation of P-I-N diode I-V curve for various I-region lengths. 68
2-11 Sentaurus simulation of injected charge concentration vs. voltage. . . 69
2-12 EOS8 Carrier-Injection Ridge Modulators for 1550 nm and 1280 nm with cross-sections [25] 70
2-13 EOS12 Carrier-Injection Modulator Device 71
2-14 Micron platform depletion-mode ring modulator diagram (not to scale) 72
2-15 Depletion-mode ring modulator dimensions (heater ommitted; not to scale) 73
2-16 Silicon free carrier absorption change at $\lambda = 1200\,\text{nm}$[33] 78
2-17 Device requirements to reach the target $ER_{dB} = 6\text{dB}$ for various values of IL_{dB} [10]. 79
2-18 Optical eye diagram (extracted simulation) of the 456 fJ/bit D1 modulator driver at 5 Gb/s, with $ER_{dB} = 11$ dB. This speed is reaching the limit of the energy-optimized process, resulting in significant ISI at this datarate. 80
2-19 EOS16 Carrier-Depletion Modulators 81
2-20 EOS16 Carrier-Depletion Modulator Response 82

3-1 Electrical model of modulator device and driver. 84
3-2 Energy costs to reach the target $ER_{dB} = 6$ dB for various values of IL_{dB}. 85
3-3 First order charge rise time schematic. 87
3-4 A simulation demonstrating how pre-emphasis can improve eye quality, raising the bitrate (1 ns carrier lifetime on random data). 88
3-5 Load line showing pre-emphasis driver and forward-bias driver. 90
3-6 Conceptual schematic of the EOS-series carrier injection modulator driver. 91
3-7 Conceptual schematic of the EOS-series carrier injection modulator driver. Segments A and B are active at the beginning of a 0-1-bit transition. 92
3-8 Conceptual schematic of the EOS-series carrier injection modulator driver. Segment B is active for the full duration of a 1-bit. 93
3-9 Conceptual schematic of the EOS-series carrier injection modulator driver. Segment C is active at the beginning of a 1-0 transition, reverse-biasing the diode and sweeping carriers out of the intrinsic region. . . 94
3-10 Conceptual schematic of the EOS-series carrier injection modulator driver. Segment D is active for the duration of a 0-bit after the pre-emphasis pulse has ended. The weak forward bias pre-charges the P-I-N diode without changing the optical logic state. 95
3-11 Configurable parameters of the P-I-N current profile. 96
3-12 Carrier injection driver final DCVS driver heads 97
3-13 Delay element schematic 98
3-14 Micrograph of carrier-injection modulator driver and device. 98
3-15 EOS12 pre-emphasis profile, unconditioned and conditioned 100
3-16 EOS12 Step Response, unconditioned and conditioned 101
3-17 Table setup 102
3-18 Dual-chip setup 102
3-19 Optical eye diagrams of the injection-based modulator driver. 103

3-20 Measured electrical power of 1280 nm injection-based modulator . . . 104
3-21 Optical 2.5 Gb/s eye diagram of signal-conditioned EOS carrier-injection modulator at 1550 nm . . . 105
3-22 Measured electrical power of signal-conditioned EOS carrier-injection modulator at 1550 nm . . . 105
3-23 Depletion-mode driver schematic . . . 107
3-24 Depletion-mode driver schematic - 1 bit - forward bias . . . 108
3-25 Depletion-mode driver schematic - 0 bit . . . 109
3-26 Simulated depletion-mode modulator driver eye diagram at 10 Gb/s, Monte Carlo simulation on a pseudo-random data sequence . . . 109
3-27 DCVS-NAND Thick-Oxide Carrier-Depletion Driver Circuit . . . 110
3-28 DCVS-NAND Thick-Oxide Modulator Energy and Eye Diagrams . . 111
3-29 Micron modulator driver schematic. . . . 112
3-30 Micrograph of Micron modulator device. . . . 113
3-31 Energy-per-bit of the Micron modulator driver. . . . 114
3-32 The EPHI DDR modulator driver . . . 115
3-33 Extracted eye diagram of the modulator driver output (with 20fF load) at 5 Gbps. . . . 116
3-34 The 50 μm x 50 μm modulator driver. . . . 117
3-35 Simplified modulator driver. . . . 118
3-36 "Ultralite" modulator driver. . . . 118
3-37 Modulator driver, fed directly from a CML-to-CMOS converter. . . . 119
3-38 Layout of 150m x 50m CML-to-CMOS (on the left) and modulator driver (on the right). . . . 119
3-39 CML-to-CMOS converter attached to the lite inverter-based modulator driver. . . . 120
3-40 CML-to-CMOS converter attached to the ultralite single-inverter driver. 120

4-1 An example of XeF_2 over-etch causing catastrophic circuit damage. . 123
4-2 Substrate undercut using XeF_2 in a bulk-CMOS process [12]. . . . 124

4-3 The die after release. 127
4-4 A glass slide has been diesawed into replacement substrates. 129
4-5 Completed substrate-transferred die. 130
4-6 A fully released, flip-chip-bonded EOS 18 die 132

5-1 An integrated WDM photonic link. A continuous-wave (CW) multi-λ laser is coupled onto the chip through a vertical-coupling grating structure. Once on chip, frequency selective ring-resonant modulators encode digital bitstreams onto their resonant wavelengths. Each wavelength propagates along the waveguide (and possibly off-chip) until it is routed through a matching drop ring to an integrated photodiode. An optical receiver forms a bit decision based upon the photodiode photocurrent. Clock signals are routed both optically along the waveguide and electrically through local H-trees. Ring tuning circuits are used to tune the resonance of the modulator and drop rings. 136
5-2 Resistive receiver sensitivity. 138
5-3 Transimpedance Amplifier. 139
5-4 TIA design example at C_p=25 fF, DR=5 Gb/s. 139
5-5 TIA performance. 140
5-6 Integrating receiver design and performance. 140
5-7 Data rate tradeoffs for a single photonic link for 4 integration scenarios. C_P=5 fF represents monolithic integration, while C_P=25 fF is expected for a TSV connection to an optical die. Channel losses of 10 dB and 15 dB correspond to on-chip and chip-to-chip links, respectively. . . . 142
5-8 Optimized power vs. data-rate for different aggregate link throughputs for Loss=10 dB, C_P=5 fF. For tuning, we assume a bit-reshuffler backend and electrically-assisted tuning with local variation σ_{rL}=40 GHz and systematic variation σ_{rS}=200 GHz. Note that the number of WDM channels changes with data rate (Channels = Throughput / Data-Rate). 143

List of Tables

1.1 Area and Energy Comparison of Electrical and Optical Systems . . . 38
1.2 Optical Power Budget [4] 39
1.3 Optical Data Transmission Latency 39

5.1 Link Evaluation Parameters 141

A.1 Optical Power Budget 147
A.2 Optical Data Transmission Latency 148
A.3 Symbol description 150

Chapter 1

Motivation

1.1 Introduction

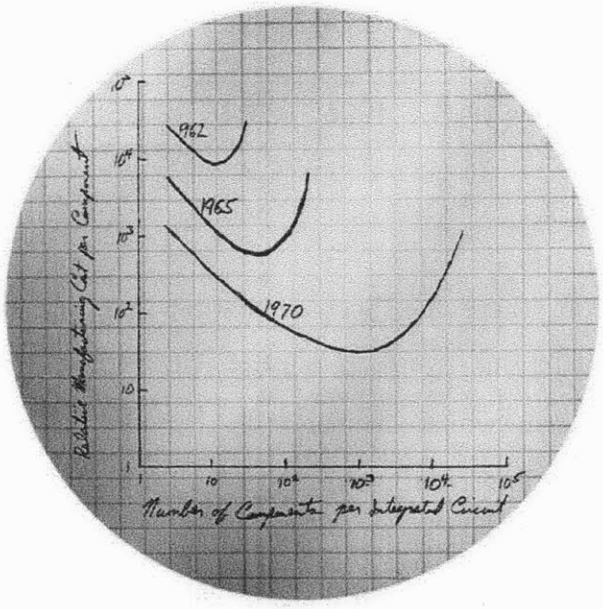


Figure 1-1: Gordon Moore's 1965 Observation[6]

In 1965, Intel co-founder Gordon Moore famously observed that semiconductor minimum-cost-component complexity – i.e., transistor density – was doubling regularly[21]; by 1975, he projected a doubling at roughly every two years[22]. Ever since, the semiconductor manufacturing industry has been driven by this self-fulfilling

"Moore's Law" (Figure 1-1). From the introduction of the microprocessor in the 1970s through the early 2000s, processor manufacturers achieved performance gains largely through clock speed scaling, made possible through the inexorable shrinking of transistors.

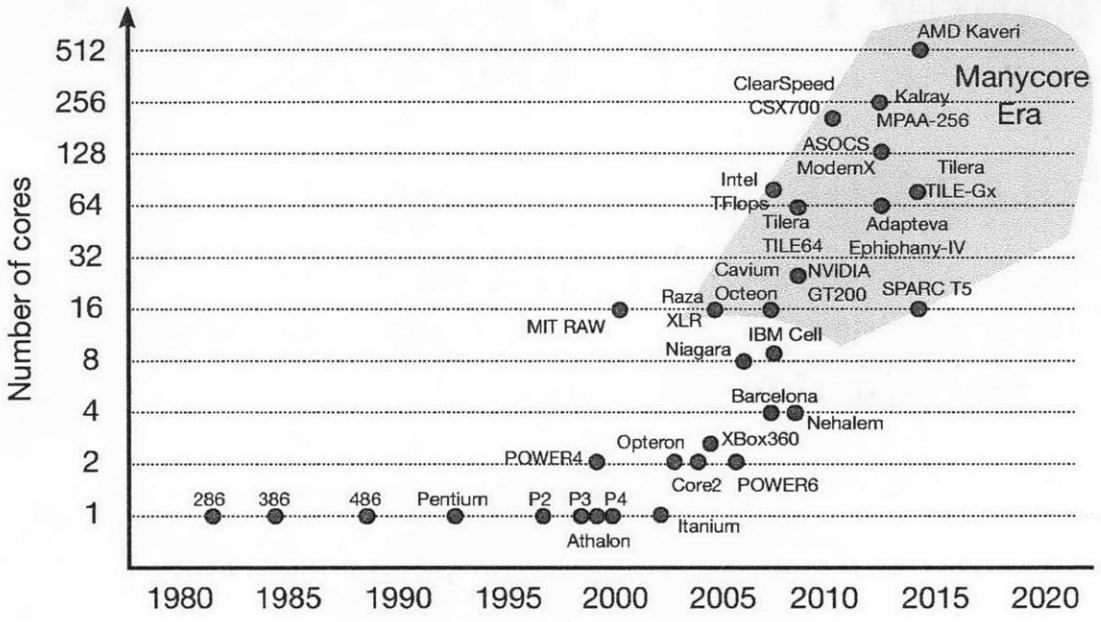


Figure 1-2

Since the year 2000, there has been a slowdown in the rate of uniprocessor performance improvement due to power density constraints, diminishing returns of instruction-level parallelism, wire delays, and direct-random-access-memory (DRAM) access latency [35]. This slowdown has prompted all major processor manufacturers to embrace parallelism and shift to multi-core architectures, where multiple simpler cores exploit thread and data level parallelism [28]. The number of cores on a single die has been steadily rising, bringing us into a "Manycore Era" (Figure 1-2) with dozens or even hundreds of cores on a single die. The core count is expected to continually increase to maintain a steady performance improvement, and a corresponding increase in main memory bandwidth is necessary if the increased core count is to result in improved application performance [4].

These manycore systems have two main interconnect bottlenecks, shown in Figure 1-3. An off-chip core-to-memory interconnect network must feed data back and forth

from DRAM to the processor die; in a modern desktop, a DDR3 SDRAM interface is used, with a maximum bisection bandwidth peaking at just over 100 Gb/s. Additionally, an on-chip core-to-core interconnect fabric connects the cores to each other. In a traditional electrical system, these are two separate interconnects with separate energy costs.

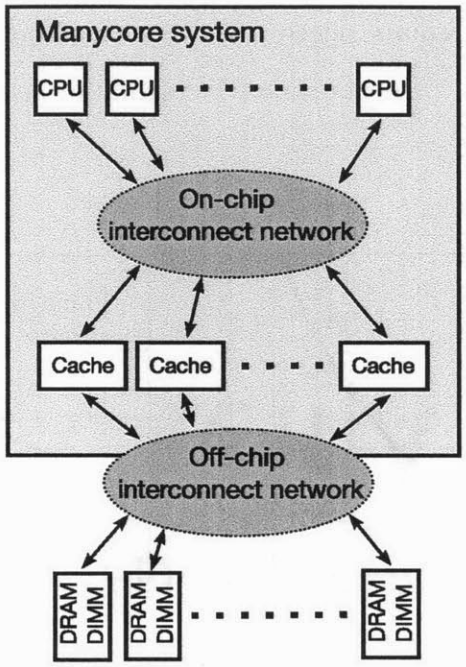


Figure 1-3: Interconnect bottlenecks in a manycore system

Existing electrical high-speed chip-to-chip interconnects will be unable to cope with the increasing communication demands of systems with multi-core architectures. Hybrid Memory Cube technology is emerging offering 1 Tb/s per channel (8-15 Gb/s per pin), but at very high cost per bit due to its low capacity and expensive SOI logic base chip. On the other hand, new DDR standards like DDR4 deliver improved capacity but still lack the required per-chip-bandwidth, with per pin data rates below 4 Gb/s. Increasing copper interconnect bandwidth involves increases in circuit complexity (resulting in higher energy cost) and material cost; achieving order-of-magnitude gains in copper in both energy efficiency and bandwidth simultaneously is extremely challenging, if not impossible.

Figure 1-4 shows the trend of several metrics over a ten-year window. The compu-

tational performance per chip is expected to reach 10 teraflops by 2017; at 1 Byte/flop, such a system would require a memory bandwidth of 80 Tb/s. The total off-chip I/O power is hard-limited in high performance computing to 140 W, simply because the melting point of silicon is fixed and does not scale with transistor length. With this hard limit, the off-chip I/O energy cost would need to stay below 2 pJ/b by 2017. Such a system would also require nearly 16,000 package pins to meet its bandwidth demands, but the ITRS [2] projects that the maximum package pin count will stay below 6000 by 2017.

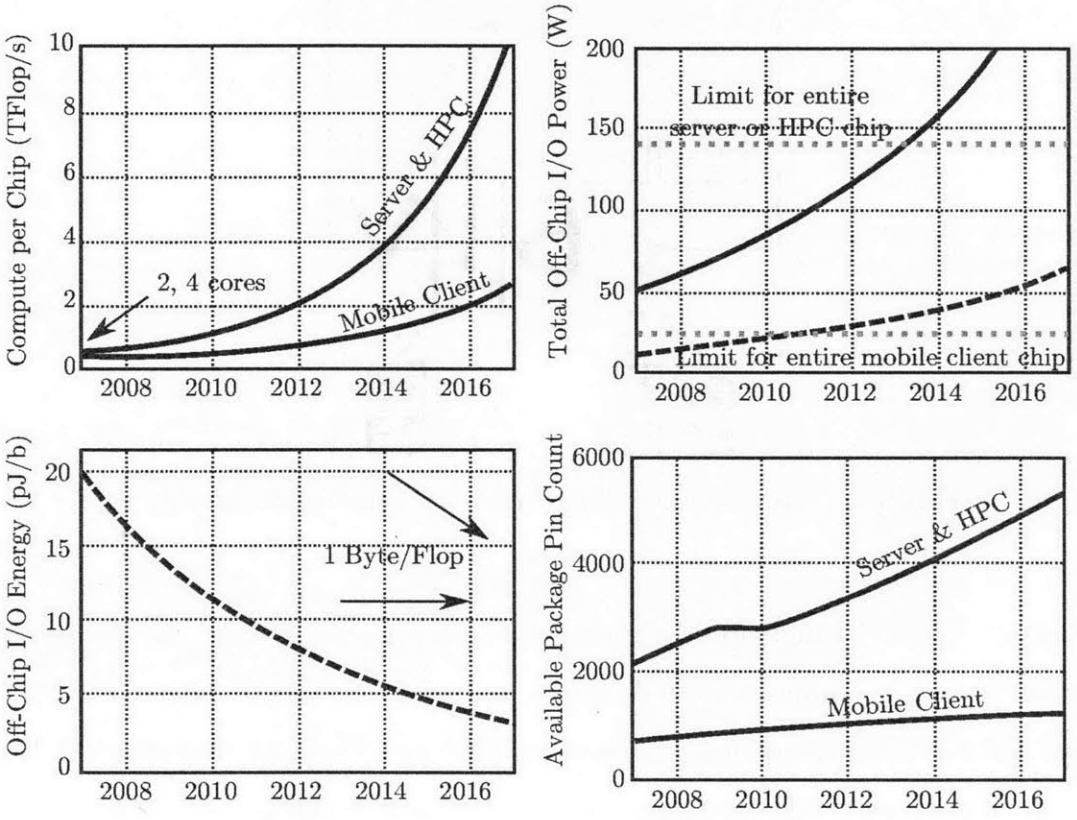


Figure 1-4: Interconnect trends in high-performance computing.[2]

It is extremely difficult for copper interconnects to simultaneously achieve huge improvements in both energy-efficiency and bandwidth density. Because the bandwidth provided by future electrical interconnects will be insufficient for these multi-core systems, the industry is exploring alternate technologies to alleviate these bottlenecks. The system architects and circuit designers seeking an alternative technology are

subject to a number of practical restraints. Given the power-constrained and cost-sensitive nature of the problem, only technologies that offer high bandwidth density (or, equivalently, low area), good energy efficiency, and in the short term, compatibility with existing mainstream manufacturing methods such as bulk-complimentary-metal-oxide-semiconductor (bulk-CMOS) and thin-silicon-on-insulator (thin-SOI) will be feasible contenders to strictly electrical designs in the commercial marketplace [4]. This problem spans many layers, and new technologies must be evaluated at the system, circuit and device levels.

1.2 A New Technology: Silicon Photonics

As future many-core architectures begin demanding bandwidths well over 1.0 Tb/s, new technologies at various levels of maturity are rising to the task of meeting this manycore bandwidth challenge. Basic requirements for a feasible contender to electronic interconnects include high bandwidth density (or, equivalently, low area and packaging), good energy efficiency, and compatibility with existing processor design processes, i.e. bulk-CMOS and thin buried oxide (thin-BOX) SOI. One promising new technology which is quickly gaining industry traction is integrated silicon photonics. Silicon is patterned into traditional electronics as well as optical devices, which gives the resulting interconnect the ability to leverage the enormous bandwidth advantage of optical dense wavelength-division multiplexing (DWDM) to offer exceptional bandwidth density. The integrated nature of having optical devices in close proximity to the driving electronics lowers the energy cost of these devices. Using silicon for integrated optical interconnects is an idea that has existed for decades and has been suggested since the 1970s [37]. The first polysilicon waveguides with losses in the range of 34 dB/cm were reported in [8] in 1996; by 1999, splitters had been demonstrated as well in [18]. The high propagation losses due to scattering, low electrooptic coefficient, low light-emission efficiency and high coupling losses prevented this technology from being adopted until closer to the 2000s, as deep-submicron fabrication techniques became advanced enough to create usable photonic devices [17].

1.2.1 A Conceptual Photonic Integrated Link

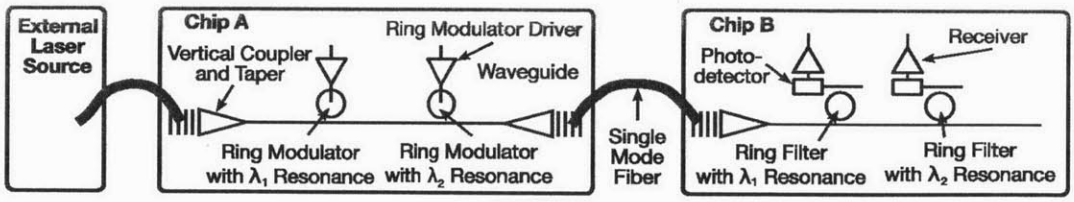


Figure 1-5: A conceptual integrated link [4]

Silicon photonics shows promise in reducing the energy cost and improving the bandwidth density of both on-chip interconnects and off-chip I/O. Figure 1-5 illustrates the concept of a fully integrated silicon photonics system, where a multi-core processor (Chip A) communicates with off-chip memory (Chip B). The simple DWDM link is used for both on-chip (core-to-core) and chip-to-chip (core-to-memory) communication. Due to the indirect silicon bandgap, there are no known high-efficiency laser sources in silicon, so without extensive process customization, unmodulated light must come from an off-chip laser source [4].

In this example, two continuous wavelengths of light (λ_1,λ_2) originate from an off-chip laser and are carried by a single-mode optical fiber. The fiber is perpendicular to the surface of chip A, where a vertical-coupler-grating tapers the light into an on-chip photonic waveguide. This photonic waveguide is designed in the poly-silicon layer (which is traditionally used for transistor gates) or the silicon-body layer (traditionally used for transistor sources and drains). The photonic waveguide carries the unmodulated light past a series of resonant ring modulators [15], which modulate the intensity of the light at the resonant wavelength. Modulated light continues through the waveguide, exits chip A through a vertical-coupler-grating into another optical fiber, and is then coupled via a vertical-coupler-grating into an on-chip waveguide on chip B. On chip B, each wavelength is "dropped" onto a separate waveguide by a tuned resonant ring filter. Each data stream is then converted from the optical domain back into the electrical domain by receiver circuit [11]. Although not shown in Figure 1-5, information can be simultaneously sent in the reverse direction using different wavelengths (λ_3,λ_4) coupled into the same waveguide on chip B and received

by chip A.

Comparing Figure 1-3 to Figure 1-5 also reveals another key advantage of this architecture; two separate interconnect networks have now been unified into a single network, further lowering the overhead necessary. This conceptual diagram also shows how DWDM drastically increases the bandwidth density of the system by combining several wavelengths onto one waveguide without increasing the area overhead.

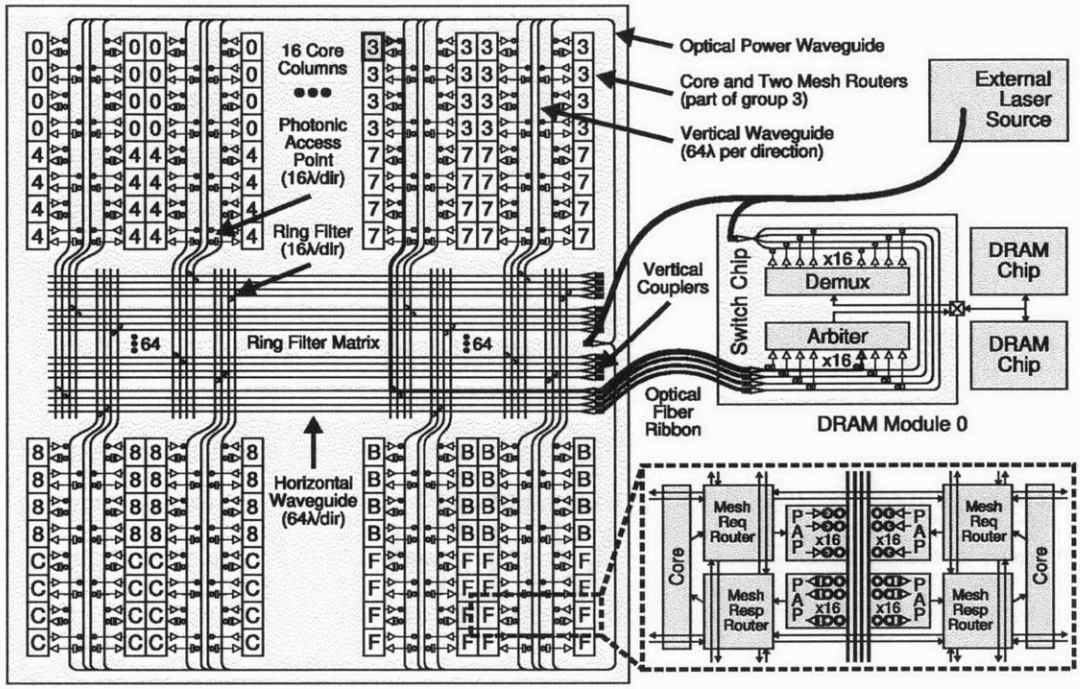


Figure 1-6: A 256-Core integrated photonic system [4]

Figure 1-6 diagrams a more sophisticated 256-core integrated photonic system based on the conceptual DWDM link presented earlier. DWDM is heavily leveraged for high density; now a single waveguide will be carrying approximately 64 wavelengths per direction for a total of 128 λ, achieving over 100x improvement in bandwidth density compared to an off-chip electrical link and over 10x compared to an on-chip electrical link. From a system architecture standpoint, the cores have been combined into groups of 16, with each core and its corresponding two mesh routers displaying a label indicating its group.

64 unmodulated wavelengths of light from an external laser source couple onto the

chip. After the light passes through a splitter, the ring modulator on the emphasized core modulates the light on a particular λ. This modulated light continues down the chip until it meets the ring filter matrix. A ring filter drops the λ onto a perpendicular waveguide, and light exits the chip onto an optical fiber ribbon, where it enters the off-chip DRAM module. It is then filtered into a photo diode and the data reaches the electronic arbiter module. The return path of the data is identical. The total area overhead of the photonic structures in this design is 7% of the active chip area, which is feasible and tolerable. There are approximately 200 waveguides and 40,000 rings, with a total throughput of 40 Tb/s.

To create such a system, photonics must be tightly integrated with electronics, which is not a trivial problem.

1.2.2 Different Approaches to Silicon Photonics

Silicon photonics has piqued the interest of several major companies and research institutions. There are several competing approaches toward optical integration with electronics. Each approach has its own strengths and drawbacks. A table containing a summary of these methodologies follows on the next page.

Silicon Photonics Methodology Comparison

Methodology	CMOS Process Modifications	Electrical-Only Performance	Parasitic Capacitance	Thermal Management	Power Delivery	Post-Processing
Discrete Integration	None	Unhindered	Very high; severely limits performance	Easily mitigated	No issue	None
Hybrid Integration; Package Interposer	None	Unhindered	High	Easily mitigated	No issue	None
Hybrid Integration; Die-to-Wafer	Microbump capability required	Unhindered	Moderate	Problematic due to second die	Potentially problematic due to second die	None
Hybrid Integration; Wafer-to-Wafer	TSV required	Unhindered	Moderately low	Problematic due to second optics die	Potentially problematic due to second die	None
Monolithic Integration; Thick-BOX SOI	None*	Slow	Very low	Easily mitigated	No issue	None
Monolithic Integration; Thin-BOX SOI	None*	Fast	Very low	Easily mitigated	No issue	XeF_2 localized etch
Monolithic Integration; Bulk-CMOS	Necessary process module addition	Slow (Memory-appropriate)	Low	Easily mitigated	No issue	None

* In this table, "none" implies that it is possible (and has been demonstrated) that certain processes have been used to create monolithic silicon photonics without process changes; it should not be implied that all processes share this capability.

Discrete Integration

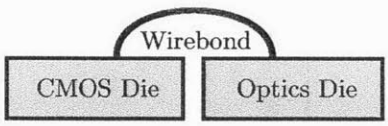


Figure 1-7: Discrete integration

The simplest approach is to have discrete optical components individually wirebonded to a traditional CMOS integrated circuit. While this approach does not require any sophisticated integration with a CMOS process, ultimately this approach has a number of drawbacks preventing its adoption in large systems. The additional capacitance added by the pad and bondwire between the driver circuit and optical device will be on the order of hundreds of femptofarads (for instance 320 fF in [14]), limiting the speed and energy-efficiency of the link. The lack of integration will also cause significant area overheads, making the resulting bandwidth and energy-efficiency inappropriate for processor-to-memory links. The added cost of having two discrete components has economic implementations in large systems as well.

Hybrid Integration: Multi-Die Packaging

Other research and industry efforts have focused on hybrid integration, where there is more sophistication in the interface between an electronics-only CMOS die and a separate optics-enabled die. There are several hybrid approaches.

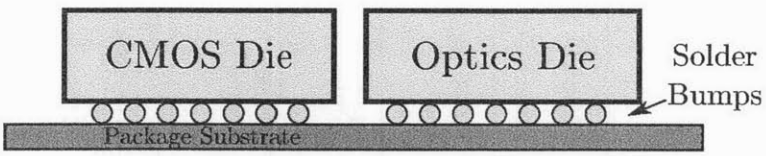


Figure 1-8: Hybrid integration examples.

One approach is to flip-chip bond both die to a more sophisticated package substrate [41], as shown in Figure 1-10. This more multi-die package is sometimes referred to as an interposer. Intel has done this by combining 90 nm CMOS transceiver circuits

with VCSEL and detector arrays with polymer waveguides through an interposer with optical capability, integrating CMOS with discrete optical components [40].

This approach is a stepping stone between fully discrete components and more sophisticated integration. The primary downsides are that it requires three components - a CMOS die, an optics die, and a sophisticated custom package. Although parasitic capacitances are improved over the discrete case, there are still higher parasitics than would exist in a fully monolithic solution. For this approach, it would be easy to attach a heatsink to the electronics-only die to mitigate thermal issues under heavy load.

Hybrid Integration: Die-to-Die, Die-to-Wafer, and Wafer-to-Wafer Bonding

Some works are focusing on integrated silicon photonics through microbump bonding of an electronics-only die to a optics-enabled die [16].

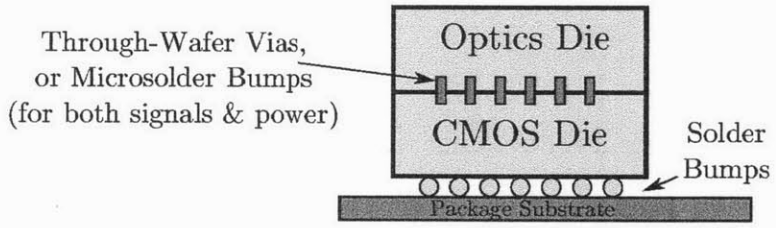


Figure 1-9: Hybrid integration examples: Die-to-die bonding.

Oracle has also shown interest in hybrid integration, integrating a 130 nm SOI photonics chip with an electronic 40 nm CMOS chip through flip-chip integration [43]. The 130 nm SOI si-Photonic chips are provided by a collaboration with Luxtera. Micro solder bumps provide the electrical connection between the two chips. They have reported a modulator power consumption of 400 fJ/bit at 5 Gbps with a BER of under 10^{-15} [42].

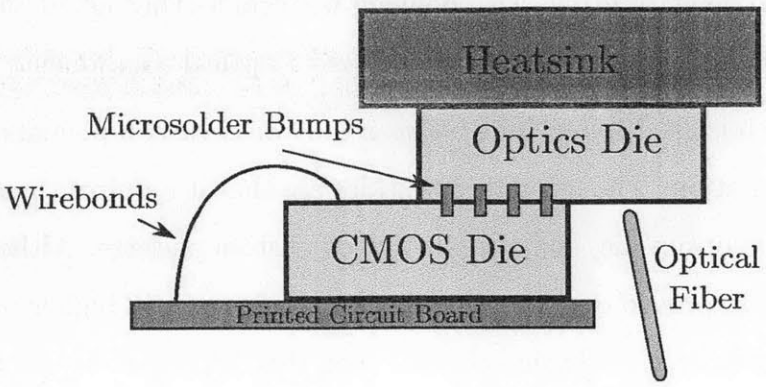


Figure 1-10: Hybrid integration examples: Oracle's die-to-die solution [?].

Other groups use wafer-level bonding with through-silicon-vias (TSVs) to achieve a similar end result. The wafer-to-wafer via capacitance in these approaches can be sub-10 fF, allowing for good energy efficiency. Michael Watts, a reader to this work, leads a research team using this wafer-level bonding method to explore low-power modulation through novel modulator devices such as microdisc modulators and adiabatic resonant microrings [38].

Our research has gravitated away from hybrid integration because of the commercial challenges involved in adopting the technology. For each component, two dice must be manufactured and then attached together, adding overhead cost and manufacturing complexity.

Monolithic Integration

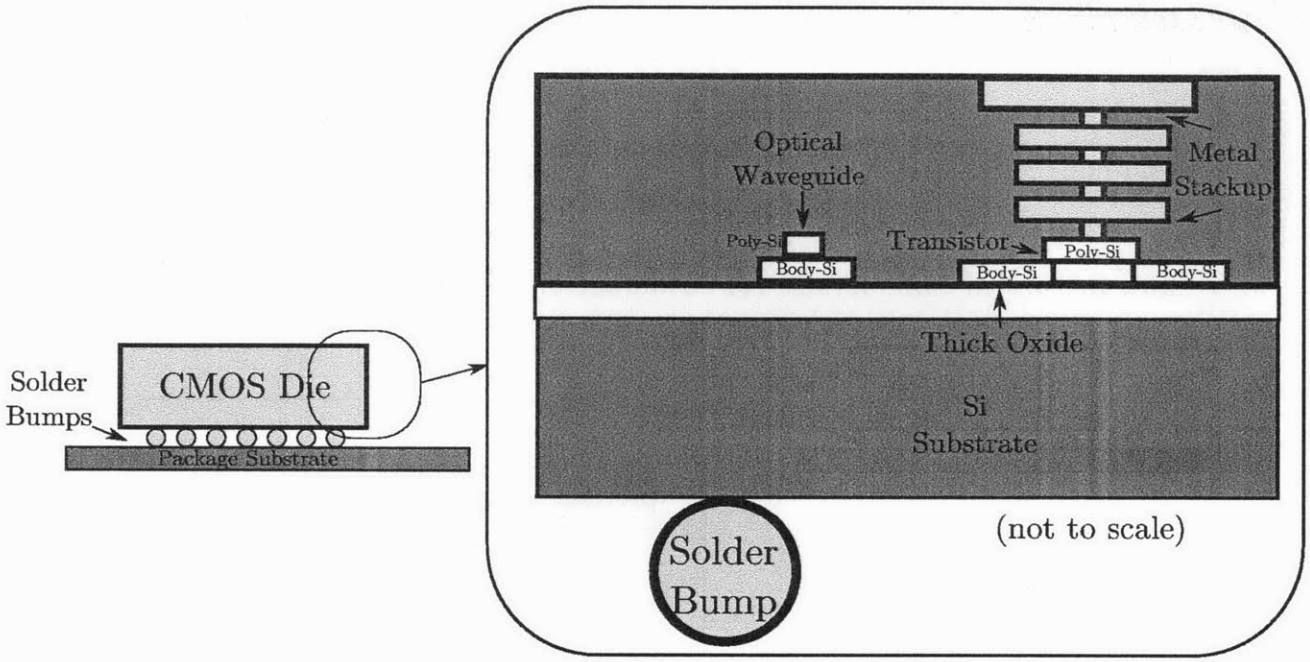


Figure 1-11: Monolithic integration with thick-BOX SOI.

Monolithically integrated silicon photonics has several advantages over other integration approaches. Because the electronic driving circuitry is immediately adjacent to the optical devices in the same process layers, parasitic capacitances are extremely low, offering very high energy efficiency. Commercially, this approach is very attractive as well, because only one die must be manufactured, keeping the component cost low. In addition, monolithic integration also reduces the area overhead of interfacing electrical and optical components [4]. All aforementioned hybrid approaches, whether with a package interposer or die-to-die/die-to-wafer bonding, will suffer from increased parasitics over a truly monolithic solution.

Luxtera has published work on monolithically-integrated silicon photonics, using a CMOS process with thick-BOX SOI to prevent the waveguide loss from being intolerably high [20]. This demonstration of integrated silicon photonics [3] [24] relies on specialized thick buried oxide SOI processes. These processes are used because the thick buried oxide provides a strong index contrast on the bottom of all optical

components, ensuring that the optical mode does not leak into the substrate. However, these processes are not suitable for processor-to-memory applications due to the thermal isolation properties of thick oxide and the effect on transistor electrostatics. Additionally, the technology also requires integration of 100% epi-Ge in process front-end, complicating the FET source and drain anneal and threshold control doping steps. Moreover, the mainstream VLSI technology is bulk-CMOS, which does not have any buried oxide layers. An advantage of the thick-BOX SOI approach is that thermal issues are easily mitigated by the traditional method of attaching a heatsink.

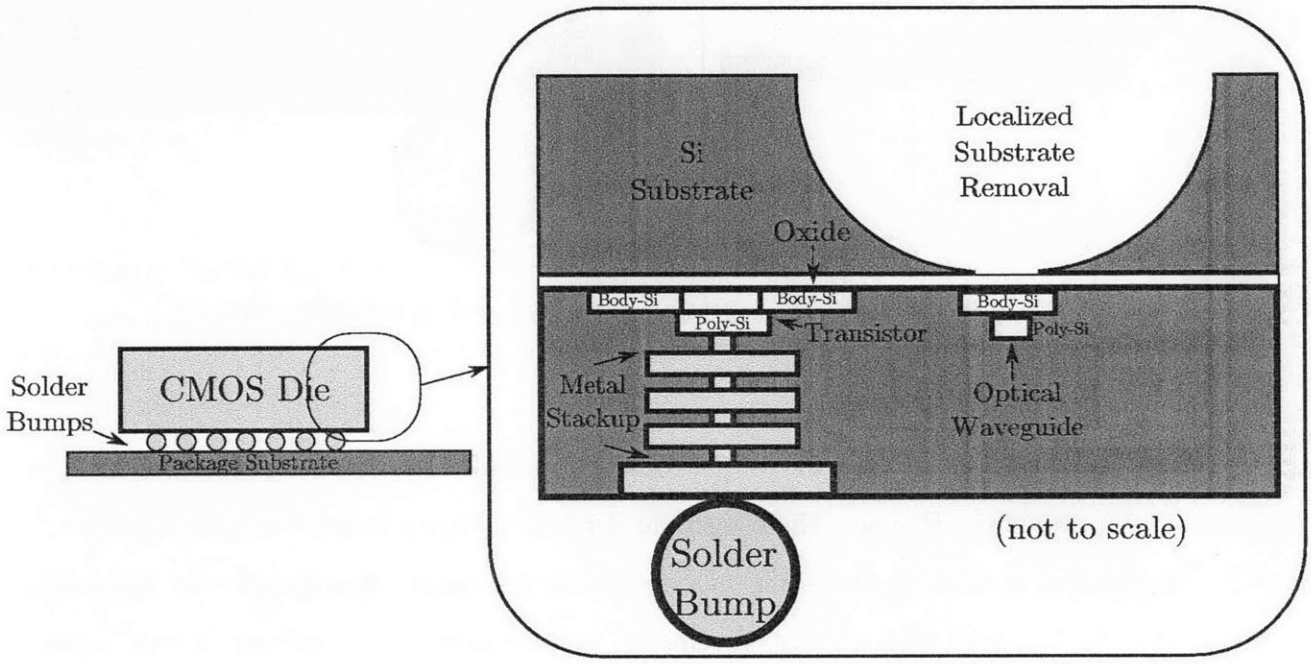


Figure 1-12: Monolithic integration (this work).

While other silicon photonics efforts have been based on hybrid integration, or monolithic approaches with thick-BOX SOI, our research is focusing primarily on monolithically integrated silicon photonics in high-performance processes. Adding extensive process customizations to mainstream processes can be prohibitively expensive and raises the barrier for adoption, especially considering that the technology is still in developmental stages which would make it a riskier investment. Our research group has participated in the development of a general method of enabling silicon photonics in thin-BOX-SOI without changes at the process foundry, which is

a significant step forward for the technology [26].

To demonstrate the feasibility of monolithically-integrated silicon photonics, we have created technology development platforms in both thin-BOX-SOI and bulk-CMOS processes.

1.3 Monolithic Integration Technology Development Platforms in Commercial Processes

We have created technology development platforms in both commercial high-performance processes (notably, IBM 45nm 12SOI) and a DRAM flash-periphery process (Micron 0.18 μm bulk-CMOS) to demonstrate the feasibility of a processor-to-memory optical link. The technology development platforms we developed serve two main purposes: to characterize optical properties of the photonic devices in bulk-CMOS and thin-BOX-SOI, and to demonstrate a full electrical-to-optical-to-electrical link circuit in a processor-to-memory environment. A motivation for this work is the understanding of the required interplay between circuit and device design and monolithic integration for these technology development platforms.

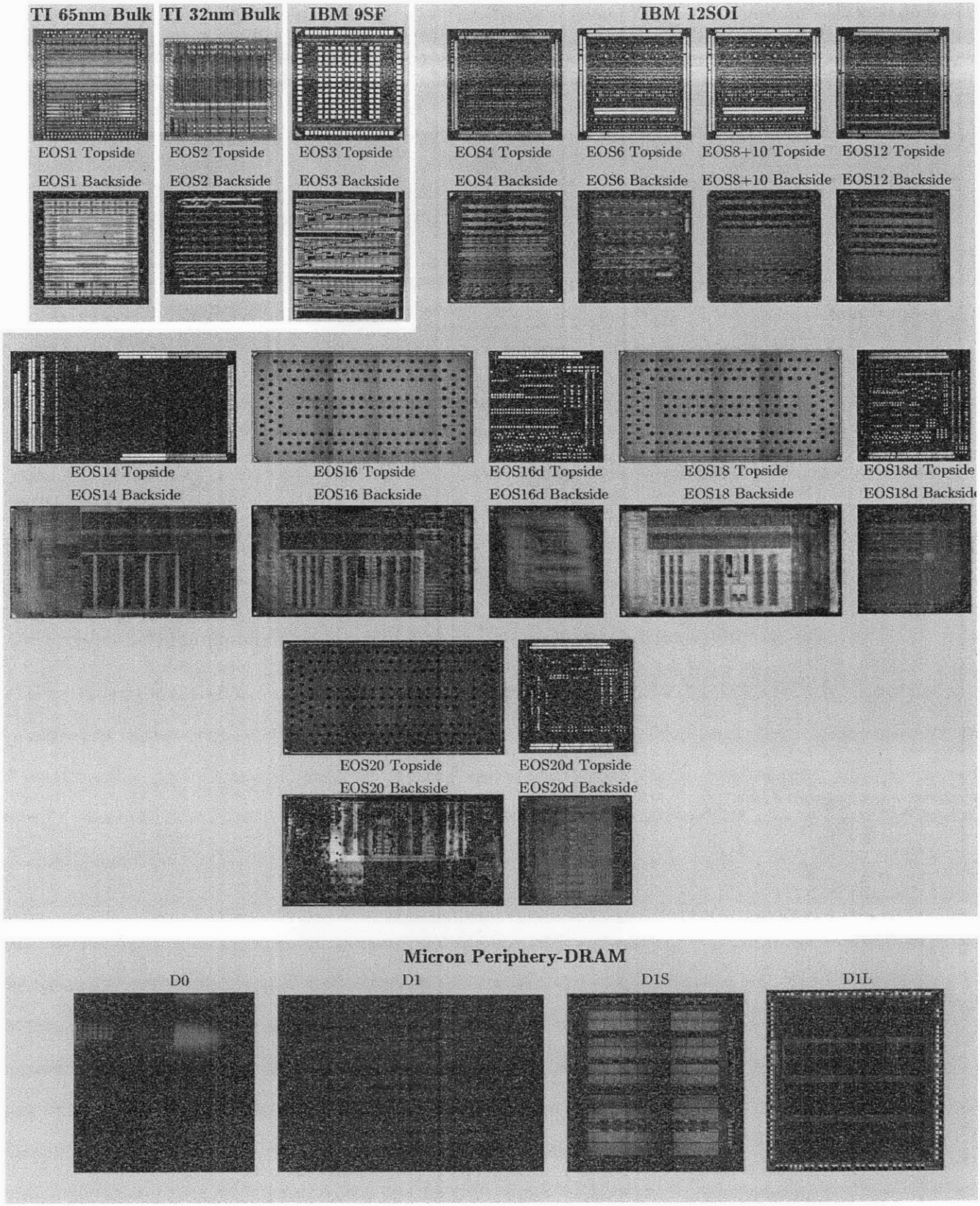


Figure 1-13: Technology Development Platforms in both Processor and Memory Processes

1.3.1 Processor-Side

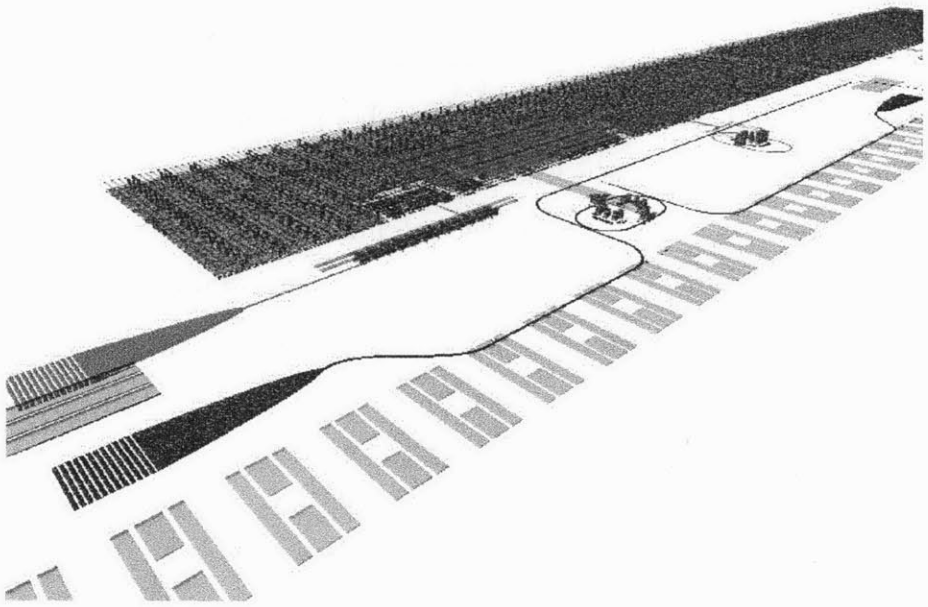


Figure 1-14: EOS: A monolithically-integrated photonic technology development platform in CMOS. A cell in a multi-λ WDM link is shown. This figure is a 3-D ray-traced representation of the actual taped-out GDS electronic-photonic circuit layout.

To demonstrate the processor end of a processor-to-memory link, we have created a technology development platform named "EOS" in a commercial IBM 45nm thin-SOI process (12SOI) to demonstrate the feasibility of zero-change monolithically-integrated photonics in a process suitable for creating high-performance processors (EOSn refers to the n^{th} generation design). This is the same process used for the IBM Cell and Power7 processors, famously powering the Playstation 3, Nintendo Wii U, and IBM's Jeopardy-winning Watson computer system. Our technology platform development die are manufactured without making any changes to the native process; we are a regular fab customer and our designs are aggregated on a multi-project wafer, sharing the reticle with other ASIC designers. The designs adhere to the process design rules, with only minor DRC waivers required. Optical waveguides and devices are made using the poly-Si layer and body-Si layers, which traditionally are intended to create transistor gates, sources and drains. Figure 1-14 shows a scale ren-

dering of the technology development platform, displaying photonic structures such as vertical grating couplers and ring modulators alongside custom analog circuitry and traditional placed-and-routed digital electronics.

To enable optics in the IBM 45nm process, a postprocessing step is necessary. The optical devices and waveguides made using body-Si and poly-Si are sitting on top of the thin silicon dioxide, which does not sufficiently confine the optical mode. After we receive the die from the foundry, we must do an additional coarse etch step, releasing the substrate underneath the optical structures using Xenon Diflouride (XeF_2) gas, which selectively etches silicon isotropically. In an SOI process such as IBM 45nm, the buried oxide provides a natural etch-stop for the XeF_2, which etches silicon preferentially to oxide at a rate of 1000:1. The XeF_2 can eat away the substrate from the back-side of the die. [1] The etch time must be carefully controlled so that the XeF_2 does not over-etch and destroy circuits, as happened in Figure 4-1. After etching, the resulting air gap underneath the optical devices provides the index contrast necessary to achieve optical mode confinement. While this etch step does add cost and complication to final assembly, this complication is on par with normal packaging considerations such as wire- or flip-chip bonding. Companies such as Akustica [9] have used a similar XeF_2 localized wafer-level release process for CMOS MEMs microphones; however, with this work, we are performing the release at the die-level.

1.3.2 Memory-Side

To demonstrate the memory end of a processor-to-memory link, we have developed a similar monolithic technology development platform in a modified flash periphery 0.18μm bulk-CMOS process with no Germanium. Our research group collaborated with Micron Technology to develop this process. In developing this technology platform, we were able to make significant process changes so that the resulting optics have low loss and high performance. Unlike high-performance processes such as IBM

[1] In a bulk-CMOS process, etch holes must be created through the STI to expose the XeF_2 gas to the substrate, from the front-side of the die [12].

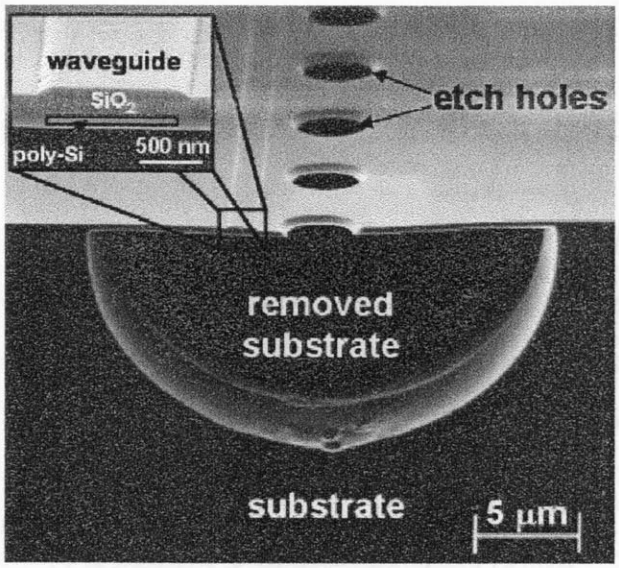


Figure 1-15: Substrate undercut using XeF_2 in a bulk-CMOS process [12].

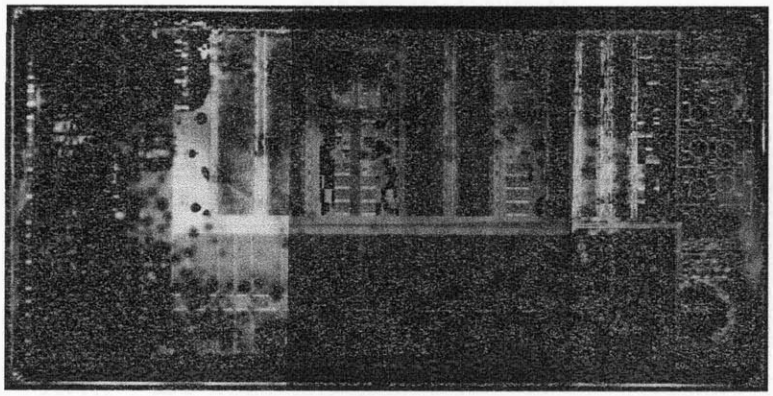


Figure 1-16: An example of XeF_2 over-etch causing catastrophic circuit damage.

12SOI, commercial DRAM processes are fine-tuned to mass-produce a single component. Making process changes in this type of specialized application is very common, and is also commercially viable. The photonics aspect of the project was implemented as a drop-in process module, adding an additional five mask sets.

Fortunately, because the research team was able to tightly collaborate with Micron, no post-processing is necessary on the Micron-platform chips, significantly simplifying packaging.

1.3.3 Silicon Photonics Area and Energy Advantage

To system designers developing the next generation of technology, a fair comparison is necessary to determine if the apparent advantages of an optical system will outshine the predicted performance of upcoming electrical links. Integrated silicon photonics microprocessors are still years away from becoming mainstream, so comparing a projected photonic system to the current state-of-the-art electrical case would be unfair. For this reason the full system from Figure 1-6 is compared to projected electrical links in the 22 nm CMOS process node, which is now state-of-the-art.

An analysis of the optical power budget for this full system, shown in Table A.1, demonstrates the importance of optimizing the designs of critical optical structures. Summing the various losses along the datapath in the unoptimized case and taking into account the estimated receiver sensitivity yields an outrageous estimated external laser power of 3.3 kW. Optimizing the designs to improve crossing loss, on-chip waveguide loss, and drop loss reduces the necessary external laser power to a more manageable 6.38 W. Notice that the modulator insertion loss, while not one of the largest sources of loss, is not to be neglected. In this link budget, the modulator carries the biggest footprint and must be carefully optimized. The modulator sets the extinction ratio and insertion loss, trading off the electrical modulation energy for optical energy required in the link. This tradeoff motivates the deeper analysis in the following chapter.

Metric	Energy (pJ/b)	Bandwidth Density (Gb/s/μm)
Global on-chip photonic link	0.25	160-320
Global on-chip optimally repeated electrical link	1	5
Off-chip photonic link (50 μm coupler pitch)	0.25	13-26
Off-chip electrical SERDES (100 μm pitch)	5	0.1
On-chip/off-chip seamless photonic link	0.25	

Table 1.1: Area and Energy Comparison of Electrical and Optical Systems

Component	Preliminary Design	Power Loss	Optimized Design	Power Loss
Coupler loss	1 dB/coupler	3 dB	1 dB/coupler	3dB
Splitter loss	0.2 dB/split	1 dB	0.2 dB/split	1dB
Non-linearity	1 dB	1 dB	1 dB	1 dB
Through loss	0.01 dB/ring	3.17 dB	0.01 dB/ring	3.17 dB
Modulator insertion loss	1dB	1dB	0.5 dB	0.5 dB
Crossing loss	0.2 dB/crossing	**12.8 dB**	0.05 dB/crossing	**3.2 dB**
On-chip waveguide loss	5 dB/cm	**20 dB**	1 dB/cm	**4 dB**
Off-chip waveguide loss	0.5×10^{-5} dB/cm	0 dB	0.5×10^{-5} dB/cm	0 dB
Drop loss	2.5 dB/drop	**5 dB**	1.5 dB/drop	**3 dB**
Photodetector loss	0.1 dB	0.1 dB	0.1 dB	0.1 dB
Receiver sensitivity	-20 dBm	-20 dBm	-20 dBm	-20 dBm
Power per wavelength		26.07 dBm		-1.03 dBm
		(0.40 W)		(0.78 mW)
Power required at source		**3.3 kW**		**6.38 W** (100 mW/λ)

Table 1.2: Optical Power Budget [4]

Component	Latency
Serializer/Deserializer (50 ps each)	50 ps
Modulator driver latency	108 ps
Through latency (2.5 ps/adjacent channel)	7.5 ps
Drop latency (20 ps/drop)	60 ps
Waveguide latency (106.7 ps/cm)	427 ps
SM fiber latency (48.3 ps/cm)	483 ps
Photodetector + TIA latency	200 ps
Total latency	**1.385 ns**

Table 1.3: Optical Data Transmission Latency

The analysis, shown in Table 1.1, suggests that the total energy overhead for the various electrical back-end components of a global on-chip photonic link will be less than 250 fJ/b, which is fivefold more efficient than a global on-chip optimally repeated electrical link. This electrical backend is shown in Figure 1-17. To achieve the target energy overhead of 250 fJ/b, an optically-distributed clock is necessary to avoid the energy cost of a phase-locked loop (PLL). This optical energy overhead is also is 1-2 orders of magnitude less than state-of-the-art photonic devices [24] (not displayed in the table) because the monolithically integrated approach lowers the capacitance of

photonic structures.

The anticipated bandwidth density of photonics is 32-64x higher for the on-chip case. Similar gains are seen for the off-chip case, with photonics projecting a 25x energy improvement and a 130-260x improvement in bandwidth density. Furthermore, as shown in Figure 1-5, the on-chip and off-chip photonic links can be unified with a single energy cost, only further justifying the development of this technology. It is important to note that external laser power is not factored into this energy estimate because modern processors are on-chip power-density limited, and the laser will be powered from an external supply.

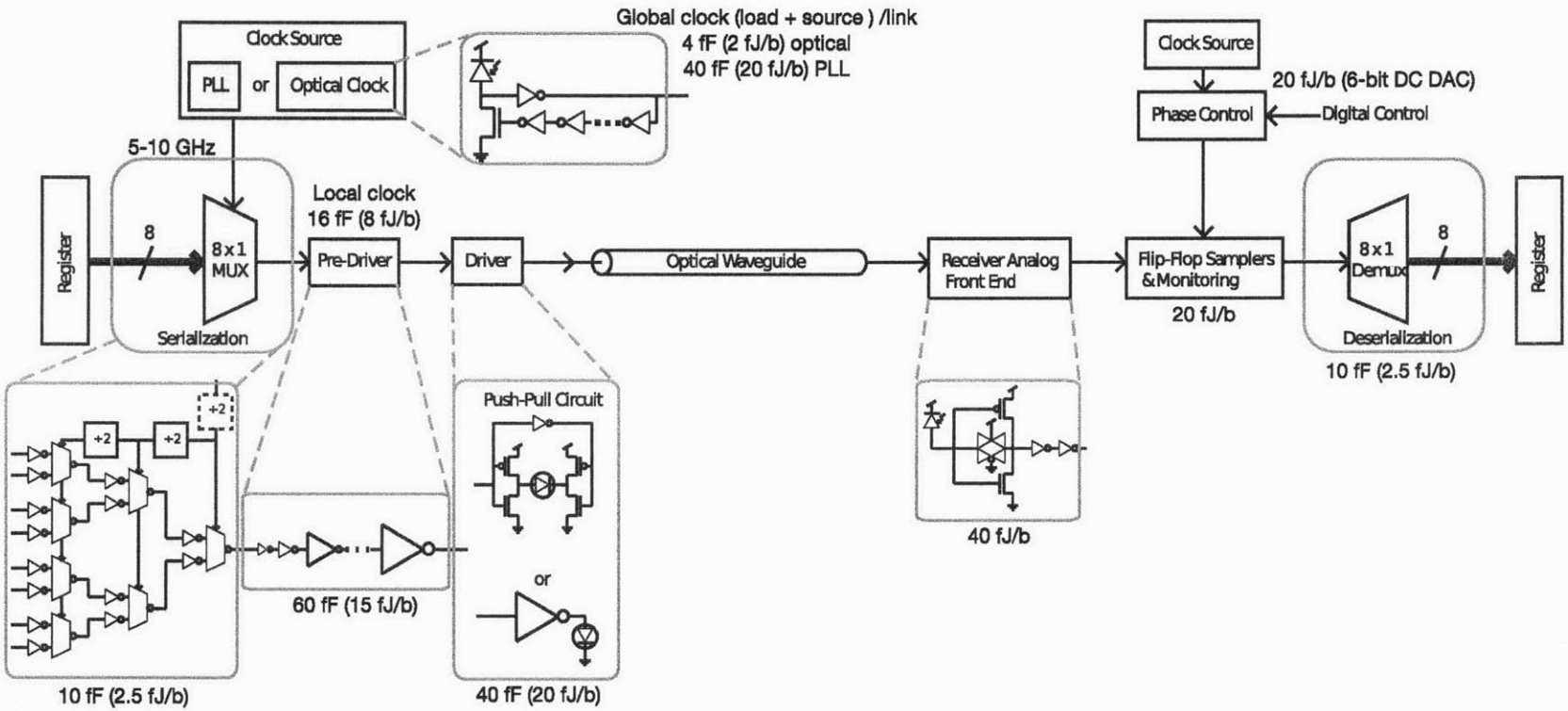


Figure 1-17: Electrical backend of a photonic link

A brief analysis of the latency of the full photonic system (shown in Table A.2) indicates that the latency of electrical and photonic systems is comparable.

The result of these analyses suggests that an integrated photonic system will need to be optimized from multiple layers (system, circuit and device) to become competitive with electrical links. The modulator is one of the most fundamental parts of this system and is critical both from an energy efficiency and bandwidth perspective, which is the motivation for this thesis.

1.4 Thesis Focus

As multi-core systems become more prevalent, new technologies will be needed to solve the bandwidth problem. Silicon photonics is a budding technology with the potential to have a large impact on the computing market if it can be economically integrated with current CMOS processes. While previous efforts in the field have relied on esoteric processing techniques such as thick-oxide SOI, our MIT team has developed a method to enable silicon photonics in commercial bulk and SOI CMOS. We have described two technology development platforms to demonstrate the feasibility of this technology.

Several building blocks are necessary to create a working integrated optical link. The focus of this work is on the simulation, design, and testing of one of the most critical system components, the electrical-to-optical modulator and its driver circuit. Traditionally this device has been a speed, area, and power limiter.

My specific contributions began with an exploration of two modulator driver circuit design topologies, including circuit design, fabrication, and testing. One circuit topology, designed to drive a carrier-injection-based ring modulator, uses a digital push-pull topology with pre-emphasis to inject charge into the modulator device's P-I-N diode structure. The use of pre-emphasis greatly increases the operating speed of the modulator device while lowering power consumption, reducing the energy per bit of the modulator and preventing the ring's optical passband from shifting to the next optical channel. The driver circuit is designed to inject the right amount of

charge quickly into the P-I-N diode structure by using pre-emphasis to take advantage of the nonlinear optical transfer function. The second circuit topology drives a reverse-biased depletion-mode modulator device through a novel voltage-boosting charge pump. A flexible driver circuit is necessary because some critical device parameters, such as the diode's carrier lifetime, are poorly controlled by the CMOS foundry and are unknown at design time. For both circuits, the link between optics and electronics is bridged by analyzing the charge injection or depletion necessary to achieve a target optical extinction ratio. For best link performance, it is essential to optimize the device and circuit jointly, hence the need for the monolithic technology development platform where such experiments can be conducted while both circuits and devices are calibrated and adjusted to the standard CMOS process. My contributions also included a Verilog-A model for optical/electrical co-simulation of the modulator driver circuit and device (outlined in Chapter 2), as well as packaging solutions to enable the first successful demonstrations. These contributions allowed the research to progress from a single modulator to transceiver-scale demonstrations.

Chapter 2

The Modulator Device

The modulator is a fundamental building block of any electrical-to-optical system. In contrast to passive photonic devices such as vertical couplers, splitters, and ring filters, the modulator is an active device which can electrically alter the amplitude of light flowing through it at a particular wavelength-channel.

For the smallest area and best energy efficiency, we have chosen to use a ring modulator over other modulator topologies. Alternatives such as a Mach-Zehnder interferometer require many times more area and as a result have far greater parasitic capacitance, limiting energy efficiency and bandwidth density.

The ring modulator is a tunable resonant ring filter. Light that can traverse the ring filter in an integer number of wavelengths will constructively interfere with itself and stay in the ring; otherwise, the light will pass through the ring filter. The resulting optical frequency response of this filter follows a Lorentzian distribution, shown in Figure 2-1 [30]. Light at the target optical channel is modulated by shifting the stopband of the resonant filter in and out of the optical channel. Often, for a maximum extinction ratio, the optical channel frequency f_0 will not be centered directly at the ring's passive resonance f_s, but slightly off to one side where the slope of the Lorentzian is highest. However, for different extinction ratio vs. insertion loss tradeoffs, the optical channel may be positioned at a different spot on the Lorentzian.

As the stopband shifts by some Δf (shown as $f_s - f_n = -20$ GHz in the figure) for a '1'-bit, the transmissivity at the channel center rises from T_0 to T_1. The extinction

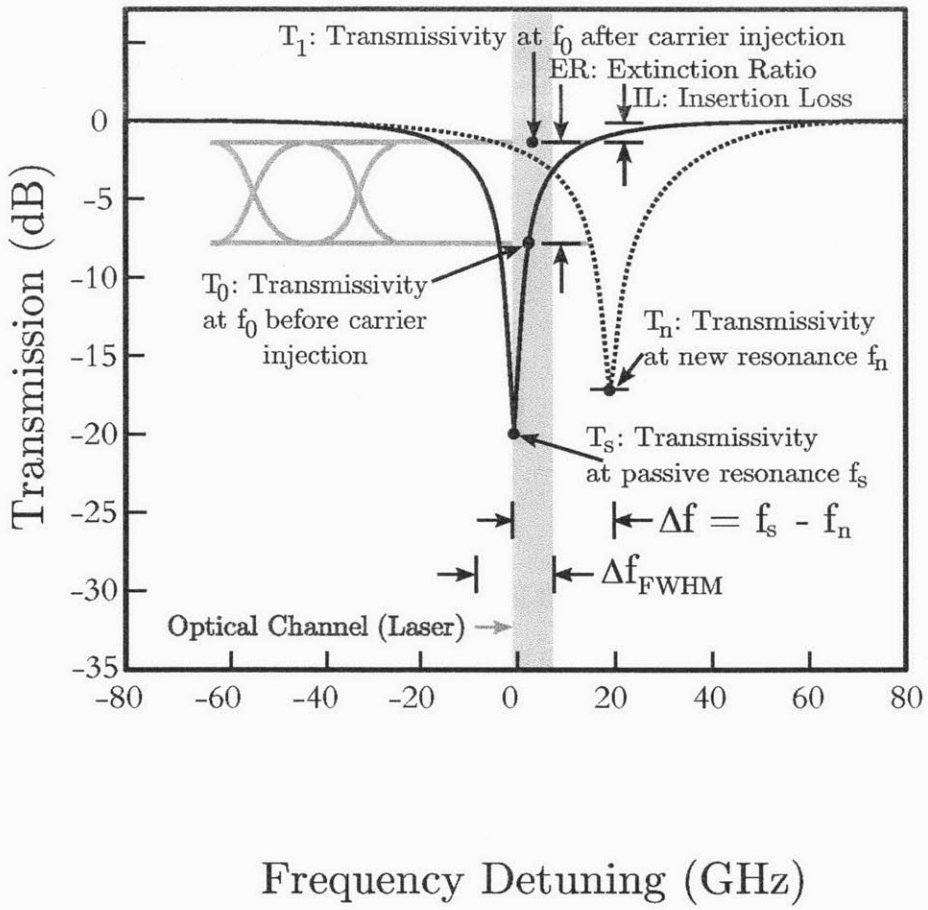


Figure 2-1: Ring modulator Lorentzian frequency response, both passive and under carrier-injection (under carrier-depletion, T_n would shift to the right instead of the left).

ratio of the modulator is defined as $ER = T_1/T_0$ (or, in decibels, $ER_{dB} = T_{1dB} - T_{0dB}$). The insertion loss, IL, is the reciprocal of the shifted-state transmissivity, so $IL = 1/T_1$ (in decibels, $-T_{1dB}$).

The stopband is shifted by changing the refractive index of the ring's waveguide. The modulator's waveguides are fabricated in the body-Si layer, the poly-Si layer, or a combination of both (in the case of a ridge waveguide). The waveguide is doped as little as possible and is unsilicided to preserve its optical properties. Changing the refractive index of the waveguide can be accomplished by several different mechanisms, of which carrier injection and thermal heating are the most dominant effects. The pioneering work of Soref [33] demonstrated that the free-carrier plasma dispersion ef-

fect can change the refractive index of silicon (shown later in Figure 2-5). This effect was originally used to provide a phase shift in branches of Mach-Zehnder modulators [24] and more recently to change the resonance of ring modulators [39] such as the ones used in this work [26]. There are two common mechanisms for changing the free carrier concentration within the ring: by forward-biasing a diode structure to inject charge, and by reverse-biasing a diode structure to deplete charge.

A system-level tradeoff is evident from Figure 2-1: a larger Δf results in a higher extinction ratio and a lower insertion loss, but comes at a higher energy cost for the modulator driver circuit. Depending on the topology, there may also be a speed penalty for shifting a larger Δf. For a given modulator device, a large extinction ratio requires less optical laser power for a given receiver sensitivity, and results in less insertion loss. However, a large extinction ratio requires higher charge injection and thus higher current and power, making it electrically more expensive and slower. A lower extinction ratio is electrically cheaper but requires higher optical laser power. A reasonable minimum extinction ratio target is 3 dB because this is a twofold change in optical power between a '1'-bit and a '0'-bit, however, the tradeoff between optical laser power and modulator power is only apppropriate to approach at the system architecture level rather than at the device level.

In addition to charge concentration, ring resonance is also sensitive to temperature. Although far too slow for data transmissions, this effect is useful to adjust the ring's resonance to account for process variation. Polysilicon heaters are designed in-plane with the ring modulators for improved thermal efficiency to both heat the rings and serve as temperature sensors. Compact analog-to-digital (ADCs) and digital-to-analog converters (DACs) can be used to digitize the temperature information.

Two topologies of ring modulators will be covered; first, a carrier-injection device is described, then a carrier-depletion device is compared.

2.1 Carrier-Injection Modulator Device Analysis

The first type of modulator device to be considered is a resonant ring filter with a stopband that shifts due to charge injection.

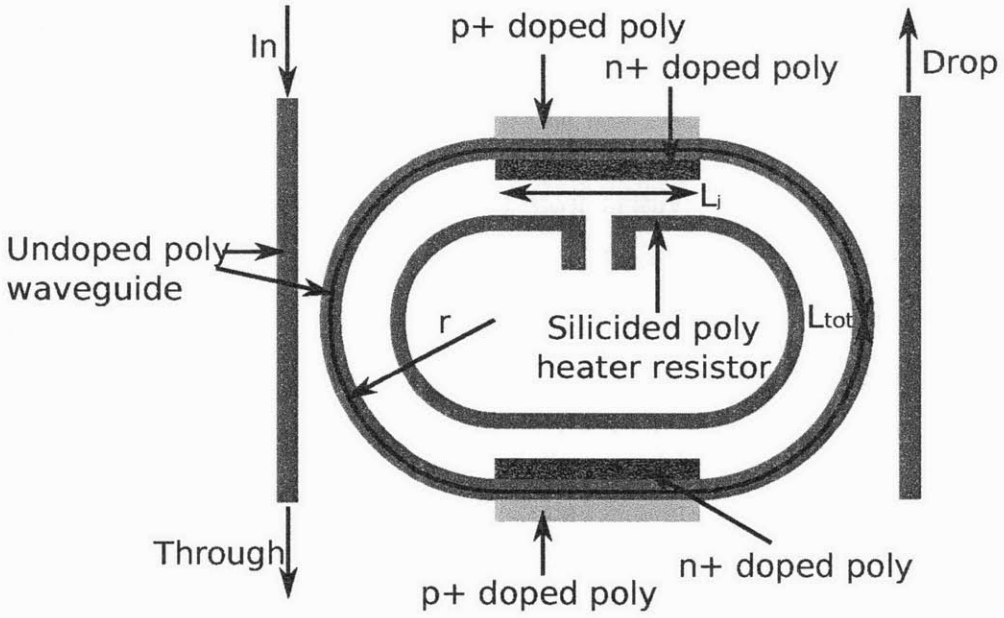


Figure 2-2: Carrier-injection ring modulator diagram top view (not to scale)

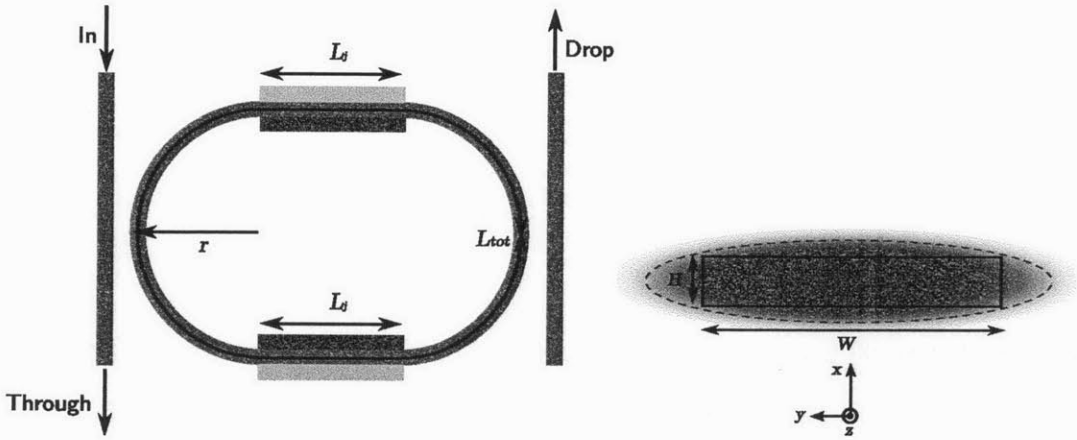
A top-view diagram of a carrier-injection ring modulator is shown in Figure 2-2. Unmodulated light at the target optical channel λ_0 travels down the waveguide until it reaches the In port of the modulator. Because λ_0 is at or near the ring's resonant wavelength, the majority of the light couples into the ring and travels counter-clockwise until it exits the Drop port. The other optical channels pass to the Through port unaffected (aside from a small amount of loss). For an optical channel centered at an ideal modulator's resonance with a perfect extinction ratio, no light at λ_0 would continue down the Through port. A certain amount of time is necessary for the resonance to build up inside the ring, but the ring's optical response is fast in comparison to its electrical response and can be neglected for the sub-10 Gb/s transmit speeds in this work, because the loaded $\mathbb{Q}$-factor of the rings in this process range between 4 and 15k, resulting in optical bandwidths of 15-60 GHz.

To easily inject large amounts of charge into the ring's waveguide, a P-I-N diode

structure is created, with highly p+ and n+ doped regions on either side of the waveguide, and with the intrinsic silicon waveguide acting as the diode's intrinsic region. The P-I-N diode is a meeting point between the circuit and device designers. Both need accurate models of the diode to simulate and verify the behavior of the device. Due to the high-injection mode of operation, accurate device simulations and sophisticated circuit equivalent models are necessary to model and analyze the numerous design tradeoffs. The development of a valid model across the optical, device, and circuit domains requires iteration between multiple simulators. A solid understanding of the interaction between the electrical and optical domains is necessary to simulate and design a working circuit to modulate light. The following sections of the thesis will develop an analytic method to approach this link between optics and electronics.

2.1.1 Carrier Injection Ring Modulator Dimensions

Some ring "racetrack" dimensions must first be defined, shown in Figure 2-3. The length of a single junction in the direction of optical propagation is L_j. There are two such junctions in the ring. The ring's radius in the curved section is r and its total circumference is $L_{tot} = 2L_j + 2\pi r$. The waveguide's height is H and its width is W. The volume of the waveguide diode junction is $V_j = 2 \cdot L_j \cdot W \cdot H$.



(a) Top view

(b) Waveguide cross-section with mode overlap concept

Figure 2-3: Carrier-injection ring modulator dimensions (heater ommitted)

Some terminology is necessary to clarify a sticking point of confusion regarding the diode dimensions. The junction length L_j (a parameter important in the optical domain) is equivalent to the diode width W_d (a parameter important in the circuit and device domains). Similarly, the waveguide width W is equivalent to the diode length L_d. This terminology preserves the intuition that a wider diode will draw more current at a given voltage.

There is also a design tradeoff involved in choosing the length of the diode junction, L_j, for a given ring circumference. A longer L_j will lower the series resistance of the diode, but it will also raise the junction capacitance. A longer L_j will also require more current from the circuit and increase the loss due to mode overlap with the highly doped regions of the diode and the metal contacts, decreasing the $\mathbb{Q}$-factor and thereby increasing the necessary charge required for a given extinction ratio.

2.1.2 Free Spectral Range

If the optical channel is centered at the ring's resonance prior to charge injection; i.e., if $f_0 = f_s$, then by definition the light at wavelength λ_0 traveling through the ring will travel an integer number of wavelengths m in one round trip L_{tot} [17], given by:

$$\frac{L_{tot}}{\lambda_0} = \frac{k_0 \cdot L_{tot}}{2\pi} = m \tag{2.1}$$

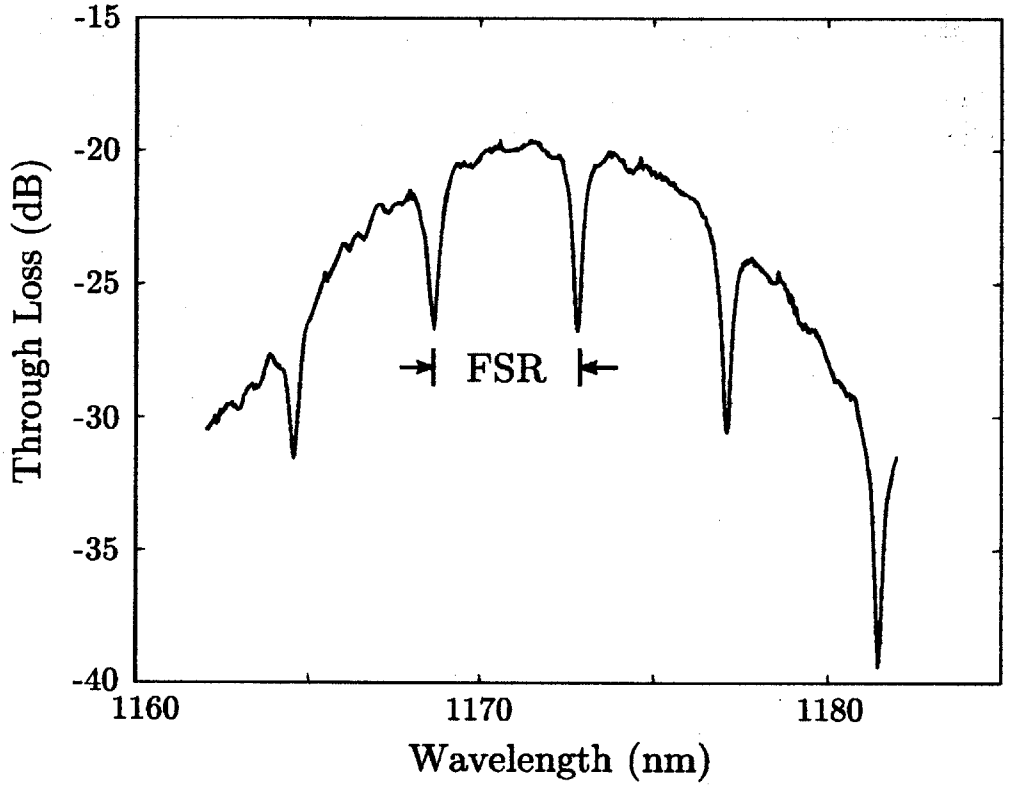


Figure 2-4: Measured EOS18 ring response; note the Lorenzian response repeating, for an FSR of 4 nm for this particular ring.

The wavenumber of the light is $k_0 = \frac{2\pi}{\lambda_0} = \frac{\omega_0}{c_s}$. The frequency spacing between the ring's resonances, the free spectral range (FSR), is determined by the frequency change required to shift the integer number of wavelengths by one. A measured ring response from EOS18 is shown in Figure 2-4, clearly showing an FSR of 4 nm for this particular ring.

$$\frac{k_{FSR} \cdot L_{tot}}{2\pi} = 1 \tag{2.2}$$

$$\frac{\omega_{FSR} \cdot L_{tot}}{c_s \cdot 2\pi} = 1 \tag{2.3}$$

$$\frac{FSR \cdot L_{tot}}{c_s} = 1 \tag{2.4}$$

$$FSR = \frac{c_s}{L_{tot}} = 1 \tag{2.5}$$

$$FSR = \frac{c}{L_{tot} \cdot n_g} \tag{2.6}$$

The speed of light in free space is c and the speed of light within the waveguide is $c_s = \frac{c}{n_g}$, n_g being the group index of the ring's waveguide. For light in the 1220 nm band, $n_g = 3.3$.

It is desirable to have a large FSR to densely pack as many waveguides as possible onto a single waveguide for high bandwidth density. System designers can also leverage the repeating resonances of the ring to account for process variation by heating the ring to tune the resonance to a fixed wavelength-channel. The repeating resonances are convenient because it will never be necessary to rely on intentionally cooling the ring.

Now that the ring has been evaluated in its passive state, the next step is to evaluate what happens under charge injection.

2.1.3 Free Carrier Refractive Index Change

As free electrons and holes are injected into the two I-regions of the carrier-injection modulator device, the index of refraction (n) in those regions will decrease. The index change (Δn) is sensitive to the change in both free electron concentration ΔN_e and free hole concentration ΔN_h.

$$\Delta n = \Delta n_e + \Delta n_h = n_{fe}\Delta N_e + n_{fh}(\Delta N_h)^{0.8} \tag{2.7}$$

Experimental data by Soref [33] has empirically determined the values of n_{fe} and n_{fh} at $\lambda = 1300\,\text{nm}$. The ring modulators we are evaluating in this example are designed for $\lambda = 1220\,\text{nm}$, so Soref's data must be extrapolated. The theoretical model for refraction due to free carriers is shown in Equation 2.8 and shows a dependence on λ^2, where q is the electronic charge, ϵ_0 is the permittivity of free space, n is

the refractive index of unperturbed silicon, $m_{ce}*$ is the conductivity effective mass of electrons, and $m_{ch}*$ is the conductivity effective mass of holes [33].

$$\Delta n = -(q^2\lambda^2/8\pi^2c^2\epsilon_0 n)[\Delta N_e/m_{ce}* + \Delta N_h/m_{ch}*] \tag{2.8}$$

The λ^2 dependence is used to adjust the $\lambda = 1300\,\text{nm}$ n_{fe} and n_{fh} coefficients to be appropriate for $\lambda = 1220\,\text{nm}$. The data from Figure 10 in [33] was transcribed, and each value of $-\Delta n$ was multiplied by $(1220\,\text{nm})^2/(1300\,\text{nm})^2$. A linear regression algorithm was used to find the new values of n_{fe} and n_{fh} at $\lambda = 1220\,\text{nm}$, shown in Figure 2-5[1].

[1]This procedure was initially performed by Milos Popović.

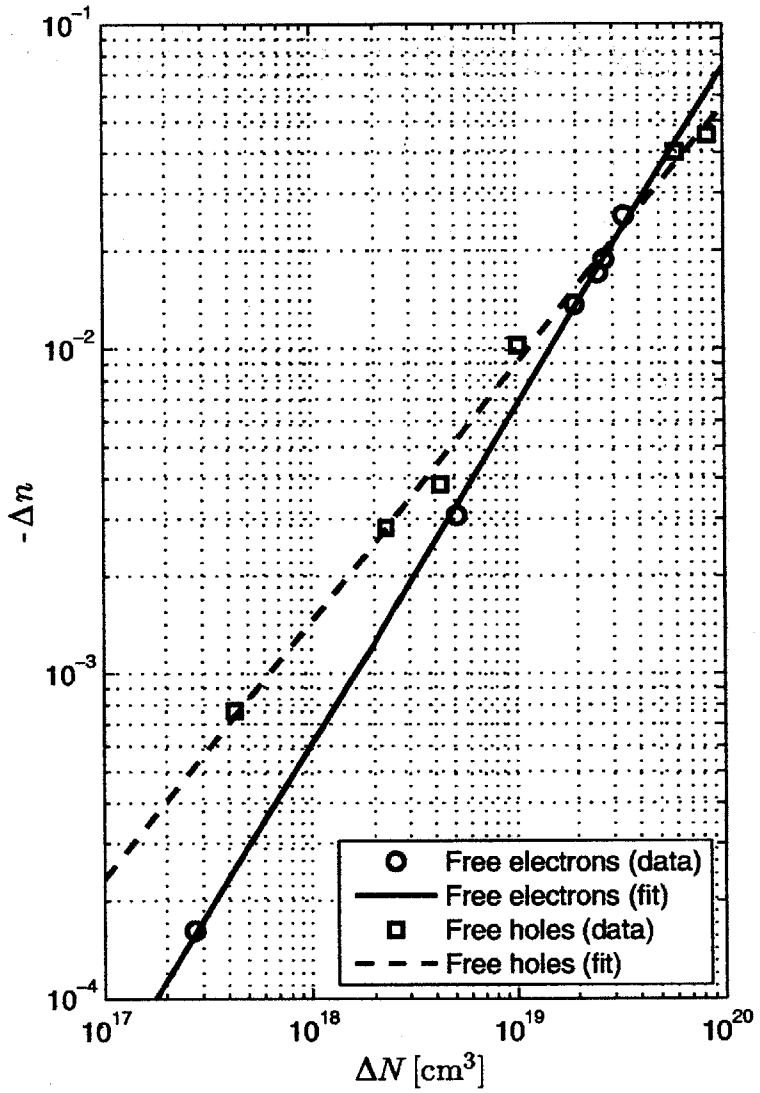


Figure 2-5: Silicon free carrier index change at $\lambda = 1220\,\text{nm}$[33]

As Soref mentions, Equation 2.8 is too simplistic to take into account the measured $(\Delta N_h)^{0.8}$ dependence in Equation 2.7, and there is also a slight exponential dependence on N_e as well which is neglected, so there is a degree of inaccuracy in the result. We use the experimental model of Equation 2.7 with the values of n_{fe} and n_{fh} in Equation 2.9 which are valid for $\lambda = 1220\,\text{nm}$ at injection levels of $\Delta N \approx 10^{19}\,\text{cm}^3$.

$$n_{fe} = 4.8 \times 10^{-22}\,\text{cm}^{-3}, \quad n_{fh} = 4.9 \times 10^{-18}\,\text{cm}^{-3} \tag{2.9}$$

The index change in the intrinsic region does not have a full effect on the optical

mode due to the mode overlapping the waveguide as shown in Figure 2-3b. This is because charge injected in the I-region does not fully overlap with the mode, but only with a portion. A correction coefficient $\Gamma = 0.43$ determined through optical simulations by the photonic device designers is added to correct for this effect.

$$\Delta n_{eff} = \Delta n \cdot \Gamma, \qquad \Gamma = 0.43 \tag{2.10}$$

2.1.4 Phase Shift Due to Refractive Index Change

The index change in the straight sections of the racetrack causes a phase shift in the light traveling in the ring, still at wavelength λ_0. The $\overline{E}$-field of single-mode light [34] through one straight section of the racetrack inside the I-region prior to charge injection can be described by a vector with the Cartesian coordinate system defined in Figure 2-3b and the light traveling in the $+\hat{x}$-direction down the waveguide.

$$\overline{E_0} = \hat{y}\underline{E_0}\sin(k_0 x)e^{-jk_0 L_j} \tag{2.11}$$

$\underline{E_0}$ is a phasor at the frequency of the light; the magnitude is unimportant because we only care about the phase of the light for this analysis. The $\overline{E}$-field of single-mode light through one straight section of the racetrack <u>after</u> charge injection is $\overline{E_1}$.

$$\overline{E_1} = \hat{y}\underline{E_0}\sin(k_1 x)e^{-jk_1 L_j} \tag{2.12}$$

$$\overline{E_1} = \hat{y}\underline{E_0}\sin(n_{eff}k_0 x)e^{-jn_{eff}k_0 L_j} \tag{2.13}$$

The index change is seen in the term $k_1 = k_0 n_{eff} = \frac{\omega_0 n_{eff}}{c_s}$. The phase change of the light $\Delta\phi'$ can be described by comparing the phase of these two cases in Equations 2.11 and 2.13 (neglecting any magnitude changes).

$$\Delta\phi' = (n_{eff}k_0 L_j) - (k_0 L_j) = \Delta n_{eff} \cdot k_0 \cdot L_j \tag{2.14}$$

The light passes two of these straight regions as it traverses the ring, each with

identical charge concentrations, which doubles the phase change.

$$\Delta\phi = 2 \cdot \Delta\phi' = \Delta n_{eff} \cdot 2L_j \cdot k_0 \tag{2.15}$$

This phase change leads to a resonance shift in the ring.

2.1.5 Resonance Shift

It is useful to know how much the resonant frequency of the ring has changed by this charge injection. Prior to charge injection, because the light was at the ring's resonant wavelength, an integer number of trips are made by the light around the waveguide as $k_0 \cdot L_{tot} = 2\pi m$. Because the ring's resonant wavelength has perturbed, this relation no longer holds for an integer number of wavelengths.

$$k_0 \cdot L_{tot} + \Delta\phi = 2\pi(m + \Delta m) \tag{2.16}$$

There are, however, the same m integer number of wavelengths in the ring for a slightly different wavelength of light λ_n, which is the ring's new resonant wavelength. Assume the ring has not shifted by a full free spectral range (FSR) through charge injection, but rather by a small perturbation. We can solve for the new wavelength, which will initially be represented by the perturbed wavenumber $k_0 + \Delta k$.

$$(k_0 + \Delta k) \cdot L_{tot} + \Delta\phi = 2\pi m \tag{2.17}$$

$$k_0 \cdot L_{tot} + \Delta k \cdot L_{tot} + \Delta\phi = 2\pi m \tag{2.18}$$

By subtracting $k_0 \cdot L_{tot} = 2\pi m$ from Equation 2.1 earlier,

$$\Delta k \cdot L_{tot} + \Delta\phi = 0 \tag{2.19}$$

This equation is not yet in a useful form to see how the ring's resonance shifts as a result of the phase change of the light. From the general form of the dispersion

relation, $k = \frac{\omega n}{c}$:

$$\Delta k = \Delta(\frac{\omega n_{eff}}{c_s}) \tag{2.20}$$

By definition,

$$\Delta k = \frac{\Delta k}{\Delta \omega} \cdot \Delta \omega \tag{2.21}$$

$$\Delta k = \frac{\Delta(\omega n_{eff})}{\Delta \omega} \cdot \frac{1}{c_s} \cdot \Delta \omega \tag{2.22}$$

The group index can be written as $n_g = \frac{\Delta(\omega n_{eff})}{\Delta \omega}$.

$$\Delta k = \frac{n_g}{c_s} \cdot \Delta \omega \tag{2.23}$$

This expression is combined with Equation 2.19 and simplified.

$$\frac{n_g L_{tot}}{c} \cdot \Delta \omega + \Delta \phi = 0 \tag{2.24}$$

The expression for FSR from Equation 2.6, $FSR = c/(L_{tot} \cdot n_g)$, is used to simplify further.

$$\frac{1}{FSR} \Delta \omega = -\Delta \phi \tag{2.25}$$

$$\frac{1}{FSR} \frac{\Delta \omega}{2\pi} = -\frac{\Delta \phi}{2\pi} \tag{2.26}$$

$$\frac{\Delta f}{FSR} = -\frac{\Delta \phi}{2\pi} \tag{2.27}$$

Equation 2.27 describes how much the ring's resonant frequency will shift (Δf) based on the phase shift of the light inside of the ring ($\Delta \phi$). This is useful because it is one step closer to an analytical link between the charge in the I-region and the transmissivity of the ring in the optical channel. As mentioned earlier, the receiver

sensitivity and the laser power will determine the minimum extinction ratio that can be tolerated. T_{sdB} determines the maximum extinction ratio that is theoretically possible. The transmissivity cannot exceed 0 dB, so if T_{sdB} is -10 dB, for instance, the maximum possible extinction ratio would be 10 dB.

The Lorentzian distribution in Equation 2.28 describes the relationship between T and $\Delta\phi$. The full-width-half-max phase shift $\Delta\phi_{HWHM}$ will change the transmissivity from T_0 to $2T_0$, or 3 dB. Note from Figure 2-1 that the ring's transmissivity at its shifted resonant frequency is higher than T_0 due to the optical loss introduced by the injected carriers. This behavior decreases the extinction ratio, but it is not modeled in this analysis. The insertion and through loss is also affected with higher carrier injection.

$$T(\Delta\phi) = 1 - \frac{1 - T_0}{1 + (\frac{\Delta\phi}{\Delta\phi_{HWHM}})^2} \tag{2.28}$$

The full-width-half-max bandwidth of the ring Δf_{FWHM} (shown graphically in Figure 2-1) and full-width-half-max phase shift $\Delta\phi_{FWHM}$ are related by Equation 2.27.

$$\frac{\Delta f_{FWHM}}{FSR} = -\frac{\Delta\phi_{FWHM}}{2\pi} \tag{2.29}$$

The half-width-half-max phase shift is half of the full-width-half-max phase shift. The negative sign can be neglected because the Lorentzian distribution is symmetric.

$$2\Delta\phi_{HWHM} = \Delta\phi_{FWHM} = 2\pi \cdot \frac{\Delta f_{FWHM}}{FSR} \tag{2.30}$$

$$\Delta\phi_{HWHM} = \pi \cdot \frac{\Delta f_{FWHM}}{FSR} \tag{2.31}$$

Therefore, when the phase shift $\Delta\phi$ reaches $\Delta\phi_{HWHM}$, the optical extinction ratio will be 3 dB. The amount of charge necessary to create this phase shift is determined by combining the expressions for $\Delta\phi$ from Equation 2.14 and $\Delta\phi_{HWHM}$ from Equation

2.31.

$$\frac{\Delta\phi}{\Delta\phi_{HWHM}} = \frac{\Delta n_{eff} \cdot 2L_j \cdot k_0}{\pi \cdot \Delta f_{FWHM}/FSR} \tag{2.32}$$

A few substitutions are necessary. From Equation 2.10, $\Delta n_{eff} = \Delta n \cdot \Gamma$; from Equation 2.6, $FSR = c/(n_g \cdot L_{tot})$, and $k_0 = 2\pi/\lambda_0$.

$$\frac{\Delta\phi}{\Delta\phi_{HWHM}} = \frac{\Delta n \cdot \Gamma \cdot 2L_j \cdot 2\pi}{\lambda_0 \cdot \pi \cdot \Delta f_{FWHM}} \cdot \frac{c}{n_g \cdot L_{tot}} \tag{2.33}$$

From Equation 2.7, $\Delta n = n_{fe}\Delta N_e + n_{fh}(\Delta N_h)^{0.8}$, and the ring's total length L_{tot} is a combination of the two straight sections and the two curved sections; $L_{tot} = 2L_j + 2\pi r$.

$$\frac{\Delta\phi}{\Delta\phi_{HWHM}} = \frac{[n_{fe}\Delta N_e + n_{fh}(\Delta N_h)^{0.8}] \cdot \Gamma \cdot 4L_j \cdot c}{\lambda_0 \cdot \Delta f_{FWHM} \cdot n_g(2L_j + 2\pi r)} \tag{2.34}$$

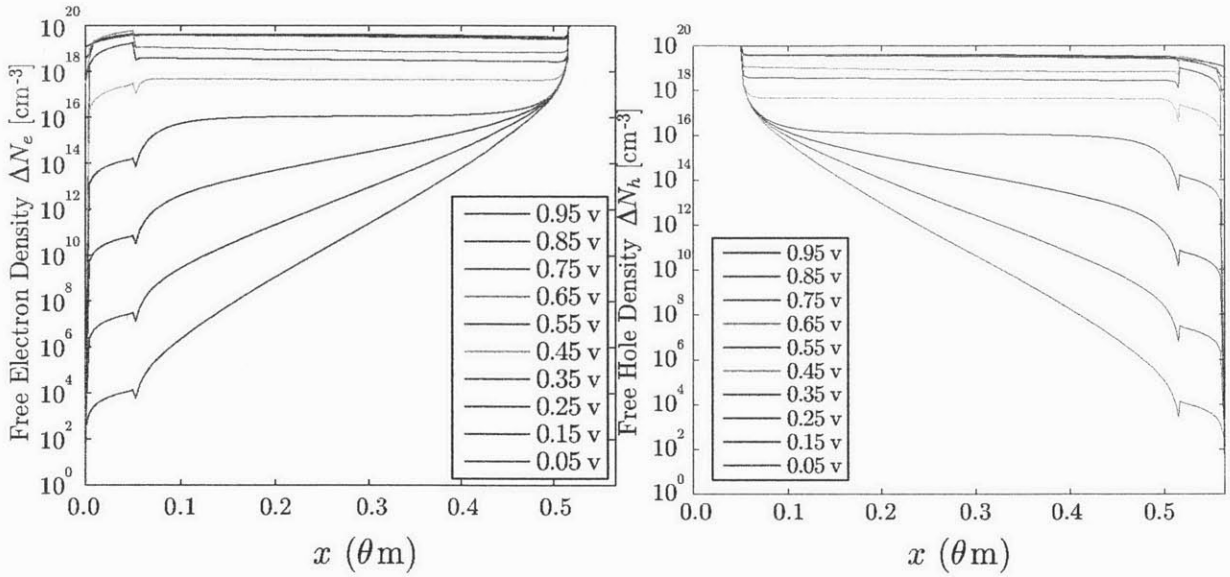
$$\gamma = \frac{\Delta\phi}{\Delta\phi_{HWHM}} \tag{2.35}$$

The ratio $\Delta\phi/\Delta\phi_{HWHM}$ is referred to as the injection ratio γ. The next step is to understand what amount of charge injection will reach a given γ, and to do so, it is critical to understand the charge distribution througout the P-I-N diode.

2.1.6 Uniform Carrier Injection

Figure 2-6 shows the free electron and hole concentration in the I-region cross-section with varying voltage (from Sentaurus simulations) showing the uniform charge density above $V \approx 0.4\,\text{V}$. The I-region extends from $x = 0.05\,\mu\text{m}$ to $x = 0.515\,\mu\text{m}$. Equation 2.34 makes the assumption that the carrier injection across the I-region is uniform. A Sentaurus[2] device simulation of the cross-section of the diode is shown in Figure 2-6. In this figure, the light propagates into the page, and the anode and cathode are on the left and right sides of the graph, respectively. The result of the simulation confirms that the carrier injection is uniform across the diode's length for both ΔN_e

[2]Sentaurus is a software suite developed by Synopsys for device simulation.



(a) Electron Density

(b) Hole Density

Figure 2-6: Free electron and hole concentration

and ΔN_h for I-region lengths of 0.5 to 2.0 μm and forward-bias voltages above 0.4 V, well below the turn-on voltage of the diode. Also, a distinction has been made between the free electron concentration ΔN_e and the free hole concentration ΔN_h, but for forward-bias voltages above 0.4 V, these concentrations are identical, confirmed by the same Sentaurus simulation and shown in Figure 2-6.

$$\Delta N = \Delta N_e = \Delta N_h \text{ for } V_D > 0.4 \text{ V} \tag{2.36}$$

ΔN now represents the free electron-hole pair concentration in the I-region.

Carrier Injection to Reach Half-Width-Half-Max

When $\Delta\phi = \Delta\phi_{HWHM}$ ($\gamma = 1$), the extinction ratio will be 3 dB. The charge concentration necessary to reach this extinction ratio will be defined as ΔN_0.

$$\frac{\Delta\phi}{\Delta\phi_{HWHM}} = \gamma = \frac{[n_{fe}\Delta N + n_{fh}(\Delta N)^{0.8}] \cdot \Gamma \cdot 4L_j \cdot c}{\lambda_0 \cdot \Delta f_{FWHM} \cdot n_g(2L_j + 2\pi r)} \tag{2.37}$$

$$\gamma = 1 = \frac{[n_{fe}\Delta N_0 + n_{fh}(\Delta N_0)^{0.8}] \cdot \Gamma \cdot 4L_j \cdot c}{\lambda_0 \cdot \Delta f_{FWHM} \cdot n_g(2L_j + 2\pi r)} \tag{2.38}$$

Equation 2.38 is a transcendental expression and requires an iterative numerical method to solve for ΔN_0. A simpler expression for the refractive index change Δn could be used to linearize the equation at a loss of accuracy.

Quality Factor

It is worth noting that Equation 2.38 can also be put in terms of the $\mathbb{Q}$ of the ring, demonstrating analytically that a higher $\mathbb{Q}$ will increase modulator performance by lowering the charge injection necessary for a given extinction ratio.

$$\mathbb{Q} = \frac{f_0}{\Delta f_{FWHM}} = \frac{c}{\lambda_0 \cdot \Delta f_{FWHM}} \tag{2.39}$$

$$\gamma = 1 = \frac{[n_{fe}\Delta N_0 + n_{fh}(\Delta N_0)^{0.8}] \cdot \Gamma \cdot 4L_j \cdot \mathbb{Q}}{n_g(2L_j + 2\pi r)} \tag{2.40}$$

Junction Length Tradeoff

One more discovery from Equation 2.38 is more apparent if the equation is put in terms of total carriers ΔN_{tot} rather than carrier concentration.

$$\Delta N_0 = \frac{\Delta N_{tot}}{V_j} = \frac{\Delta N_{tot}}{2L_j \cdot W \cdot H} \tag{2.41}$$

$$\gamma = 1 = \frac{\left[n_{fe} \cdot \frac{\Delta N_{tot}}{2L_j \cdot W \cdot H} + n_{fh} \cdot \left(\frac{\Delta N_{tot}}{2L_j \cdot W \cdot H}\right)^{0.8}\right] \cdot \Gamma \cdot 4L_j \cdot \mathbb{Q}}{n_g(2L_j + 2\pi r)} \tag{2.42}$$

The numerator and denominator are multiplied by the total junction volume $V_j = 2L_j \cdot W \cdot H$ and the ring modulator volume $V_{tot} = (2L_j + 2\pi r) \cdot W \cdot H$ is substituted.

$$\gamma = 1 = \frac{\left[n_{fe} \cdot \Delta N_{tot} + n_{fh} \cdot (2L_j \cdot W \cdot H)^{0.2} \cdot (\Delta N_{tot})^{0.8}\right] \cdot \Gamma \cdot 2 \cdot \mathbb{Q}}{n_g \cdot V_{tot}} \tag{2.43}$$

There is a design tradeoff involved in choosing L_j for a given ring circumference. A longer L_j will lower the series resistance of the diode, but it will also raise the junction capacitance. The significance of Equation 2.43 is that the effect of L_j on the modulation depth (for a given circumference) is weak, while it allows us to make tradeoffs at the circuit level. Notice that a larger L_j will require more current from the circuit.

Determining Carrier Injection for a Given Extinction Ratio

A more general form of Equation 2.38 to solve for ΔN given any extinction ratio h is a more useful tool; Equation 2.38 only solves for ΔN_0 for the specific extinction ratio of 3 dB.

$$h = T/T_0 \tag{2.44}$$

$$T = h \cdot T_0 \tag{2.45}$$

$$\tag{2.46}$$

The Lorentzian distribution describing the optical transmissivity from Equation 2.28 can be solved using the given h.

$$T(\Delta\phi) = 1 - \frac{1-T_0}{1+\gamma^2} \tag{2.47}$$

$$1 - T = \frac{1-T_0}{1+\gamma^2} \tag{2.48}$$

$$1 + \gamma^2 = \frac{1-T_0}{1-T} \tag{2.49}$$

$$\gamma = \sqrt{\frac{1-T_0}{1-T} - 1} \tag{2.50}$$

$$\gamma = \sqrt{\frac{1-T_0}{1-h\cdot T_0} - 1} \tag{2.51}$$

$$\tag{2.52}$$

Equation 2.37 is combined with Equation 2.51 to find the ΔN necessary to reach this injection ratio.

$$\gamma = \sqrt{\frac{1-T_0}{1-h\cdot T_0} - 1} = \frac{[n_{fe}\Delta N + n_{fh}(\Delta N)^{0.8}]\cdot\Gamma\cdot 4L_j\cdot c}{\lambda_0\cdot\Delta f_{FWHM}\cdot n_g(2L_j+2\pi r)} \tag{2.53}$$

Again, Equation 2.53 is a transcedental equation without a simple solution. The only unknown is ΔN.

2.1.7 A Ring Modulator at $\lambda = 1220$ nm

This section will use the methods of the previous section to analyze a realistic ring modulator designed for a commercial 32 nm bulk-CMOS process.

Below are some of the optical parameters and physical constants relevant to the ring modulator.

$$c = 3\times 10^8\,\text{m/s} \quad \text{(Speed of light)} \tag{2.54}$$

$$\lambda_0 = 1220\,\text{nm} \quad \text{(Resonant wavelength)} \tag{2.55}$$

$$n_g = 3.3 \quad \text{(Waveguide group index)} \tag{2.56}$$

$$\Gamma = 0.42 \quad \text{(Mode overlap)} \tag{2.57}$$

$$FSR = 1.4\,\text{THz} \quad \text{(Free spectral range)} \tag{2.58}$$

$$\Delta f_{FWHM} = 270\,\text{GHz} \quad \text{(Ring bandwidth)} \tag{2.59}$$

$$n_{fe} = 4.8\times 10^{-22}\,\text{cm}^{-}3 \quad \text{(Refraction electron coefficient)} \tag{2.60}$$

$$n_{fh} = 4.9\times 10^{-18}\,\text{cm}^{-}3 \quad \text{(Refraction hole coefficient)} \tag{2.61}$$

$$T_0 = -10\,\text{dB} \quad \text{(Transmissivity at resonance)} \tag{2.62}$$

$$h = 3\,\text{dB} \quad \text{(Target extinction ratio)} \tag{2.63}$$

$$\gamma = 1 \quad \text{(Injection ratio for } h = 3\text{ dB)} \tag{2.64}$$

$$\tag{2.65}$$

The ring geometry is given below using the dimensioning from Figure 2-3.

$$L_j = 6\,\mu\text{m} \quad \text{(Single junction length)} \tag{2.66}$$

$$r = 8\,\mu\text{m} \quad \text{(Bend radius)} \tag{2.67}$$

$$W = 430\,\text{nm} \quad \text{(Waveguide I-region width)} \tag{2.68}$$

$$H = 90\,\text{nm} \quad \text{(Waveguide height)} \tag{2.69}$$

$$V_j = 4.128 \times 10^{-13}\,\text{cm}^{-}3 \quad \text{(Diode I-region junction volume)} \tag{2.70}$$

$$L_{tot} = 62.3\,\mu\text{m} \quad \text{(Ring circumference)} \tag{2.71}$$

$$\tag{2.72}$$

The MATLAB `fzero` command was used to solve Equation 2.38 using the above parameters to find the injected charge density necessary for an extinction ratio of 3 dB.

$$\Delta N_0 = 2.1 \times 10^{18}\,\text{cm}^{-}3 \tag{2.73}$$

$$\Delta Q_0 = 1.9 \times 10^{-13}\,\text{C} \tag{2.74}$$

$$\tag{2.75}$$

These parameters were used to explore the modulator driver design space.

2.1.8 Finger Design

Due to the limitations of fabricating curved structures on a Manhattan grid, the ring is elongated into a racetrack shape so that the P-I-N diode structure can occupy a straight section of the waveguide [26]. Figure 2-7 also shows the geometry of the P-I-N structure in finer detail. The intrinsic silicon waveguide has connections on one side to thin p+ unsilicided poly fingers and on the other side to n+ fingers. Farther

[2]The actual device parameters for the process are protected under a non-disclosure agreement but the given parameters are a reasonable substitute.

from the waveguide, the finger tapers out to a silicided section with tungsten contacts to the first metal layer.

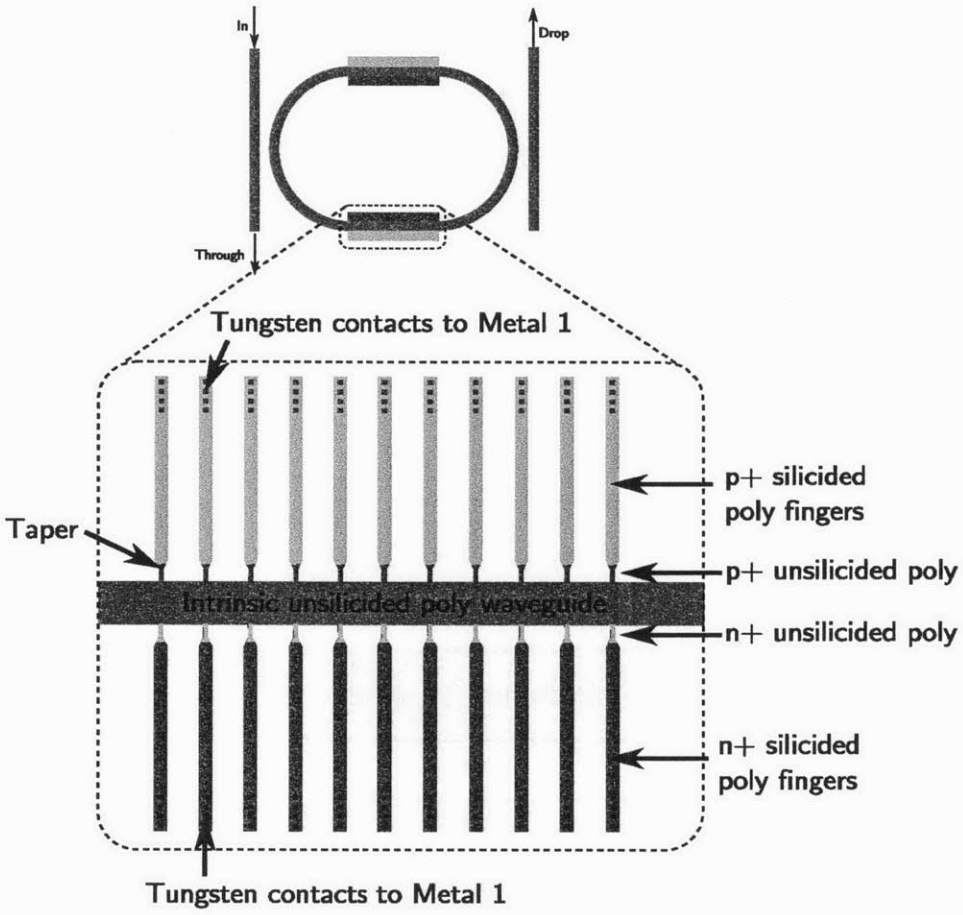


Figure 2-7: P-I-N finger structure.

There are several tradeoffs associated with the finger design. The fingers disrupt the optical mode and cause loss, lowering the ring's quality factor ($\mathbb{Q}$) and making it more difficult to achieve good modulation depth. However, the fingers have high resistance, and too few fingers will raise the series resistance of the diode, making charge injection difficult. A similar tradeoff exists with the choice of how far to place the diffusion region from the waveguide. Having the highly-doped region close to the waveguide increases loss but makes it easier to inject charge by having a lower diode body length. The finger spacing is also a critical parameter which must be carefully chosen by the device designers.

2.1.9 The P-I-N Diode

The P-I-N diode is a meeting point between the circuit and device designers. Both need accurate models of the diode to simulate and verify the behavior of the device. Due to the high-injection mode of operation, accurate device simulations and sophisticated circuit equivalent models are necessary to model and analyze the numerous design tradeoffs. The development of a valid model across the optical, device, and circuit domains requires iteration between multiple simulators. We have used a Padé approximation to create a SPICE circuit to model the diode's behavior [36]. The simulation design flow for developing a model for the P-I-N diode is shown in Figure 2-8.

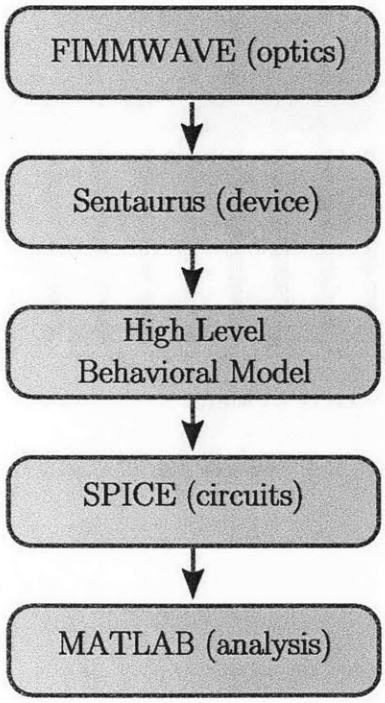


Figure 2-8: Simulation design flow.

2.1.10 P-I-N Diode Simulation Using Sentaurus

The Sentaurus device simulator can provide a more accurate physical simulation of the P-I-N diode at the expense of poor integration with complicated circuit descriptions.

A two-dimensional diode simulation setup is shown in Figure 2-9. The simulation is performed on a diode cross-section, where the light blue block in the center represents the intrinsic poly region with optical propogation going into and out of the page. The p+ doped anode is on the left of the figure and the n+ doped cathode is on the right, with oxide on the top and bottom of the polysilicon sections. A fine mesh is defined around the polysilicon sections, and to decrease simulation time, a coarser mesh is used for the oxide regions.

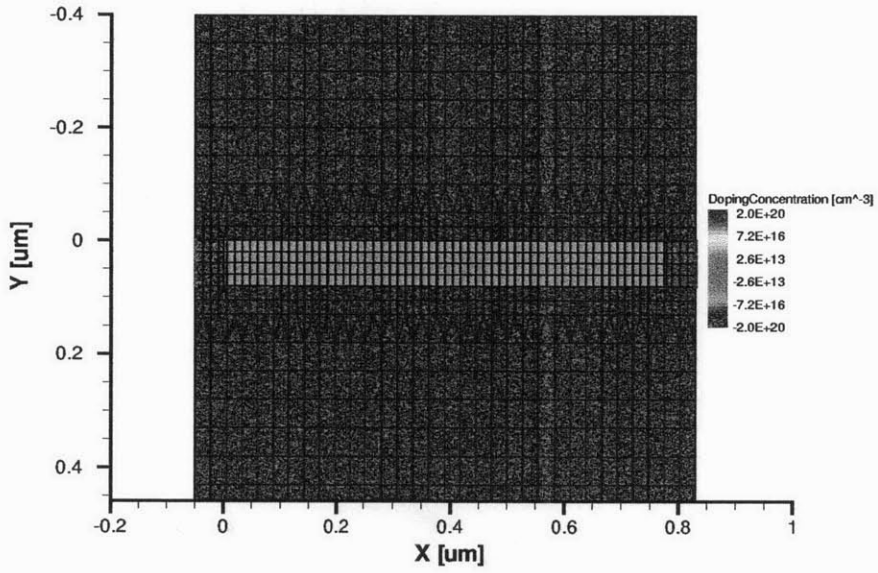


Figure 2-9: P-I-N cross-section showing Sentaurus simulation mesh.

This simple simulation setup allows several design tradeoffs to be analyzed. Waveguide thickness is set by the foundry and cannot be changed by the device designers, but the diode I-region length (shown in Figure 2-9 across the x-axis) is one parameter that can be chosen. A small diode length is good from an electrical perspective because it lowers the resistance of the device and increases the injected carrier density for a given voltage, as shown in the Sentaurus results in Figure 2-11 [3]. However, a diode body length smaller than the waveguide width will cause high through loss and

[3]Figure 2-11 shows data valid for silicon with a carrier lifetime of 1 μs. Polysilicon with a carrier lifetime of 1 ns achieves approximately 5x less carrier density at a bias of 1.0 V, but the shape of the graph (including the carrier concentration saturation) is similar

will further decrease the overlap of the injected carriers with the optical mode.

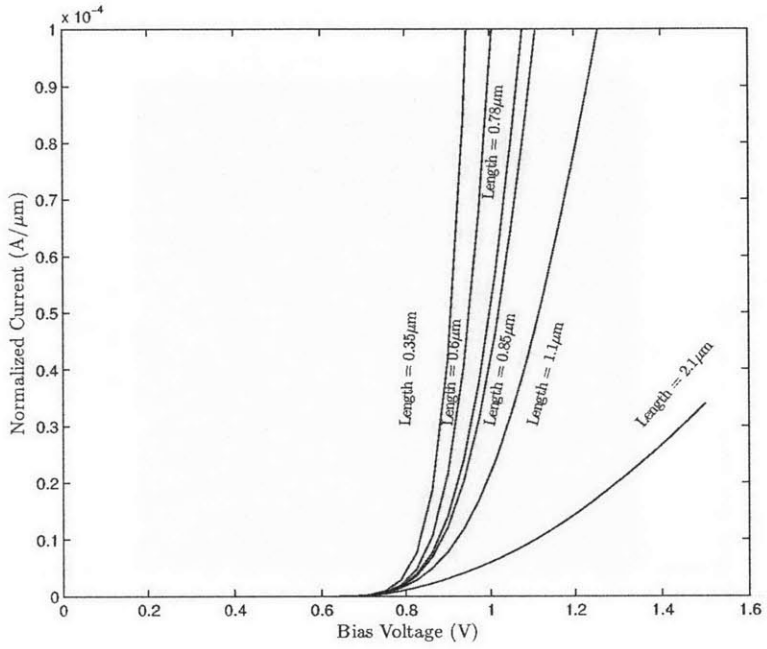


Figure 2-10: Sentaurus simulation of P-I-N diode I-V curve for various I-region lengths.

The I-V curve of the diode, shown in Figure 2-10, is used at design time to determine some first-order qualities of the driver circuit, such as the on-resistance and the required current drive necessary to reach a target voltage and charge concentra-

tion. The I-V curve is clearly dependent on the diode I-region length (where I-region "length" from a diode point-of-view is equivalent to the waveguide "width"). Note that the drawn diode width does not necessarily correspond to the fabricated width due to the dopant diffusion in the fingers.

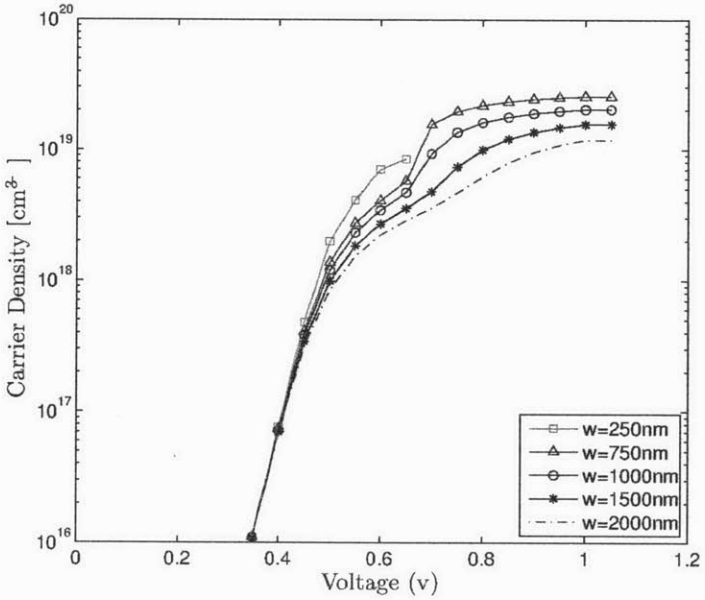


Figure 2-11: Sentaurus simulation of injected charge concentration vs. voltage.

The Sentaurus simulations of the P-I-N diode reveal other aspects of the diode behavior that will affect the design of the driver circuit. Figure 2-11 shows the simulated dependence of carrier density on the diode base (I-region) width. It is

important to note that the carrier concentration saturates at voltages above 0.8 V, meaning that driving the diode further into forward bias wastes power and can only help amortize the series resistance of the driver, with no other optical advantage.

Figure 2-6 demonstrates that we may assume that both free electron and free hole concentrations are uniform across the I-region of the diode when the diode's voltage is above the forward-bias threshold.

2.1.11 The Physical Carrier-Injection Design

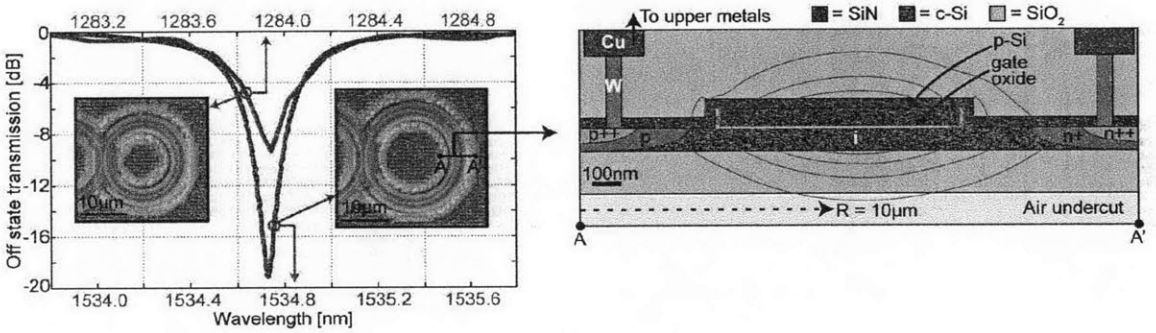
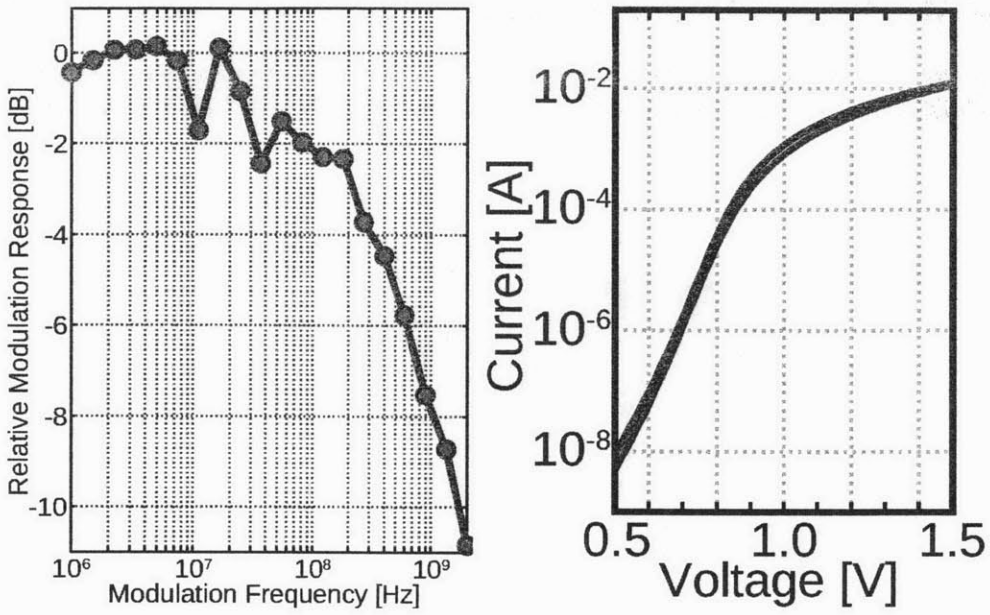


Figure 2-12: EOS8 Carrier-Injection Ridge Modulators for 1550 nm and 1280 nm with cross-sections [25]

The fabricated EOS carrier-injection ridge modulators, designed to operate via carrier injection in the 1280 nm and 1550 nm bands, are shown along with their cross-sections in Figure 2-12 [25]. Polysilicon is used to help confine the mode to the middle of the waveguide and avoid the overlapp with lossy P,N+ regions and metal contacts. Even though polySi is much higher loss (50-60dB/cm) compared to body-Si (3-4dB/cm) the polySi helps confine the mode better in the absence of partial-etch step.



(a) Frequency Response

(b) Device I-V Curve

Figure 2-13: EOS12 Carrier-Injection Modulator Device

The low 3 dB bandwidth of the modulator (approximately 200 MHz) is caused by the long minority carrier lifetime of the silicon (on the order of a few nanoseconds). The long minority carrier lifetime is because there are very few defects in crystalline silicon. This low bandwidth motivates interesting circuit design choices, outlined later in Chapter 3.

2.2 Charge-Depletion Modulator Device

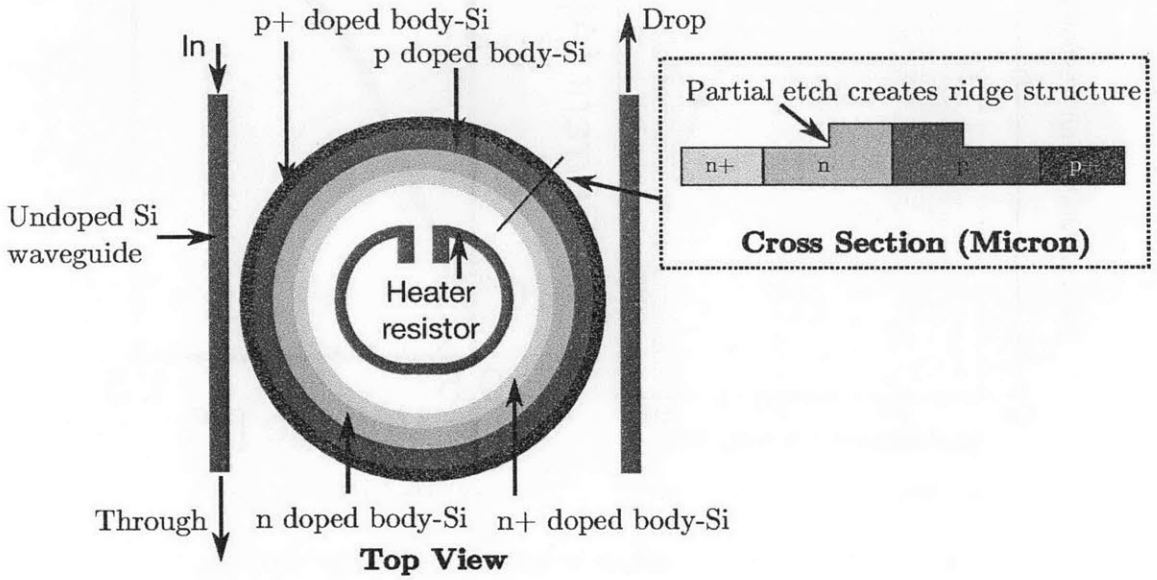


Figure 2-14: Micron platform depletion-mode ring modulator diagram (not to scale)

The second type of modulator device to be considered is a resonant ring filter with a stopband that shifts due to charge depletion rather than charge injection. A diagram of a lateral-junction depletion-mode modulator with both top and cross-section views is shown in Figure 2-15. This depletion-mode modulator (constructed in the Micron peripheral-memory process) makes use of a partial-etch process to form a ridge waveguide in the poly-silicon waveguide layer. The amorphized polysilicon layer [19] [1] is doped such that the cross-section forms a horizontal diode, which operates as a varactor in the reverse bias regime. The highly-doped n+ and p+ outer edges serve as contacts to the diode structure. As the diode is reverse biased, the depletion region grows, depleting carriers from the center of the structure where overlap with the optical mode is the highest. Another alternate design is possible in a processes where a partial etch is not available.

Again, the diode structure provides a meeting point for device and circuit designers.

2.2.1 Charge-Depletion Ring Modulator Dimensions

Before proceeding into an analysis and model of a charge-depletion modulator, it is useful to again define dimensions of the device. Variable names have been chosen to be different from the carrier-injection modulator device of the previous section.

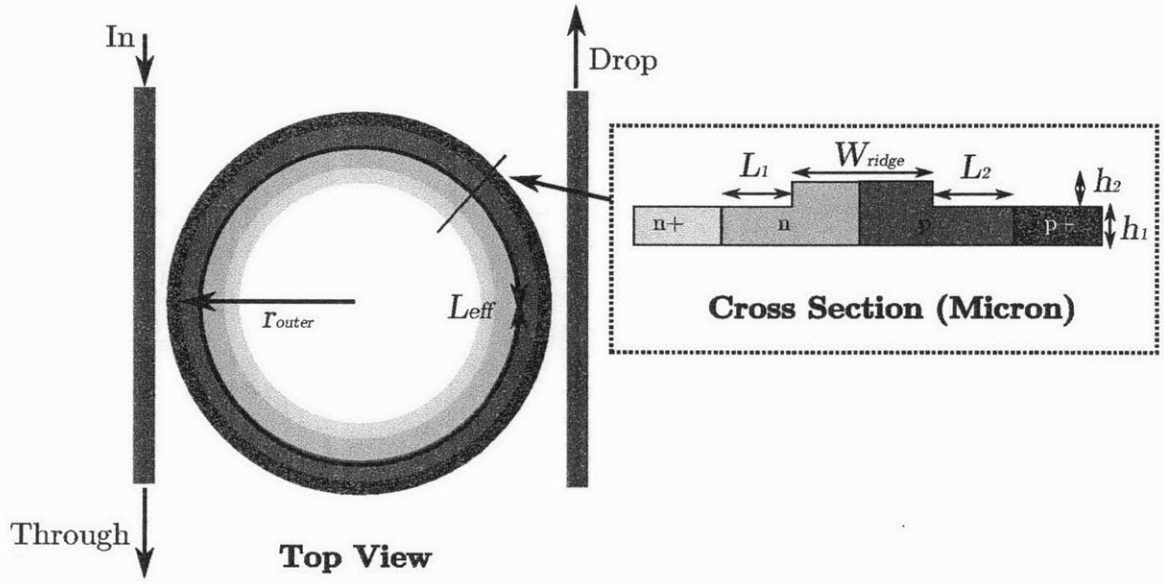


Figure 2-15: Depletion-mode ring modulator dimensions (heater ommitted; not to scale)

The radius from the center of the ring to the outermost edge of the p-type silicon is r_{outer}. In this device, there is no need for a racetrack shape with a straight section for finger contacts, so the ring is designed to be perfectly round. The effective circumference of the ring is $L_{eff} = 2\pi \cdot (r_{outer} - L_2 - W_{ridge}/2)$. The width of the ridge section (defined by the partial etch) is W_{ridge}; the partially-etched section of n-type silicon is denoted by L_1 and the partially-etched section of p-type silicon is denoted by L_2, such that the effective waveguide width is $L_1 + L_2 + W_{ridge}$. The total height of the ridge section is $h_{tot} = h_1 + h_2$.

The total effective ring volume, V_{eff}, is given below.

$$V_{eff} = h_1 \cdot \left(\pi(r_{outer} + L_2)^2 - \pi(r_{outer} - W_{ridge} - L_1)^2\right) + h_2 \cdot \left(\pi r_{outer}^2 - \pi(r_{outer} - W_{ridge})^2\right) \quad (2.76)$$

2.2.2 Determining the Depletion Region Width

The cross-section of the depletion-mode ring modulator is a standard P-N diode. Evaluating this junction in the reverse-bias regime will help produce a model that can be used by circuit designers to estimate the optical output of the device.

The Einstein relation in 2.78 describes the relationship between the diffusion constant for electrons (D_e) and holes (D_h), the mobility for electrons (μ_e) and holes (μ_h), and the thermal voltage ($V_T = kT/q$) [29, p. 102]. These parameters are all known from the foundry; we assume the ring is operating at room temperature ($T = 300\text{K}$).

$$D_e = \mu_e V_T \quad (2.77)$$

$$D_h = \mu_h V_T \quad (2.78)$$

The minority carrier diffusion lengths for electrons (L_e) and holes (L_h) is proportional to the diffusion constant for electrons (τ_e) times the minority carrier lifetime for electrons (τ_e) and holes (τ_h)[29, p. 131], shown in 2.80. These parameters are also known from the foundry.

$$L_e = \sqrt{D_e \tau_e} \quad (2.79)$$

$$L_h = \sqrt{D_h \tau_h} \quad (2.80)$$

The general form of the reverse-bias saturation current for a P-N diode is given in 2.81, where q is the charge of an electron, A is the diode cross-sectional area, N_A

is the n-type doping, and N_D is the p-type doping[29, p. 249].

$$I_s = qA\left(\frac{D_e}{L_e N_A} + \frac{D_h}{L_h N_D}\right) \cdot n_i^2 \tag{2.81}$$

In the case of this particular diode junction, the area is $A = (h_{tot}) * L_{eff}$.

$$I_s = q h_{tot} L_{eff}\left(\frac{D_e}{L_e N_A} + \frac{D_h}{L_h N_D}\right) \cdot n_i^2 \tag{2.82}$$

The reverse-bias saturation current I_s is thus a known quantity. The diode built-in potential V_{bi} is also known [29, p. 204]:

$$V_{bi} = V_t \cdot ln\left(\frac{N_A N_D}{n_i^2}\right) \tag{2.83}$$

The unbiased depletion region width can be estimated with the following formula[29, p.214].

$$x_{d0} = \sqrt{2\frac{\epsilon_s \epsilon_0}{q}\frac{(N_A + N_D)}{N_A N_D} \cdot V_{bi}} \tag{2.84}$$

The n-side and p-sides of the depletion region can be calculated separately.

$$x_{n0} = \frac{x_{d0}}{1 + \frac{N_D}{N_A}} \tag{2.85}$$

$$x_{p0} = \frac{x_{d0}}{1 + \frac{N_A}{N_D}} \tag{2.86}$$

The unbiased junction capacitance can also be calculated. This provides a useful estimate of the device capacitance to a circuit designer for fan-out analysis.

$$C_{j0} = \frac{\epsilon_0 \epsilon_s L_{eff} h_{tot}}{x_{d0}} \tag{2.87}$$

If the diode reverse-bias voltage V_{cap} exceeds the built-in voltage V_{bi}, the depletion-region width will change. The new n-side depletion-region width is x_n and the p-side

depletion-region width is x_p.

$$x_n = x_{n0}\sqrt{1 - \frac{V_{cap}}{V_{bi}}} \tag{2.88}$$

$$x_p = x_{p0}\sqrt{1 - \frac{V_{cap}}{V_{bi}}} \tag{2.89}$$

We can calculate the change in the depletion region widths for both the n-side and p-side.

$$\Delta x_n = x_{n0} - x_n \tag{2.90}$$

$$\Delta x_p = x_{p0} - x_p \tag{2.91}$$

This depletion of charge from the waveguide will lead to an index shift.

2.2.3 Determining the Index Shift, Absorption Change, and Resonance Shift

The modulating depletion region overlaps with the optical mode. As was done with the carrier-injection modulator, because the depletion region is only overlapping with a portion of the mode, an overlap factor Γ can be applied to account for the discrepancy. Γ_n is the overlap factor for the n-side of the diode, and Γ_p is the overlap factor for the p-side of the diode. Both overlap factors are close to 1 for the Micron design.

As with the carrier-injection modulator in equation 2.7, repeated below, the the index change (Δn) is sensitive to the change in both free electron concentration $\Delta N_e \approx N_D$ and free hole concentration $\Delta N_h \approx N_A$.

$$\Delta n_e = n_{fe}\Delta N_D \tag{2.92}$$

$$\Delta n_h = n_{fh}(\Delta N_A)^{0.8} \tag{2.93}$$

This index change only has an effect on the mode where the depletion region is overlapping with the mode. The mode is largely confined within the ridge region, so the effective index change is approximated as the proportion of the depletion-region change to the ridge width. This approximation should be accurate to within a factor of two.

$$\Delta n_{eff} = \Delta n_e \Gamma_n \frac{\Delta x_n}{W_{ridge}} + \Delta n_h \Gamma_p \frac{\Delta x_p}{W_{ridge}} \tag{2.94}$$

The change in the $\mathbb{Q}$-factor of the ring due to the change in charge can also be modeled. From Soref's 1987 paper, we can also extract the free carrier absorption change (Figure 2-16), and scale the data to $\lambda = 1200$ nm, as was done in 2-5 for the refraction change.

Similar to the change in refractive index, the effective change in absorption can approximated by taking into account the mode overlap factors.

$$\Delta \alpha_{eff} = \Delta \alpha_e \Gamma_n \frac{\Delta x_n}{W_{ridge}} + \Delta \alpha_h \Gamma_p \frac{\Delta x_p}{W_{ridge}} \tag{2.95}$$

This effective change in absorption is used to calculate the change in the $\mathbb{Q}$-factor of the ring.

$$\Delta \mathbb{Q} = 2\Delta \alpha_{eff} \frac{c}{n_g \lambda_s} \tag{2.96}$$

The change in refractive index leads to a shift in the ring's resonance, $\Delta f = f_s - f_n$.

$$\Delta f = f_s - f_n = \Delta n_{eff} \frac{c}{\lambda_s n_g} \tag{2.97}$$

Finally, the change in the full-width-half-maximum bandwidth of the ring can be calculated from the new ring resonance and the change in $\mathbb{Q}$-factor.

$$\Delta FWHM = f_n * \Delta \mathbb{Q} \tag{2.98}$$

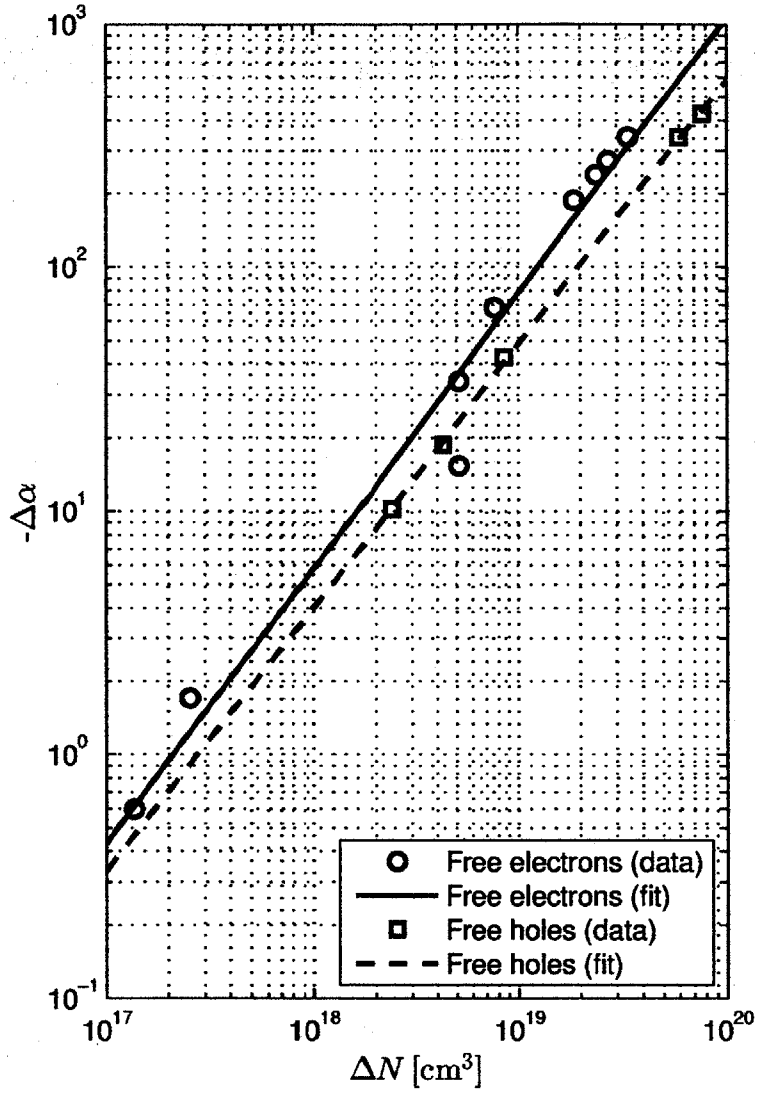


Figure 2-16: Silicon free carrier absorption change at $\lambda = 1200\,\text{nm}$[33]

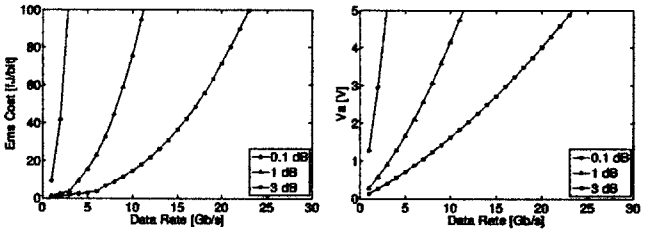
2.2.4 Determining Transmissivity

The necessary charge difference between the on- and off-states, ΔQ, is determined in Equation 2.99 by evaluating the Lorenzian transfer function, where $T_1 = 1/IL$ is the ring's transmissivity in its shifted on-state, $T_0 = 1/(ER \cdot IL)$ is its transmissivity in its off-state, and T_n is its transmissivity at the ring's resonant wavelength. Q_0 is the charge difference necessary to shift the ring by its half-width-half-max bandwidth [39], where q is the charge of an electron, $n_g = 4$ is the group index of the ring, and some typical ring modulator parameters are set as $V_{tot} = 2.6 \times 10^{-12}$ cm^3 is the total volume

of the ring, $n_f = 3 \times 10^{-21}$ cm^3 is the carrier-induced index change per unit carrier density at $\lambda_0 = 1300$ nm, and $\Gamma = 0.4$ is the overlap of the optical mode with the ring cross-section. The quality factor $\mathbb{Q}_f$ of the ring is set by the required data rate as $\mathbb{Q}_f = \frac{8 \cdot \pi}{3\lambda_0 \cdot DR}$, with a maximum value of 1×10^5, limited by practically achievable optical losses.

$$\Delta Q = Q_0 \cdot \left(\sqrt{\frac{T_1 - T_n}{1 - T_1}} - \sqrt{\frac{T_0 - T_n}{1 - T_0}} \right) \tag{2.99}$$

$$Q_0 = \frac{q \cdot n_g \cdot V_{tot}}{2 \cdot \mathbb{Q}_f \cdot n_f \cdot \Gamma} \tag{2.100}$$



(a) Device energy from supply vs. datarate.

(b) Reverse bias voltage V_a vs. datarate.

Figure 2-17: Device requirements to reach the target $ER_{dB} = 6$dB for various values of IL_{dB} [10].

The required charge difference is integrated onto the nonlinear junction capacitance (Equation 2.102) to determine the minimum reverse-bias drive voltage. V_a is plotted in Figure 2-17b as a function of data rate, at fixed ER and various values of IL [10]. The required V_a is a critical design specification for the circuit designer; anything beyond the typical process VDD of 1 V will require special voltage-boosting circuitry, because adding a separate power rail is impractical. While doing this calculation, the designer must check that the calculated width of the depletion region, x_d, has not exceeded the actual height of the waveguide.

$$x_d = \sqrt{\frac{2 \cdot \epsilon \cdot (N_A + N_D)}{q \cdot N_A \cdot N_D}(V_{bi} + V_a)}, \qquad x_d < h \tag{2.101}$$

$$\Delta Q = Q(V_a) - Q(0) = \int_0^{V_a} \frac{C_{j0}}{\sqrt{1 + \frac{V}{V_{bi}}}} dV \tag{2.102}$$

$$V_a = V_{bi} \cdot \left[\left(\frac{Q - Q(0)}{2 \cdot V_{bi} \cdot C_{jo}} \right)^2 - 1 \right] \tag{2.103}$$

The optical transmissivity as a function of diode voltage is a very nonlinear function. A modulator model that operates in conjunction with circuit simulators such as SPICE is a necessary tool to produce a working design. A modulator model, presented in Appendix A.1, is implemented in Verilog-AMS. This model can be included in a SPICE simulation to provide the circuit designer with the optical output of a modulator driver simulation.

The Verilog-AMS modulator model has both electrical and optical ports. The electrical ports are the diode P and N terminals. The optical ports are the modulator device's In, Through, and Drop ports. These optical ports are represented as electrical ports from the simulator's perspective. The "voltage" on an optical port corresponds to the intensity of light on that port. For instance, if the "voltage" at the In port of the modulator is 1 V, and the "voltage" on the Through port for a '1'-bit is 0.5 V, this corresponds to an insertion loss of 3 dB. The absolute value of the intensity of light (again, the "voltage" represented by the simulator) is not important. The relative intensity of the Through port with respect to the In port allows easy calculation of the extinction ratio and insertion loss of the modulator device.

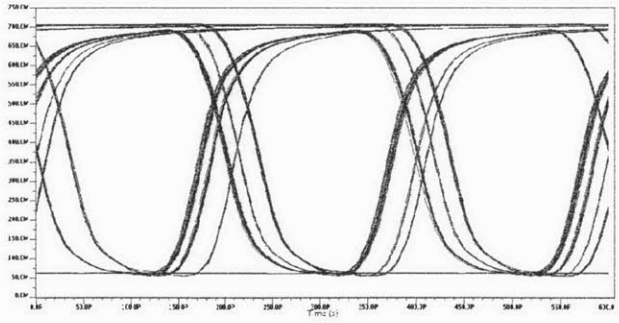


Figure 2-18: Optical eye diagram (extracted simulation) of the 456 fJ/bit D1 modulator driver at 5 Gb/s, with $ER_{dB} = 11$ dB. This speed is reaching the limit of the energy-optimized process, resulting in significant ISI at this datarate.

An optical eye diagram from an extracted simulation of a reverse-bias modulator driver (designed for the D1 chip) using this Verilog-AMS model is shown in Figure 2-18.

$$V_a = V_{bi} \cdot \left[\left(\frac{Q - Q(0)}{2 \cdot V_{bi} \cdot C_{jo}} \right)^2 - 1 \right] \tag{2.104}$$

The diode operates as a varactor in the reverse-bias regime. The required charge difference to meet a required IL, ER, and data rate DR specification can be calculated through Equation 2.102. The required reverse-bias diode voltage, V_a, is plotted in Figure 2-17b as a function of data rate, at fixed ER and various values of IL.

2.2.5 The Physical Carrier-Depletion Design

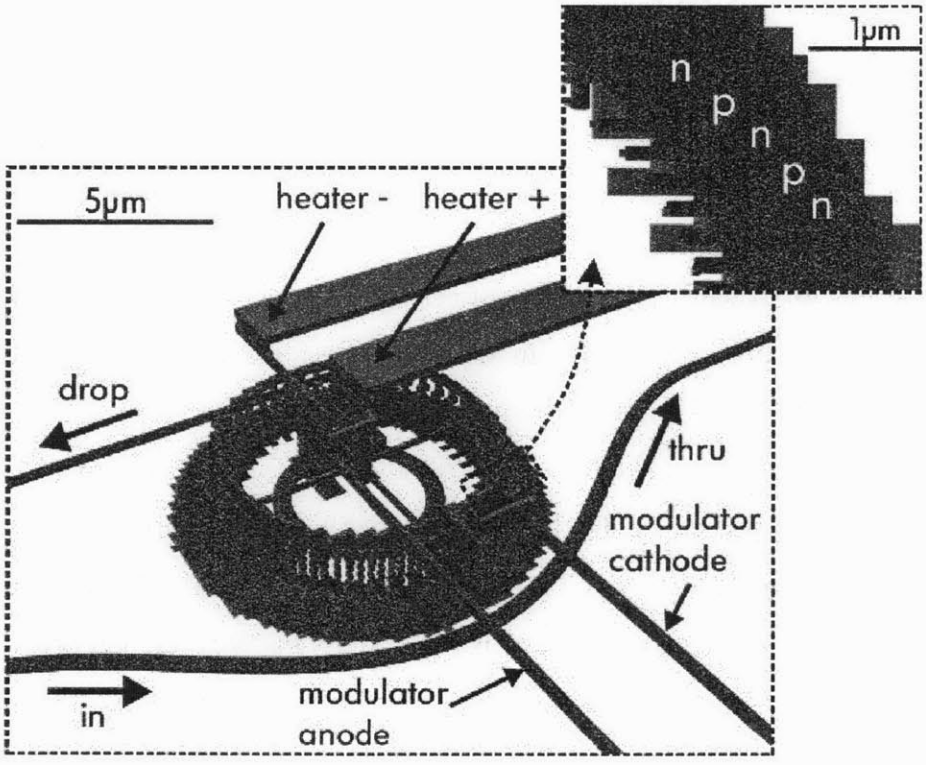


Figure 2-19: EOS16 Carrier-Depletion Modulators

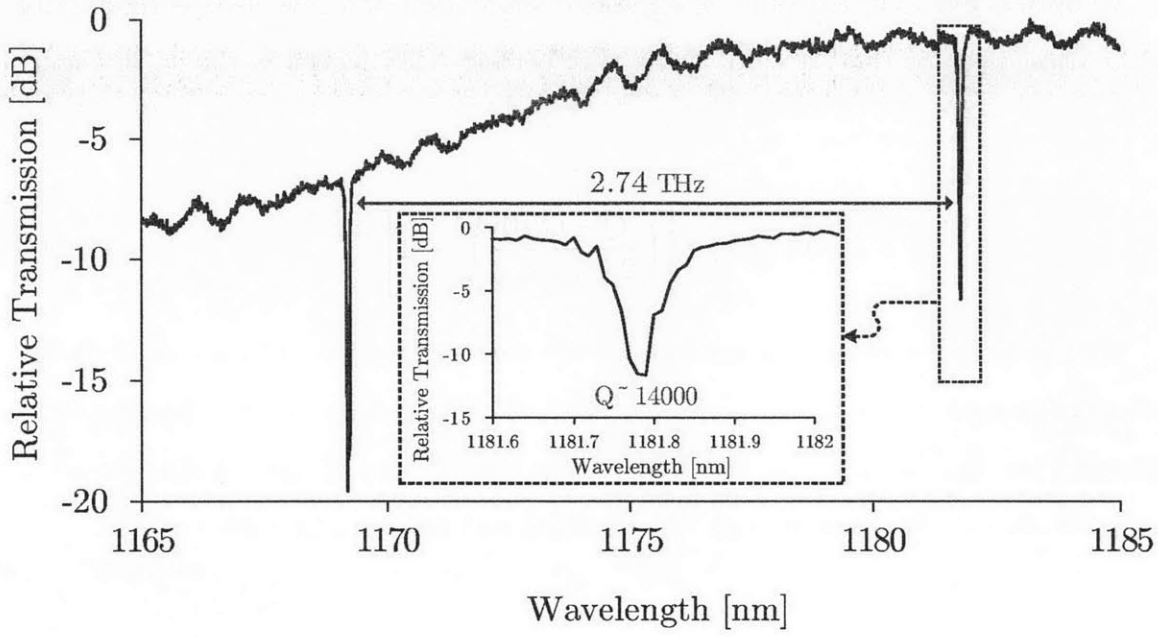


Figure 2-20: EOS16 Carrier-Depletion Modulator Response

The fabricated EOS carrier-injection ridge modulators, designed to operate via carrier injection in the 1280 nm and 1550 nm bands, are shown along with their cross-sections in Figure 2-12 [25].

2.3 Conclusion

This chapter has analyzed in detail the carrier-injection and carrier-depletion modulator topologies. For both topologies, an analytical method has been described, which circuit and device designers can use as a meeting point. In the carrier-injection case, a circuit designer must determine the amount of charge injection necessary to produce a given optical extinction ratio. In the carrier-depletion case, a circuit designer must know the necessary reverse-bias voltage to provide an optical device, and must also factor in the appropriate capacitance of the device. This chapter has provided this important tool, which was used in the development of the modulator driver circuits described in the next chapter.

Chapter 3

The Modulator Driver

3.1 Introduction

Whether in a carrier-injection or a carrier-depletion topology, the modulator driver changes the charge concentration inside the modulator diode, resulting in a change in the index of refraction inside the waveguide thus shifting the ring modulator's resonant frequency. From the previous chapter, for both designs, we have a link between a target extinction ratio and the carrier concentration change. Several driver topologies are presented, along with their applicability toward different scenarios. First we analyze a generalized model for a modulator driver, and then go into specific details for each type of design.

3.2 A General Modulator Driver Model

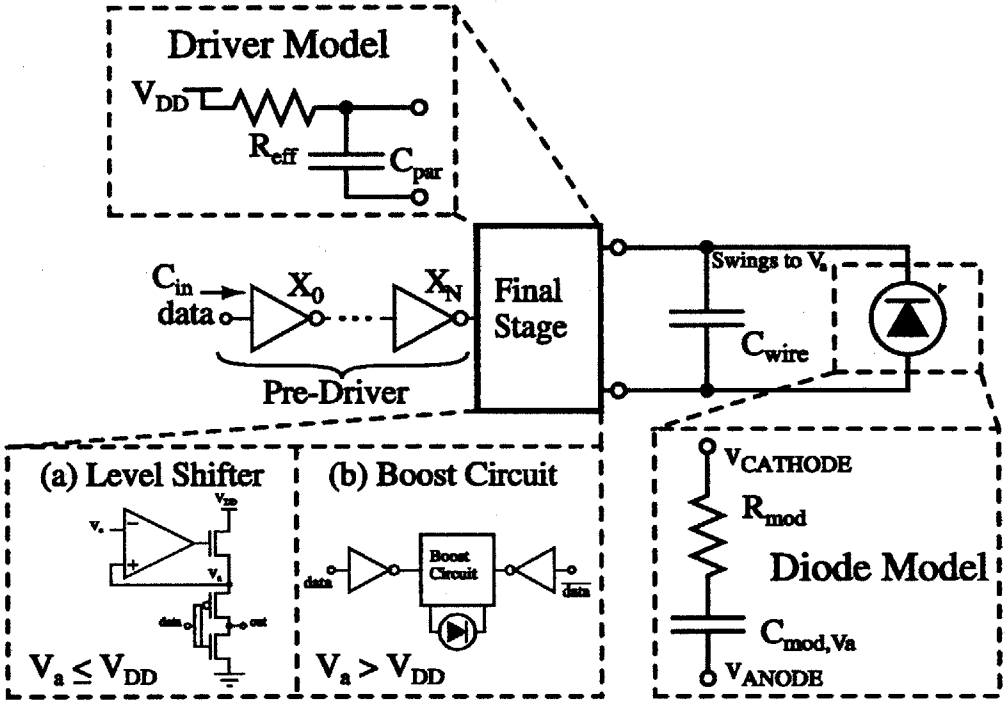


Figure 3-1: Electrical model of modulator device and driver.

The generalized driver model is shown in Figure 3-1 as an inverter chain pre-driver followed by a final driver stage. The final driver stage is attached to the modulator device, and drives the device to voltage V_a. There is some parasitic capacitance associated with the wiring device, C_w. To the first order, the diode itself (if in the reverse-bias regime) can be approximated as a series resistance R_{mod} and a reverse-bias capacitance $C_{mod,Va}$.

The circuit topology of the final drive stage should be chosen based on whether the driver is for a carrier-injection or carrier-depletion device, and it will also change depending on the diode drive voltage V_a compared to the process V_{DD}. For the carrier-depletion case, if $V_a < V_{DD}$, a low-swing topology can be used (Figure 3-1a); otherwise a voltage-boosting circuit may be necessary (Figure 3-1b). In both cases, the final stage can be modeled as an effective resistance R_{eff} and a parasitic capacitance C_{par}.

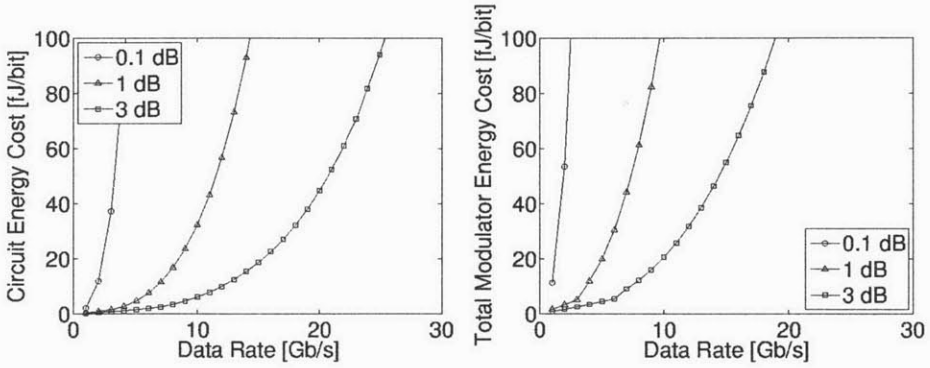
The R_{eff} will change with the width W of the final drive transistors as R_0/W. C_{par} is dominated by the gate capacitance of the final drive transistors. Because

electron mobility is typically twice that of hole mobility, PMOS transistors tend to be sized twice as large as their counterpart NMOS, making the total width $3 \cdot W$. C_{par} is modeled as $3 \cdot W \cdot \gamma \cdot C_g$, with $C_g = 1$ fF/um in the process and $\gamma = C_d/C_g$.

The final stage must be sized such that the target charge Q is achieved on C_{eff} in less than a bit-time. Logical effort analysis is used to determine an appropriate size for the driver's width (W) and pre-driver chain fan-out (FO) to meet the data rate requirements.

The CV^2 energy-per-bit of the final driver stage can be calculated based on its parasitic capacitance C_{par}, taking into consideration that for random data with 50% ones and zeros, a zero-to-one transition (a charging of that capacitance) will only happen 1/4 of the time. The "V^2" term in CV^2 will depend on whether the final driver stage is a low-swing or a boost driver. In the case of a low-swing driver, C_{par} is only charged to V_a with the energy coming from V_{DD} (thus $V_a \cdot V_{DD}$ equaling the "V^2" term); in the case of a boost driver, the "V^2" term is equal to V_a^2.

$$E_{dr} = \frac{C_{par} + C_w}{4} \cdot \max\left(V_a \cdot V_{DD}, V_a^2\right) + \frac{1}{4} \cdot \frac{3 \cdot C_g \cdot W}{1 - \frac{1}{FO}} \cdot V_{DD}^2 \tag{3.1}$$



(a) Driver energy cost E_{dr}. (b) Total modulator energy cost $E_{mod,tot} = E_{ms} + E_{dr}$.

Figure 3-2: Energy costs to reach the target $ER_{dB} = 6$ dB for various values of IL_{dB}.

Figure 3-2 shows the driver energy-per-bit cost to reach $ER_{dB} = 6$ dB for various values of IL_{dB}. For current modulator device technology, to satisfy data rates up to 30 Gb/s, the intrinsic RC time constant of the device allows designers to choose R_{mod}

up to 1 kΩ without significant energy penalty, relaxing the modulator optical losses due to contact placement. Verifying this theoretical model with measured results is one goal of this thesis.

3.3 Carrier Injection Modulator Driver

The double-data-rate modulator driver is designed as a configurable all-digital push-pull driver circuit with sub-bit-time pre-emphasis and split supplies, operating over the wide range of drive currents required for the variety of optical devices. The split power supply and level-shift drivers allow tradeoffs between energy efficiency and extinction ratio without increasing the power consumption of the back-end. Unlike previous work on sub-bit pre-emphasis [39], where different voltage values are used to pre-emphasize the modulator drive, in this work we change the on-resistance of the modulator driver connected to a fixed supply voltage. A novel signal conditioning technique allows the driver circuit to achieve an open eye from the modulator device at roughly 10x the intrinsic bandwidth of the device.

3.3.1 A Case for Pre-Emphasis

A pre-emphasized topology is used for the carrier-injection modulator driver. This design decision was made by first analyzing a simplified driver model, shown in Figure 3-3, to gain a basic understanding of the limits and requirements of the circuit. It is assumed that we have chosen a target extinction ratio (3 dB is a minimum choice) and a minimum $\Delta N_0 = 2.1 \times 10^{18}$ cm^{-3} necessary to reach that extinction ratio. The minimum injected charge necessary is $\Delta Q_0 = 1.9 \times 10^{-13}$ C for a ring modulator in an IBM 45 nm process with $L_j = 6$ μm, $r = 8$ μm, and a target extinction ratio of 3 dB at $\lambda_0 = 1200$ nm.

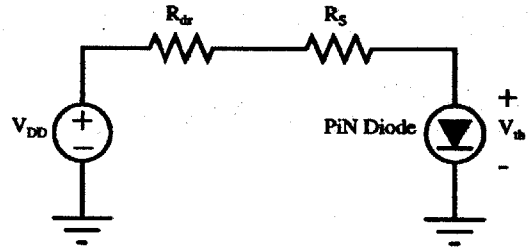


Figure 3-3: First order charge rise time schematic.

The charge $Q(t)$ in the I-region follows a first-order linear differential equation, where $i(t)$ is the diode current and τ_c is the carrier lifetime in the I-region [29].

$$\frac{dQ(t)}{dt} = i(t) - \frac{Q(t)}{\tau_c} \tag{3.2}$$

The solution to this equation makes the initial condition assumption that $Q(0) = 0$. Q_S is the steady-state charge in the I-region.

$$Q(t) = Q_S\left[1 - e^{-\frac{t}{\tau_c}}\right] \tag{3.3}$$

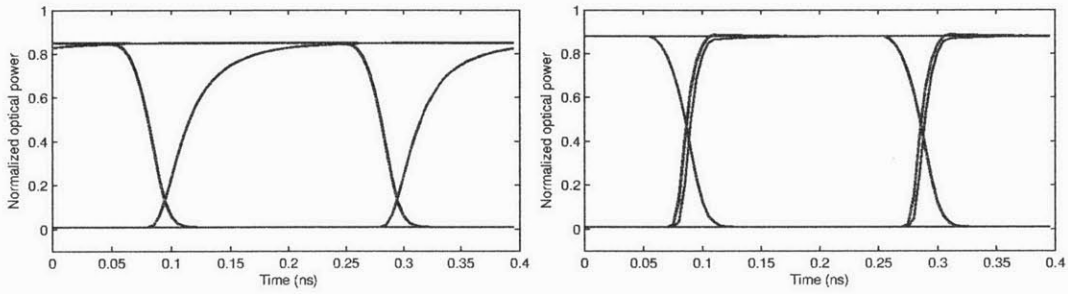
$$Q_S = \tau_c \cdot i(t) \tag{3.4}$$

The nonlinear relationship between the charge in the I-region and the optical transmissivity can be exploited to increase energy efficiency and speed. With a fixed driver resistance, on a transition from a 0 to a 1 bit, the charge in the I-region will steadily grow toward Q_S, reaching Q_S perhaps sometime in the middle of the bit. If the driver circuit provides $i(t)$ such that $Q_S = Q_0$, the resulting optical eye diagram will be poor at high speed due to the slow rising trajectory of $Q(t)$. A sub-bittime pre-emphasized driver can do much better [13]. A driver with a much lower resistance for a small portion of the bit time on a 0 to 1 transition will increase the charge in the I-region much more rapidly, improving the optical eye diagram [39]. The total driver resistance during pre-emphasis is referred to as R_A, the resistance during the remaining forward bias portion of the bit time is R_B, and the reverse-bias driver

resistance is R_C.

This pre-emphasis technique has three main advantages:

- It lowers the amount of steady-state charge necessary for a '1'-bit, increasing energy effeciency.
- Because the amount of steady-state charge has been reduced, there is less chance of shifting all the way to the next optical channel.
- The modulator can operate at higher speeds due to an improved eye as in Figure 3-4.



(a) Without pre-emphasis (b) With pre-emphasis

Figure 3-4: A simulation demonstrating how pre-emphasis can improve eye quality, raising the bitrate (1 ns carrier lifetime on random data).

To estimate the required driver resistance at a datarate of 2 Gb/s, it is assumed that the pre-emphasis driver will be active for the first quarter of a bit time for a zero-to-one transition, or 50 ps. The goal of the pre-emphasis driver will be to raise $Q(t)$ by ΔQ_0 by the end of the pre-emphasis pulse. The minority carrier lifetime, τ_c, is between 1-2ns in the IBM 45nm process.

$$Q(t = 50\,\text{ps}) = Q_0 = Q_{SA} \cdot (1 - e^{(-50\,\text{ps}/10\,\text{ns})}) \tag{3.5}$$

$$Q_{SA} = 3.9 \times 10^{-12}\,\text{C} \tag{3.6}$$

Q_{SA} is approximately 20x the steady-state charge, indicating that the transient current should be 20x larger than the steady-state. In a commercial CMOS process, V_{DD} is fixed, and is approximately 1.0 V at the 45 nm node. A digital driver circuit can be approximated to the first order by the circuit shown in Figure 3-3. R_{dr} is the total resistance of the driving transistors and R_S is the series resistance of the P-I-N diode. This simplification of the driver circuit is used to find an appropriate starting range for the driver resistance to make a first estimate at the transistor sizing.

The high V_{th} of the P-I-N diode means that a low R_S and R_{dr} are necessary to achieve a reasonable Q_S. The V_{th} of the modulator diode is measured as approximately 0.8 V for currents in the range of 1 mA. The V_{DD} limit of the process is also a concern. For demonstrating short-term functionalality of the modulator driver, V_{DD} can be temporarily raised to 1.5 or 2.0 V if the driver cannot provide enough current at 1.0 V. This solution is not feasible for mainstream adoption of silicon photonics due to the degradated reliability of the circuits. It is also important to note that above a certain bias voltage, the carrier concentration in the I-region begins to saturate, as shown in Figure 2-11, so increasing the voltage further may have little effect.

$$i(t) = \frac{V_{DD} - V_{th}}{R_A + R_S} \tag{3.7}$$

$$Q_{SA} = \tau_c \cdot \frac{V_{DD} - V_{th}}{R_A + R_S} \tag{3.8}$$

$$R_A + R_S = \tau_c \cdot \frac{V_{DD} - V_{th}}{Q_{SA}} = 51\Omega \tag{3.9}$$

With a diode contact resistance of approximately 30 Ω, the pre-emphasis driver resistance should be 20 Ω. Unfortunately, due to the high V_{th}, the transistors are likely to be in the triode region, and this target is difficult to achieve without an unreasonably large driver circuit. In the first generation of circuits we opt to raise the supply locally to 1.5 V, which relaxes the output driver requirements to at least $R_A = 150\ \Omega$. In practice, the goals can be achieved with even larger output impedance since at above a supply voltage of 1.2 V, the driver is mostly in saturation during the

transition and acts more as a current source.

Once the pre-emphasis pulse ends and the charge reaches the level of Q_0, the pre-emphasis driver will deactivate and the forward-bias driver will be active. The forward-bias driver will only sustain the charge at the level of Q_0.

$$R_B = \tau_c \cdot \frac{V_{DD} - V_{th}}{Q_0} - R_S = 1.0k\Omega \tag{3.10}$$

The forward-bias driver resistance is much more reasonable. We chose a nominal design point where the pre-emphasis driver is four times the strength of the forward-bias driver and two times the strength of the reverse-bias driver. This is just an approximate analysis to provide the design intuition.

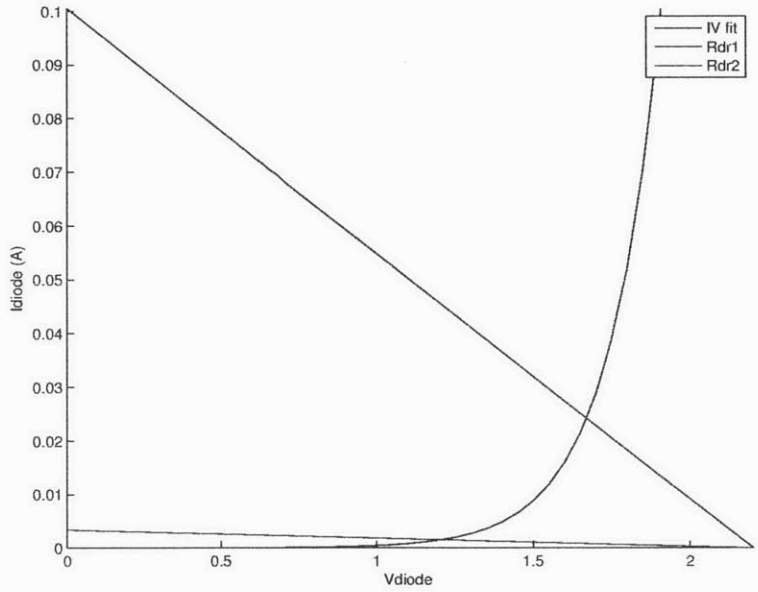


Figure 3-5: Load line showing pre-emphasis driver and forward-bias driver.

3.3.2 Carrier-Injection Driver Circuit

The modulator driver circuit is a custom digital push-pull circuit with sub-bittime pre-emphasis [23]. Two on-chip 31-bit pseudo-random-bit-sequence (PRBS) generators feed the double-data-rate (DDR) driver circuit (the PRBS generators are now shown in the schematic). The schematic for the driver is shown in Figure 3-6. The circuit

has four parallel drive segments, named segments A, B, C, and D. Each segment is responsible for either forward- or reverse-biasing the P-I-N diode depending on the appropriate drive regime. The strength of these drive segments can be digitally tuned across four bits. The inset in the schematic shows the logic levels of the control signals for each segment in the case of a 0-1-0 data pattern.

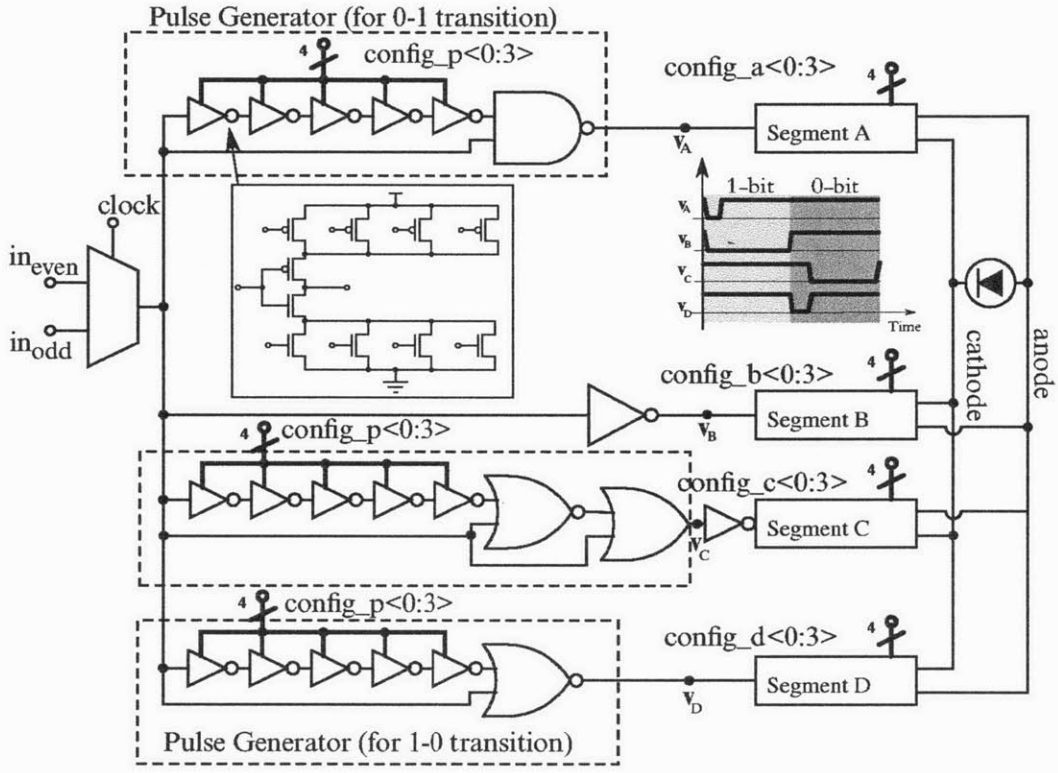


Figure 3-6: Conceptual schematic of the EOS-series carrier injection modulator driver.

Figure 3-7 illustrates which segments are active at the beginning of a zero to one transition. Both the sub-bittime forward-bias driver (Segment A) and the full-bittime forward-bias driver (Segment B) are active, maximizing the current into the I region. Segment A is activated by a tunable delay element.

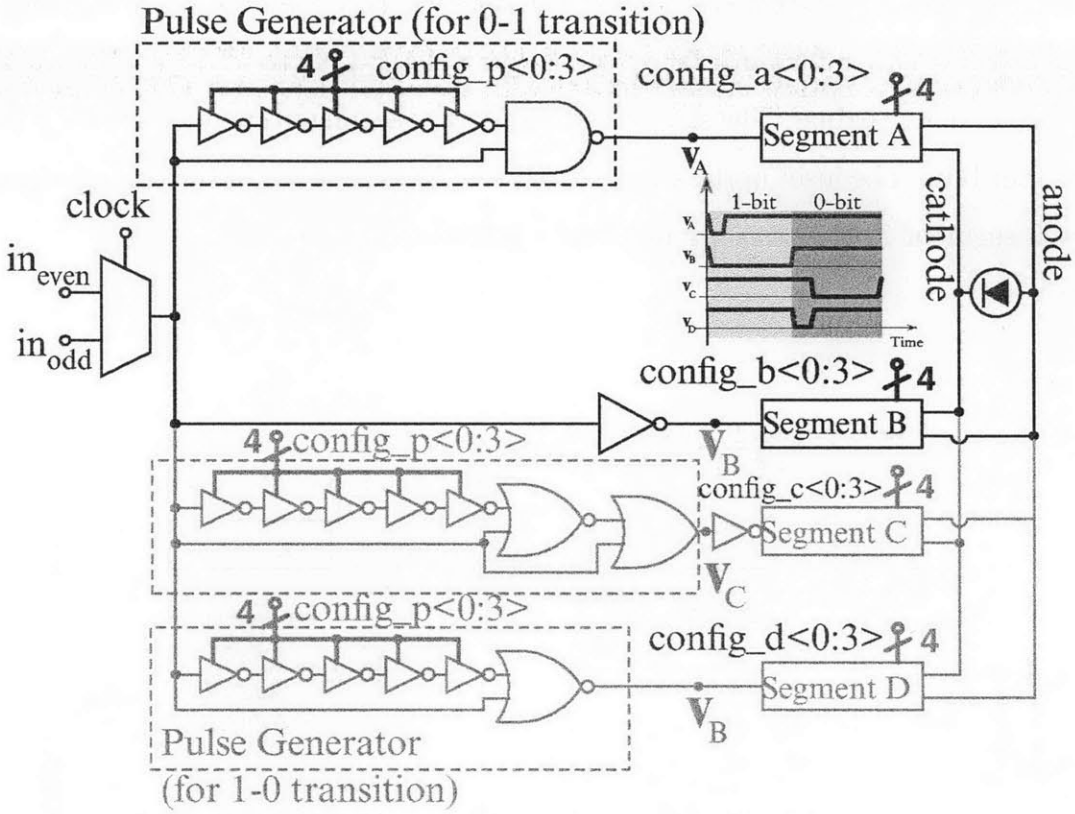


Figure 3-7: Conceptual schematic of the EOS-series carrier injection modulator driver. Segments A and B are active at the beginning of a 0-1-bit transition.

After a short time (less than a bit time), Segment A deactivates, while Segment B remains on for the duration of the one-bit, lowering the current into the diode. This is shown in Figure 3-8.

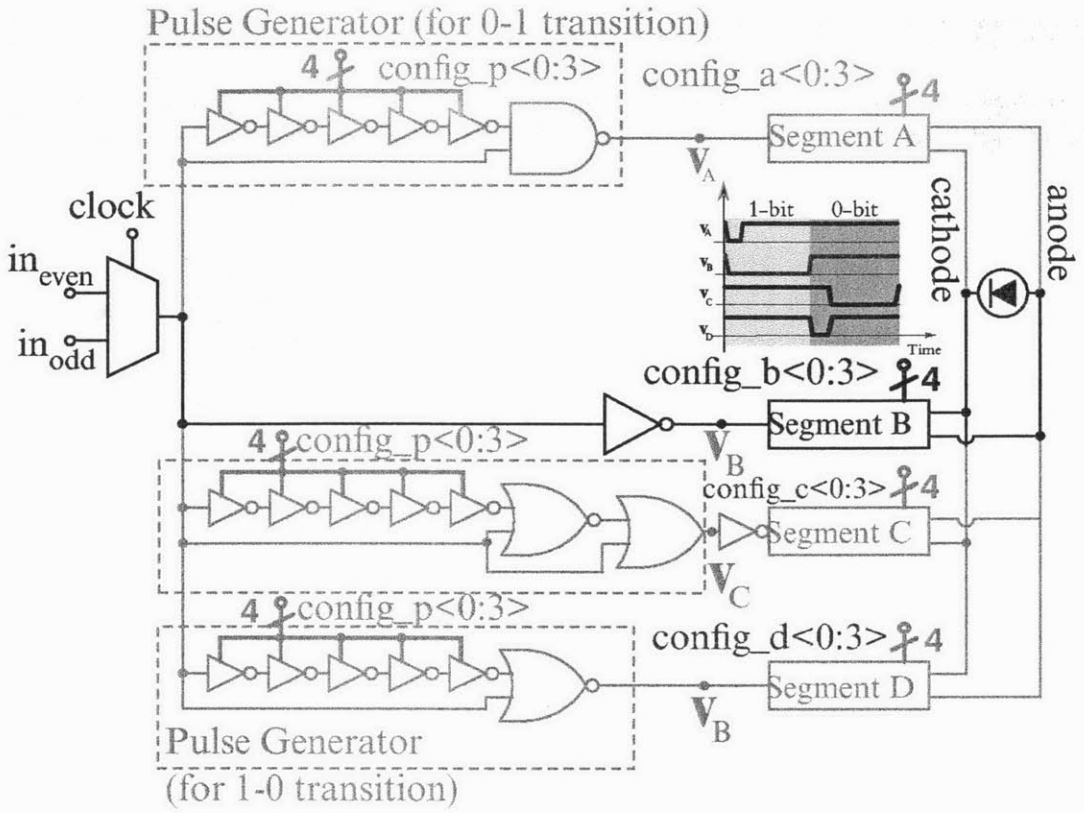


Figure 3-8: Conceptual schematic of the EOS-series carrier injection modulator driver. Segment B is active for the full duration of a 1-bit.

Figure 3-9 illustrates the beginning of a one-to-zero transition. Segment C is active, reverse-biasing the diode and sweeping the injected carriers out of the intristic region of the P-I-N diode for the first portion of the zero-bit.

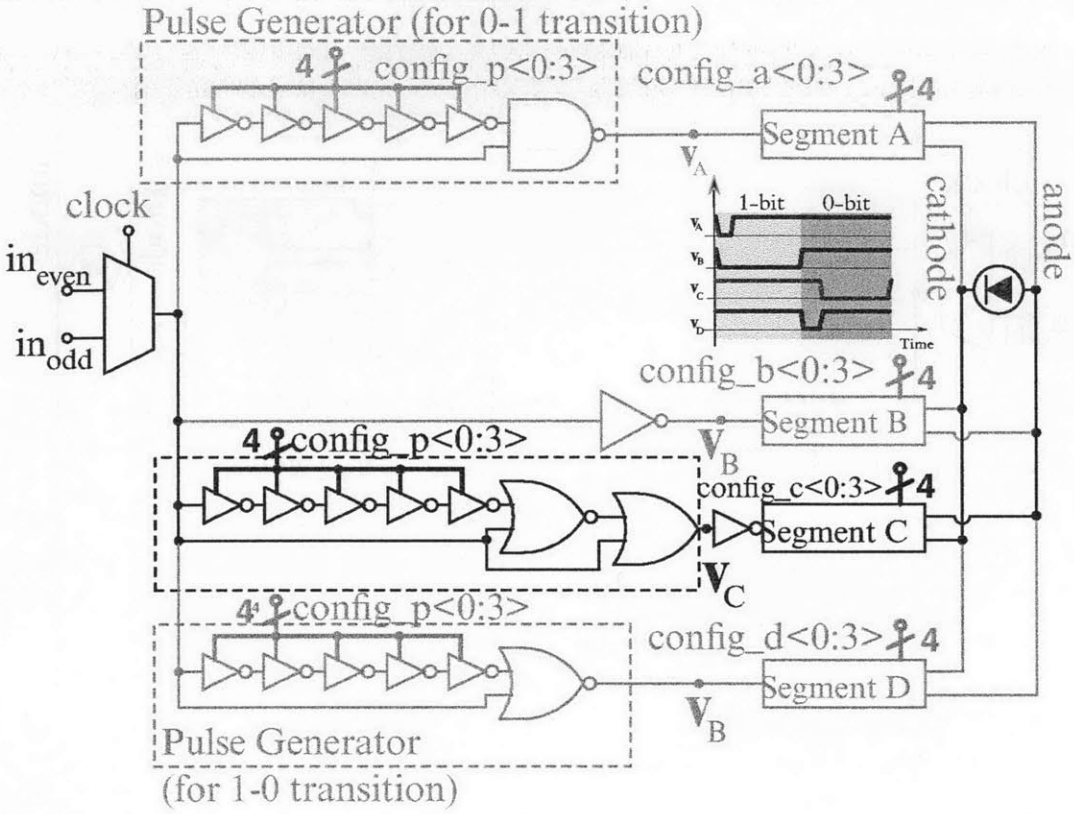


Figure 3-9: Conceptual schematic of the EOS-series carrier injection modulator driver. Segment C is active at the beginning of a 1-0 transition, reverse-biasing the diode and sweeping carriers out of the intrinsic region.

Figure 3-10 shows how Segment C deactivates, and Segment D activates, for the duration of a zero-bit after the pre-emphasis pulse. Reverse-biasing the diode at this stage would be the traditional method to use. However, note that Segment D is actually forward-biasing the diode, rather than reverse-biasing it. Segment D forward-biases the diode slightly below the diode's threshold voltage. This pre-charges the diode for a fast zero-to-one transition. It also takes advantage of the nonlinear relationship between the charge in the I-region and optical transmissivity. By forward-biasing the diode slightly below threshold, the charge inside the I-region is not sufficient to change the device's optical transmissivity; in a sense, it is below the optical threshold of the device. This novel technique was published in ISSCC [23].

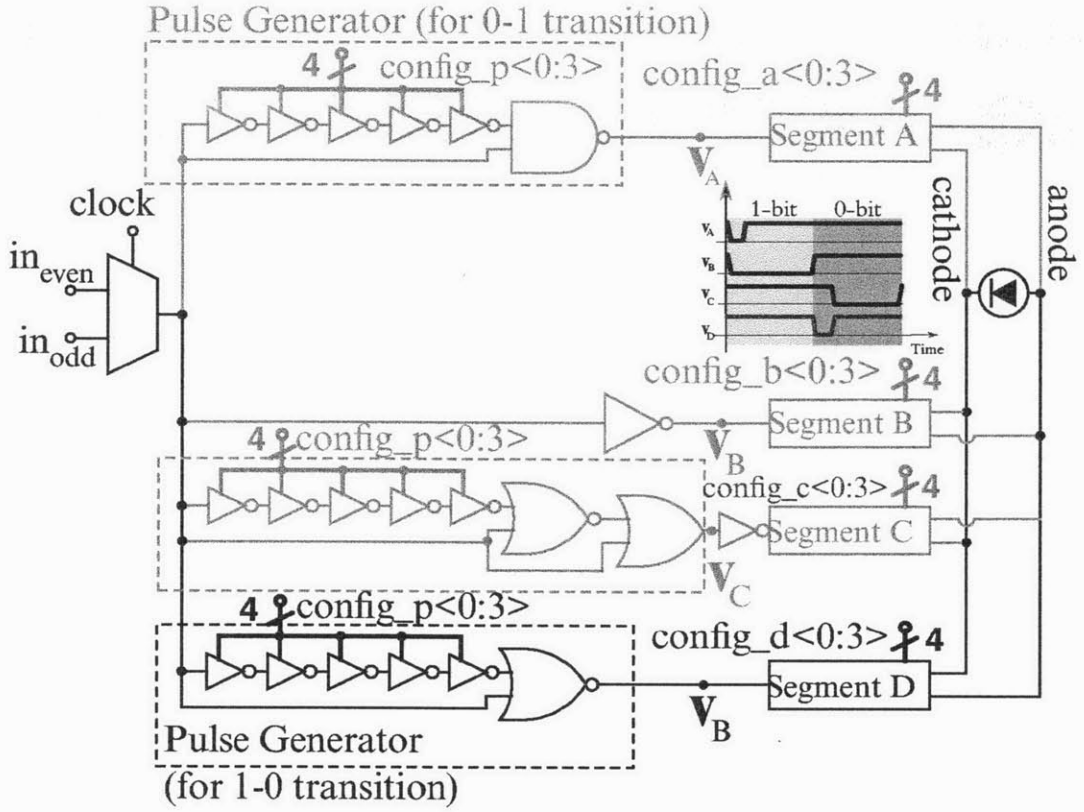
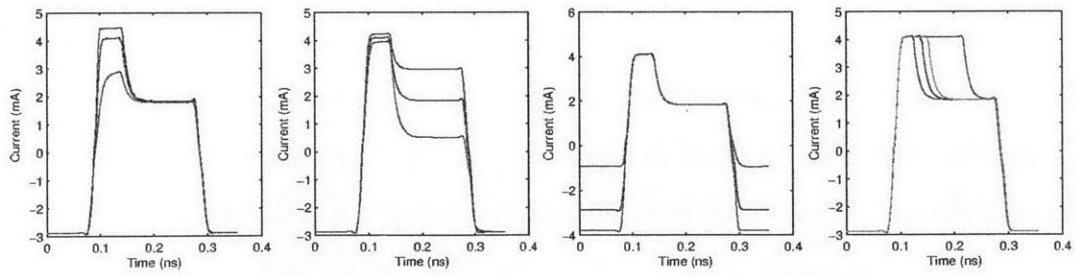


Figure 3-10: Conceptual schematic of the EOS-series carrier injection modulator driver. Segment D is active for the duration of a 0-bit after the pre-emphasis pulse has ended. The weak forward bias pre-charges the P-I-N diode without changing the optical logic state.

Each segment has been designed as a DAC, where control bits can configure the strength of the segment via the digital back-end. Some of the configurable elements of the pre-emphasis profile are seen in Figure 3-11. The unpredictable diode parameters such as carrier lifetime prompted a highly configurable design to account for process uncertainty, and the difficulty of coupling high-speed signals in the 1 to 10 GHz range on- and off- chip lead us to develop a sophisticated on-chip digital back-end infrastructure. The wide range of configurability will be able to adapt to the uncertain device properties of the P-I-N diode. The carrier lifetime of the poly-Si is unpredictable due to an unknown amount of defects and varies from fab to fab, so this parameter was difficult to estimate during design of the first generation of the circuit, but the minority carrier lifetime of the silicon estimated to be between the

ranges of 100 ps and 2 ns. It was later measured to be 1 ns. Simulations suggest the modulator can also work at 10 Gb/s if the minority carrier lifetime were shorter; the operation of the modulator depends heavily on the carrier lifetime. The driver currents for pre-emphasis, forward bias, and reverse bias can be tuned to within an order of magnitude. The duration of the pre-emphasis pulse can be tuned in the range of 10 ps to 100 ps.



(a) Pre-emphasis varying (b) Forward-bias varying (c) Reverse-bias varying (d) Pulse width varying

Figure 3-11: Configurable parameters of the P-I-N current profile.

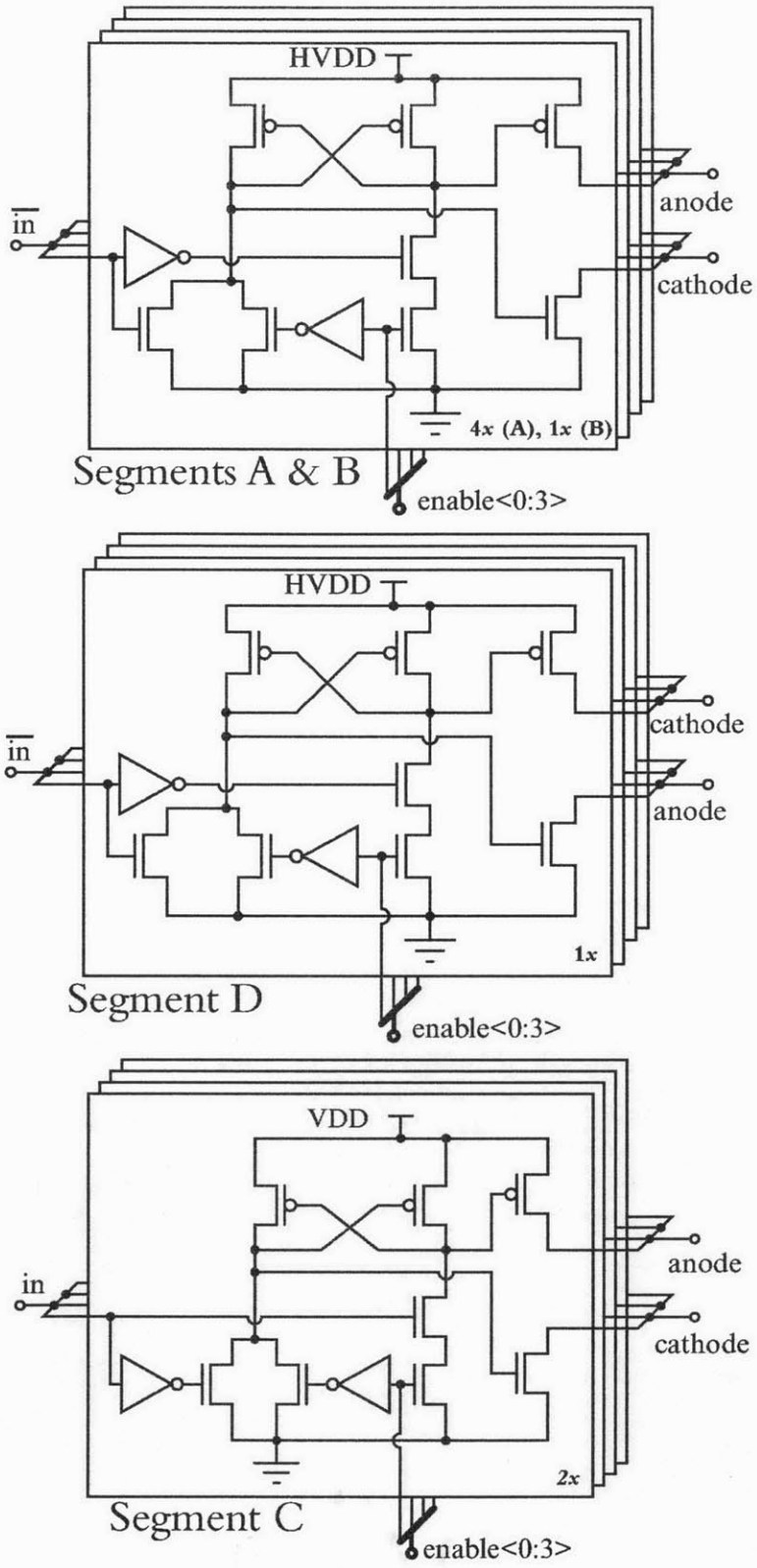


Figure 3-12: Carrier injection driver final DCVS driver heads

There are additional delay elements not shown in the schematic which can help tune the arrival of data transitions to the final stage to correct for manufacturing variation. These delay elements are also digitally configurable across four bits each. The delay element consists of a current-starved inverter, shown in Figure 3-13

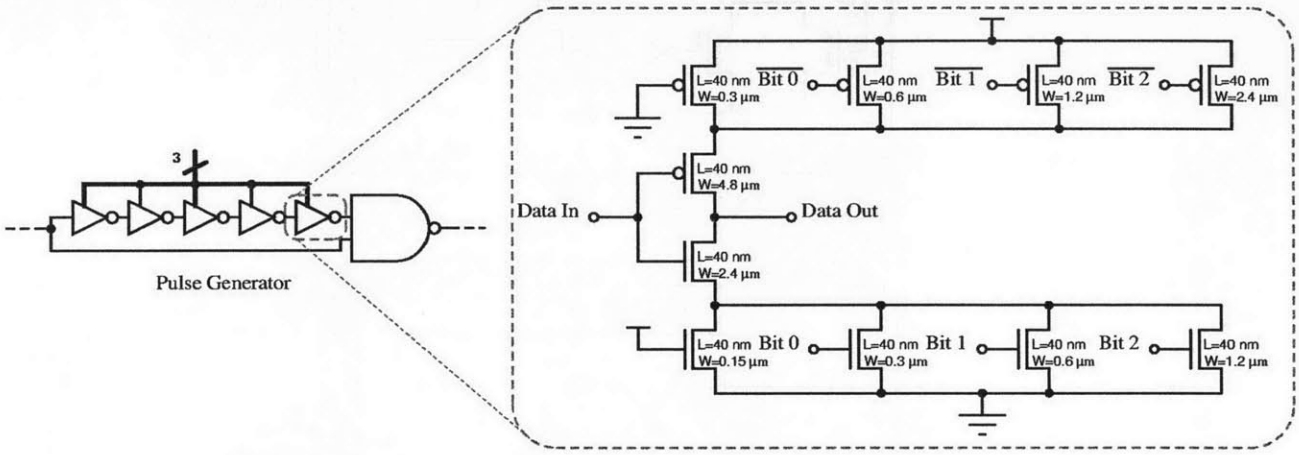


Figure 3-13: Delay element schematic

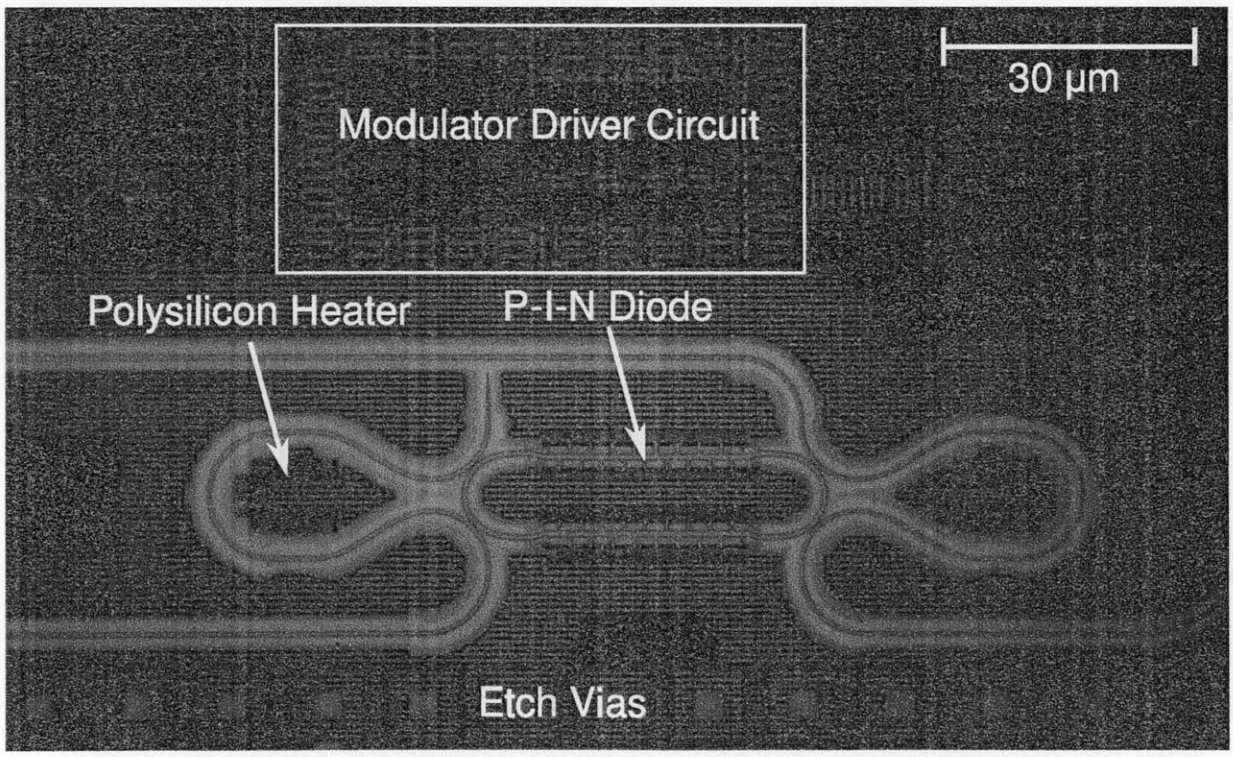


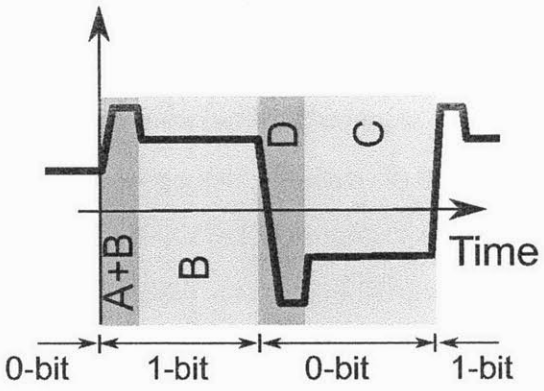
Figure 3-14: Micrograph of carrier-injection modulator driver and device.

The final stage of the EOS4 driver is a set of segmented differential cascade voltage switched (DCVS) NAND gates. These DCVS gates are connected to an isolated power rail, HV_{DD}, so that the drive voltage of the final stage of the driver can be increased above the process voltage of 1.0 V without risking damage to other transistors on the die. There are three sets of segmented DCVS NAND gates; one set is for the preemphasis, one set is for forward-biasing the diode for a full bit period, and the final set is for reverse-biasing the diode for a full bit period. The schematics for these forward and reverse-bias DCVS NAND gates is shown in Figure 3-12.

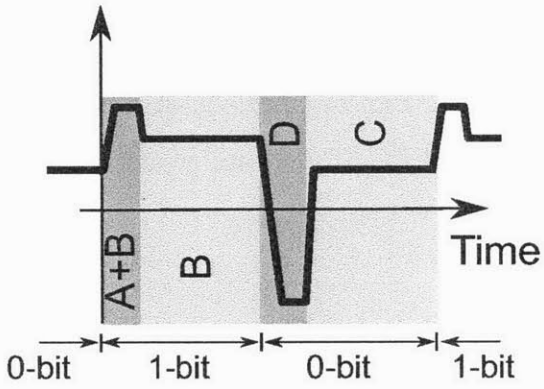
In addition to modulators, this chip also has configurable data receivers, configurable clock receivers, and data snapshot buffers. A serial scan chain allows the bits to be set and read easily by an FPGA or a similar external circuit. A micrograph of the modulator design is shown in Figure 3-14.

3.3.3 Advantages of Signal Conditioning

The signal conditioning technique mentioned in the previous section helps to overcome the inherent bandwidth limitation of the device. The device's electrical 3 dB bandwidth, shown in Figure 2-13 in the previous chapter, is limited to approximately 250 MHz due to the minority carrier lifetime of the device. A traditional "unconditioned" pre-emphasis profile is shown in Figure 3-16a. In the unconditioned profile, there is a strong forward-bias pre-emphasis at the beginning of a zero-to-one transition, and a strong reverse-bias pre-emphasis at a one-to-zero transition. After the strong reverse-bias pre-emphasis, a weaker reverse-bias persists for the duration of the zero-bit.



(a) Unconditioned

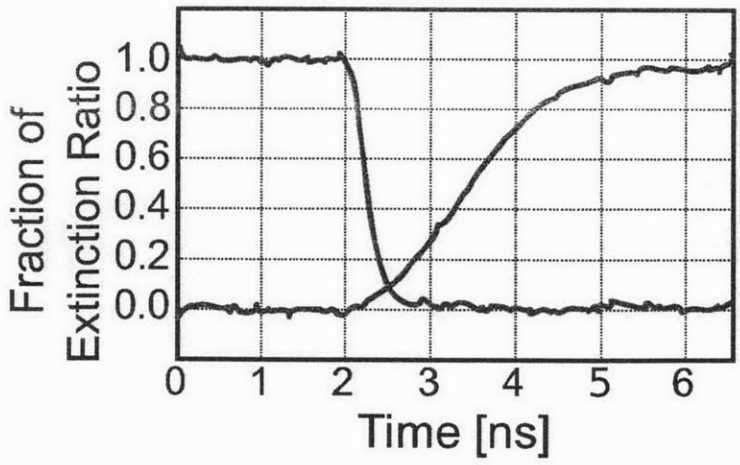


(b) Conditioned

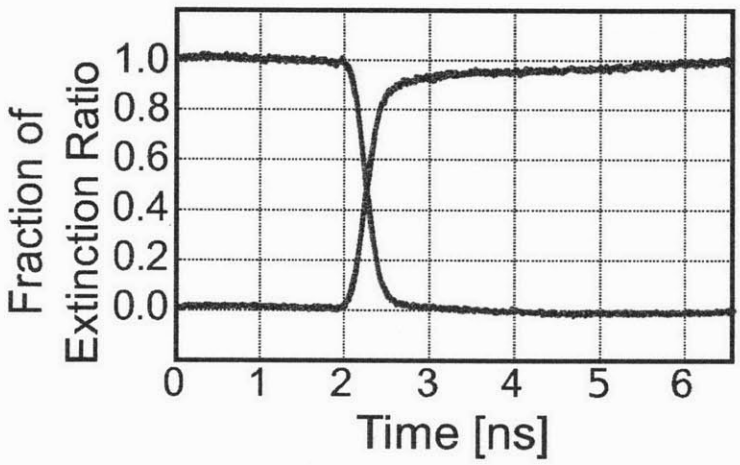
Figure 3-15: EOS12 pre-emphasis profile, unconditioned and conditioned

In the conditioned case, shown in Figure 3-15b, the profile differs for a zero-bit. After a one-to-zero transition, there is still a strong reverse-bias pre-emphasis to sweep the carriers out of the P-I-N diode's intrinsic region, snapping the ring back to its unbiased state. After the pre-emphasis pulse is over, the circuit slightly forward-biases the diode as a pre-charge for the next one-to-zero transition. This trick helps to overcome the bandwidth limitations on a zero-to-one transition caused by the long minority carrier lifetime.

The impact of this technique can be seen in Figure 3-16. The zero-to-one and one-to-zero step responses are measured and normalized in the Y-axis for fair comparison. In the unconditioned case, the zero-to-one transition is significantly slower.



(a) Unconditioned



(b) Conditioned

Figure 3-16: EOS12 Step Response, unconditioned and conditioned

In the conditioned case, the rising and falling edges are nicely matched, which will provide the most open eye and thus the optimal performance.

3.3.4 Experimental Test Setup

Our test station consists of an optical table with an infrared camera on a vertically-mounted microscope. Custom fiber holders have been machined and on each side of the die, a fiber holder sits atop a computer-controlled 5-axis stage. A second horizontally-mounted camera behind the die allows us to easily see how close the fiber is to the surface. The packaged die plugs into a custom circiuit board containing level

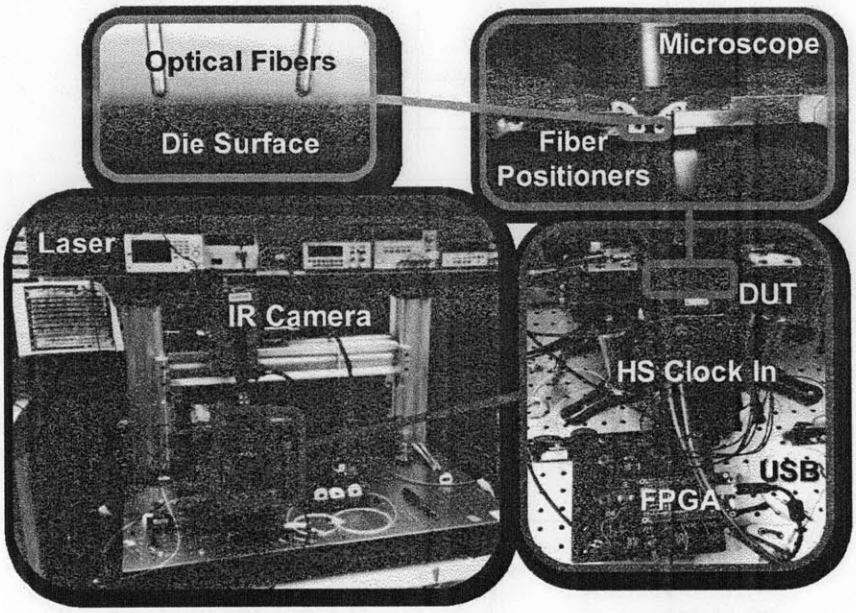


Figure 3-17: Table setup

shifters for the scan chain signals and current DACs to power the polysilicon heaters.

For experiments from chip-to-chip, a dual-table setup has been created with a trans-lab fiber running underneath the table.

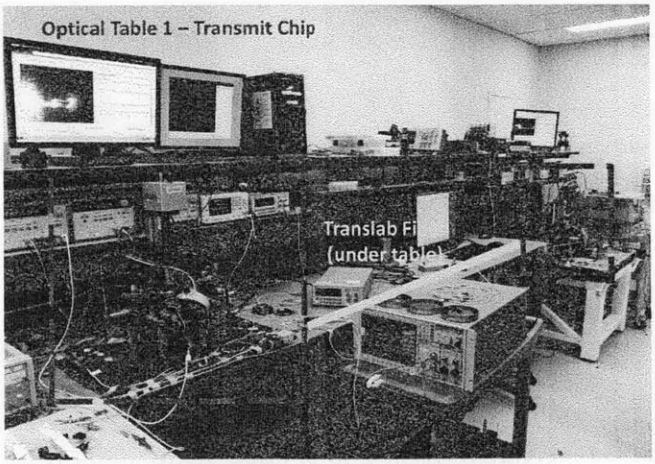


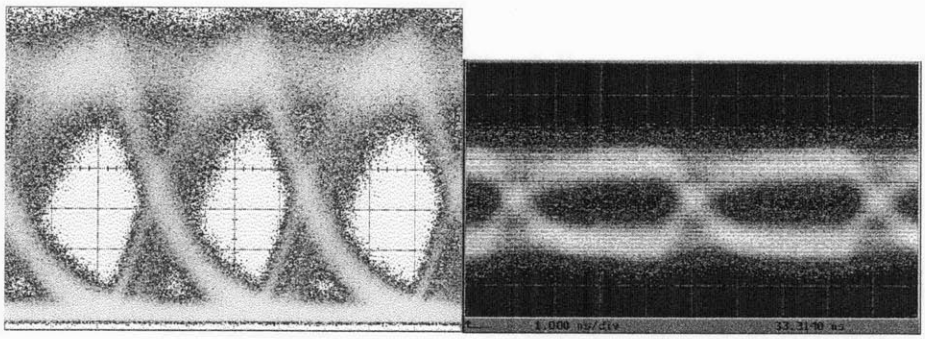
Figure 3-18: Dual-chip setup

3.3.5 Experimental Results

The most significant result thus far is a demonstration of working modulator devices in both the 1550 nm band and the 1280 nm band. The optical eye diagrams shown in the following subsections demonstrate the first measured monolithically integrated modulator driver in a convenitonal CMOS process. The speed of the driver, and thus the eye width, is limited largely by the carrier lifetime. The eye height for 1280 nm results are worse than comparable 1550 nm results due to the lack of an optical amplifier in the 1280 nm band. The noise of the optical oscilloscope significantly worsened the eye height as well.

Without Signal Conditioning

Previous generations of the carrier-injection modulator driver without signal conditioning reached maximum speeds of approximately 600 Mbps and were published in Optics Express 2012 [25]. The eye diagrams at 1550nm (reaching 600 Mbps) and 1280nm (reaching 250 Mbps) are shown in Figure 3-19. The only reason the 1550nm result is at a higher datarate is because we were able to supply more optical power at that band, which resulted in a more open eye.



(a) 1550 nm eye diagram at 600 Mbps (b) 1280 nm eye diagram at 250 Mbps

Figure 3-19: Optical eye diagrams of the injection-based modulator driver.

The measured electrical power and extinction ratio of the modulator without signal conditioning is shown in Figure 3-20. This generation of the design can achieve an open eye at 250 Mbps at approximately 4pJ/bit at an extinction ratio of around

2.3 dB.

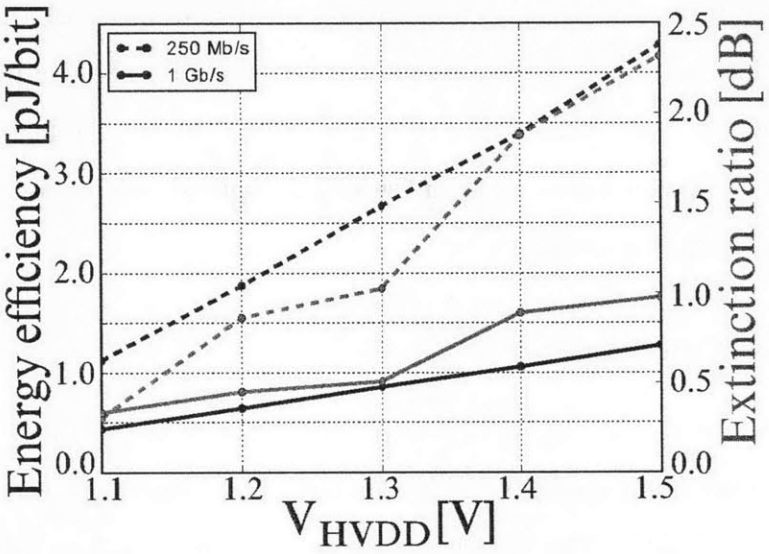


Figure 3-20: Measured electrical power of 1280 nm injection-based modulator

With Signal Conditioning

With the signal conditioning described in this section, performance increased considerably. Virtually the same circuit, with only minor changes to implement the improved pre-emphasis profile, was able to boost performance to the multi-gigabit regime, reaching 2.5 Gb/s, shown in Figure 3-21. This result was published in ISSCC 2013 [23]. This result is the fastest, most energy-efficient carrier-injection modulator in a sub-100nm monolithically-integrated CMOS process to date.

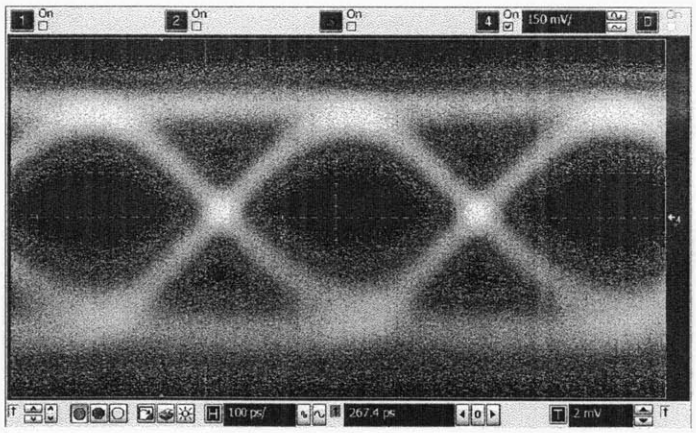


Figure 3-21: Optical 2.5 Gb/s eye diagram of signal-conditioned EOS carrier-injection modulator at 1550 nm

The energy-efficiency and extinction ratio achieved by the signal-conditioned EOS carrier-injection modulator is shown in Figure 3-22. This iteration of the design achieved 1.23 pJ/b at an extinction ratio of 3.4 dB at 2.5 Gb/s.

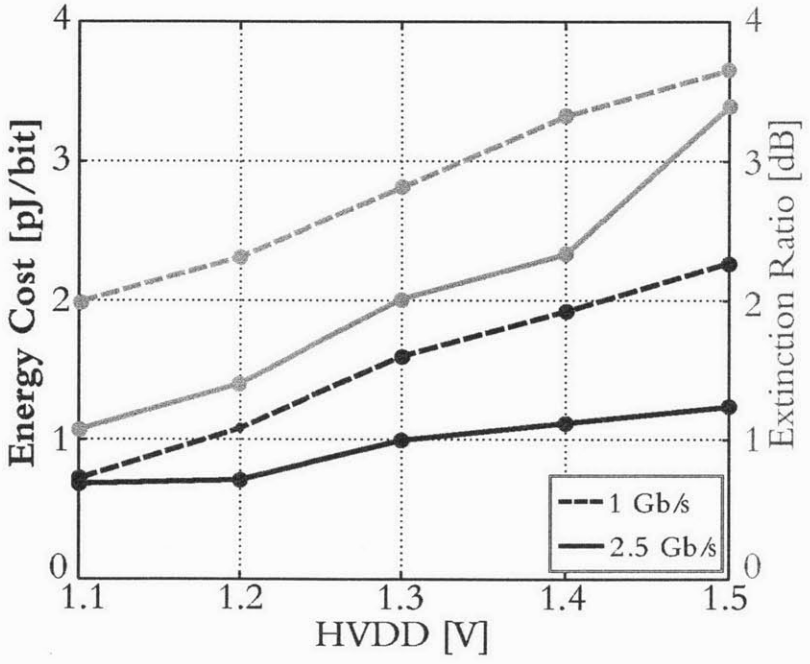


Figure 3-22: Measured electrical power of signal-conditioned EOS carrier-injection modulator at 1550 nm

The energy-efficiency improves with speed because the on-current through the

forward-bias diode gets amortized over a larger number of bits.

This concludes the analysis of the carrier-injection modulator.

3.4 Depletion-Mode Modulator Drivers

The carrier-injection modulator in the previous section has two main drawbacks. First, its energy efficiency is low due to the need to forward-bias the P-I-N diode for each one-bit. Secondly, the minority carrier lifetime of the diode limits the speed of the device to low-Gb/s speeds.

The depletion-mode modulator device does not have these drawbacks. Rather than forward-biasing the diode structure to inject free carriers, the diode structure is reverse-biased to increase the size of the depletion region. In contrast to the carrier-injection device, the depletion-mode device is more sensitive to the process doping levels and material properties, and depending on the exact device behavior, may require higher reverse-bias voltages (beyond the native process V_{DD}). However, as a topology, the depletion-mode modulator has far more potential in terms of speed and energy-efficiency than the carrier-injection approach.

3.4.1 Design Constraints

From the analysis in Chapter 2, we concluded that a reverse-bias drive voltage of $V_a = 2.0$ V is necessary to reach the required ΔQ for a 1-0 transition (in the lab, the device performed better than expected and its performance is satisfactory at 1.5 V). For a 0-1 transition, the driver circuit also needs a forward-bias conduction path. In a modern CMOS process, V_{DD} is limited to 1.0 V, and adding an extra supply rail is impractical, leading us to consider charge-pump based circuits.

Additionally, to avoid reducing the lifespan of the driver circuit through oxide breakdown, each transistor must satisfy:

- $V_{ds} \leq 1.0$ V
- $V_{gd} \leq 1.0$ V

- $V_{gs} \leq 1.0$ V

This design constraint can be mitigated by using thicker-oxide devices for the final drive transistors; however, these transitors may impose a speed constraint. Two driver designs are analyzed; first, a charge-pump-based approach, for reverse-bias drive voltages above 1.5V, and second, a simpler DCVS NAND-gate approach with thick-oxide final stage transistors.

3.4.2 Processor (EOS) Platform: Charge-Pump Carrier-Depletion Driver Circuit

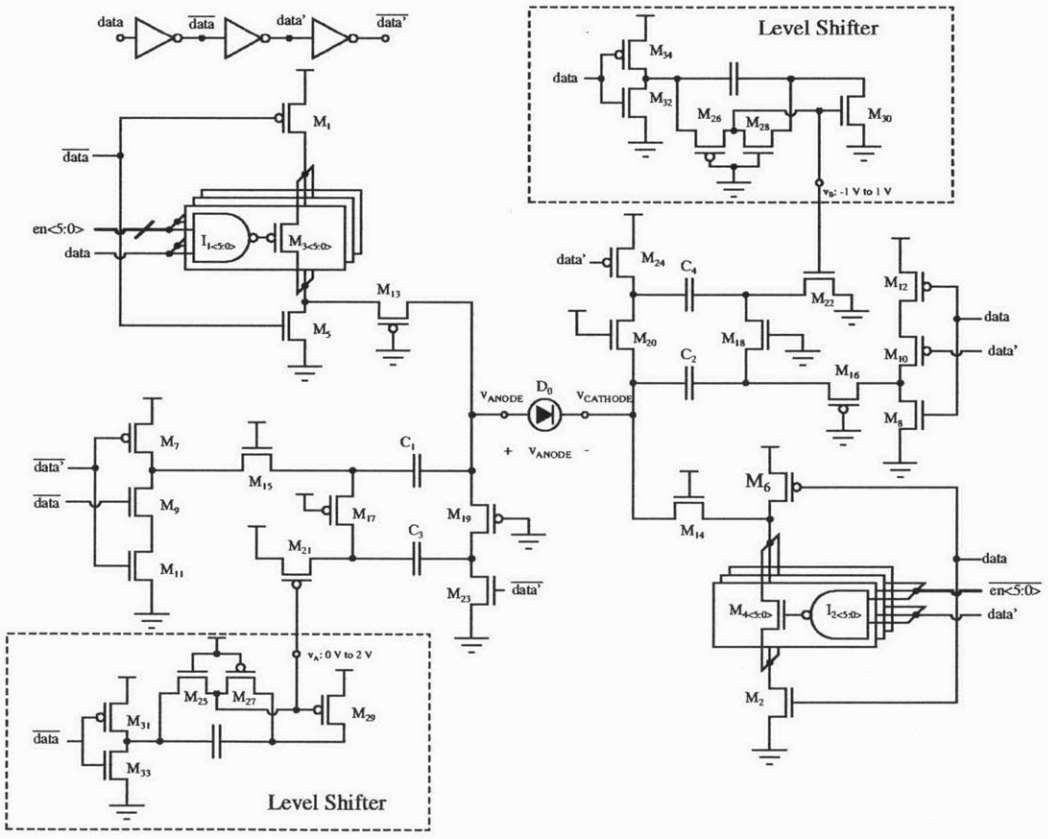


Figure 3-23: Depletion-mode driver schematic

The first depletion-mode driver design is a charge-pump circuit intended for modulator devices requiring a drive voltage much greater than the native process VDD. The full schematic for the depletion-mode driver is shown in Figure 3-23. An inverter

chain, shown at the top of the schematic, is used to generate auxiliary data signals used at various points throughout the circuit: $\overline{data}$ represents a slightly delayed, inverted data, data' is a delayed version of data, and so forth. This circuit has two modes of operation depending on whether the current bit (data on the schematic) is a 1 or a 0.

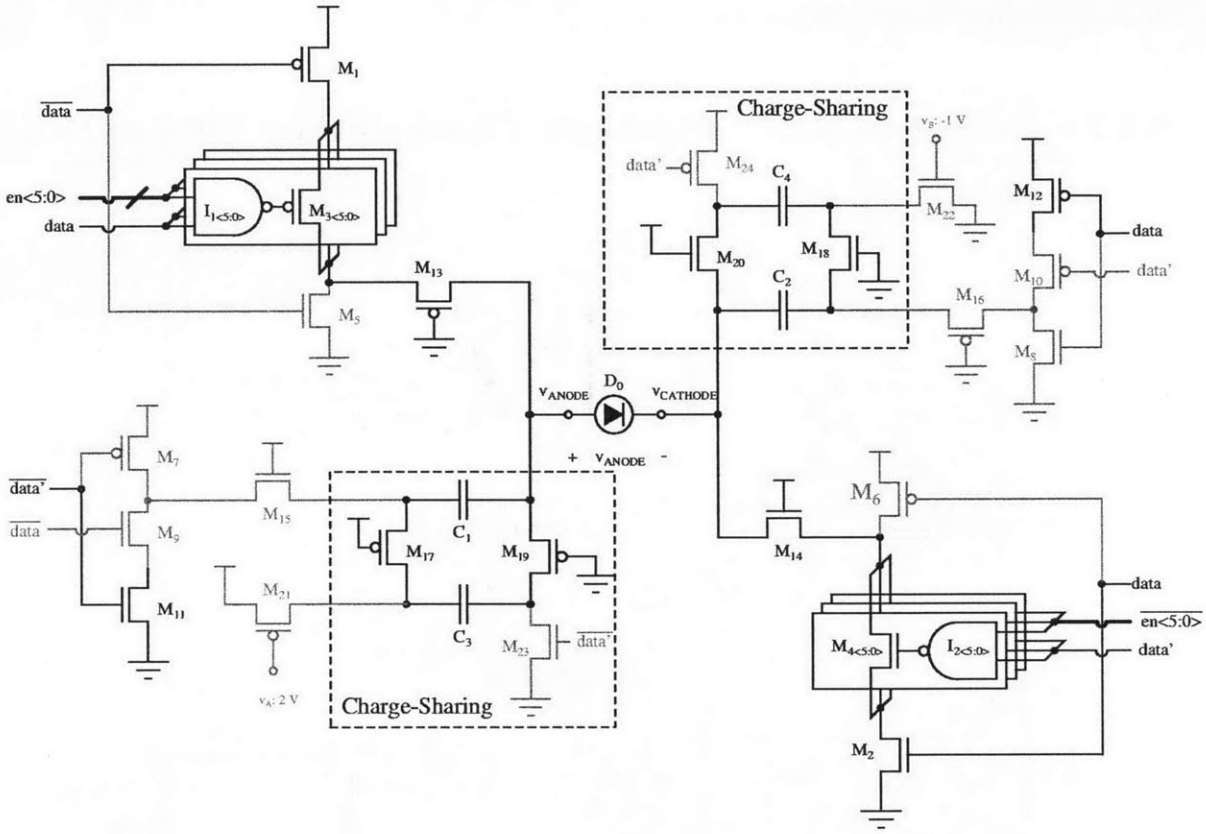


Figure 3-24: Depletion-mode driver schematic - 1 bit - forward bias

If the current bit is a 1, two things are happening simultaneously. The modulator diode is being forward-biased thorugh M_1, M_3, M_{13}, M_{14}, M_4, and M_2. The amount of forward-bias current that the driver can sustain can be digitally adjusted by the six configuration bits, labeled as en<5:0> on the schematic. At the same time, capacitors C_1 and C_3, and capacitors C_2 and C_4 are connected in parallel to share charge. Both C_4 and C_3 were pre-charged to 1.0 V in the circuit's other state, to be described shortly. In its steady state of operation, and after charge-sharing, both C_1 and C_2 will have reached nearly 1.0 V. This state is shown in Figure 3-25. Inactive transistors

are greyed out.

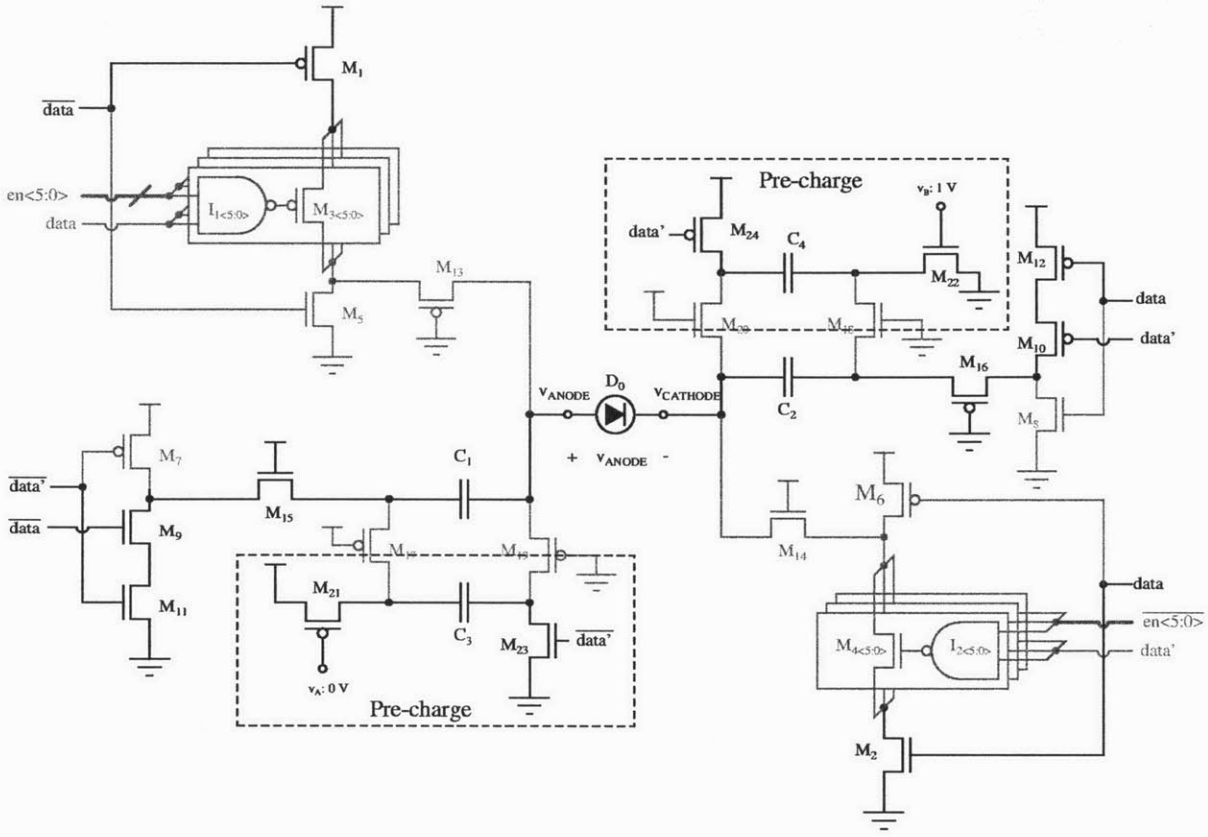


Figure 3-25: Depletion-mode driver schematic - 0 bit

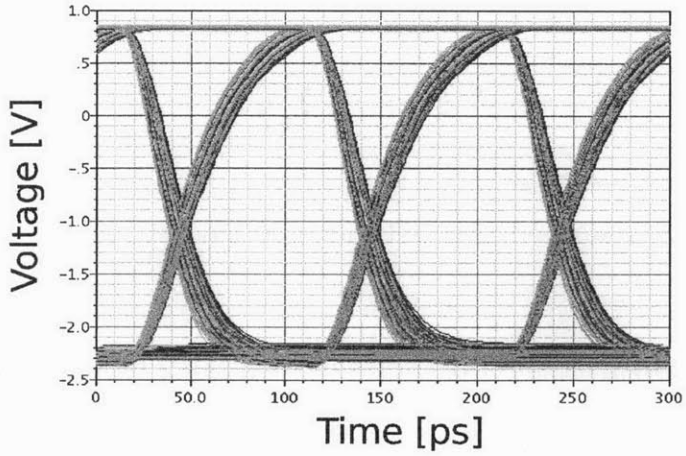


Figure 3-26: Simulated depletion-mode modulator driver eye diagram at 10 Gb/s, Monte Carlo simulation on a pseudo-random data sequence

If the current bit is a 0, again, two processes happen simultaneously. C_1 and C_2 had previously been charged to 1.0 V in the circuit's other state of operation. Now they are used to reverse-bias the diode, as shown in Figure 3-25. C_1 is pushed below M_{15} to pull node v_{ANODE} to -1.0 V. Similarly, C_2 is stacked on top of M_{16} to push node $v_{CATHODE}$ to +2.0 V. At slow speeds, this circuit will reverse-bias the modulator diode at voltages above 2.5 V. At its maximum speed of 10 Gb/s, the circuit can reverse-bias the diode near 2.0 V. A simulated eye diagram is shown in Figure 3-26.

The two level shifters shown in the full schematic (Figure 3-23) are necessary to bias transistors M_{21} and M_{22} to satisfy the oxide-breakdown voltage constraints.

3.4.3 Processor (EOS) Platform: DCVS-NAND Thick-Oxide Carrier-Depletion Driver Circuit

The second carrier-depletion design topology is intended for modulator devices which can operate with reverse-bias voltages in the 1.0 to 1.5 V range (at or slightly above the process VDD). This transmitter circuit shown in (Figure 3-27 consists of a carrier-depletion optical ring resonant modulator and a 65m x 30m driver circuit.

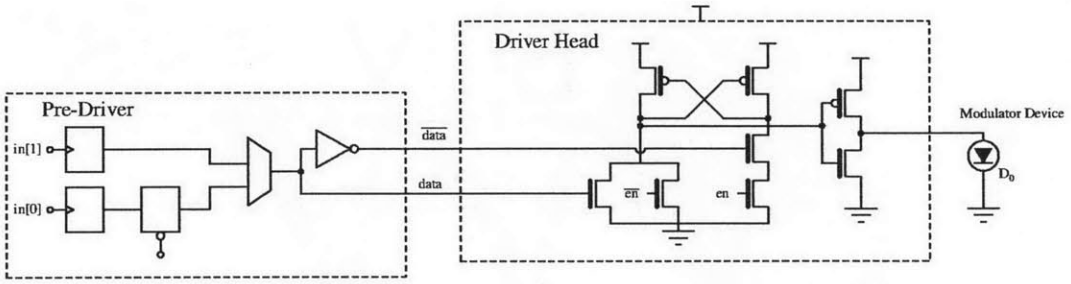


Figure 3-27: DCVS-NAND Thick-Oxide Carrier-Depletion Driver Circuit

In the driver, a 2-to-1 serializer performs muxing of a 2-bit input digital bitstream, a DCVS NAND-gate level-shifts the bits from VDD to the HVDD, and a final inverter in the HVDD domain drives the cathode terminal of the modulator diode. The diode anode is shorted to VSS, and the circuit applies a voltage of -HVDD to 0V across the diode for carrier-depletion operation. Based on the voltage swing requirements of

our previously demonstrated devices [32], we built both the DCVS NAND-gate and final inverter using thick-oxide transistors. Though slower, this allows HVDD to be increased well above the nominal process voltage of 1V - a precautionary measure taken to accommodate depletion modulator designs that often require higher driver voltages [5] [43].

The optical modulator device is a lateral junction modulator [4], with interdigitated junction definition enabled by the advanced 45nm process lithography. The modulator has a free-spectral-range of 2.74 THz and optimized coupling resulting in a measured Q-factor of approximately 14,000. Given such high Q, we note that above-nominal supply voltages are unnecessary for this improved device; a swing of just -0.7V to 0V captures greater than 8dB of extinction. Thus, both the thick-oxide devices and the level-shifter can be eliminated in the future to improve speed and energy.

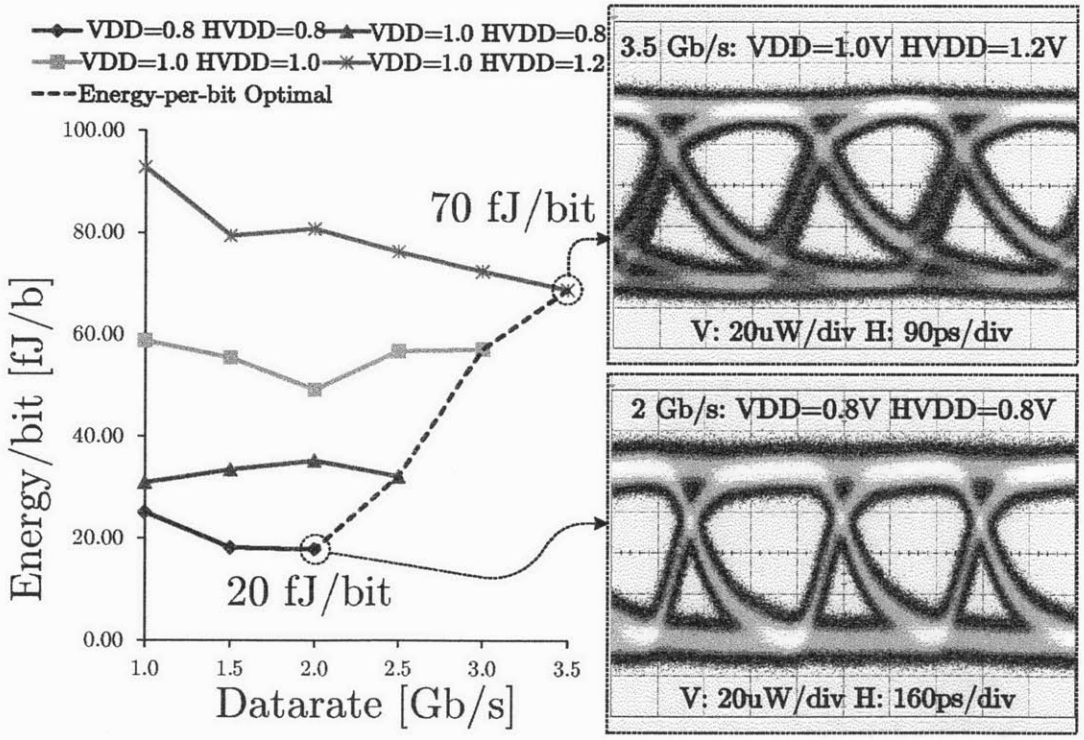


Figure 3-28: DCVS-NAND Thick-Oxide Modulator Energy and Eye Diagrams

Figure 3-28 shows modulator eye diagrams taken at 2 Gb/s and 3.5 Gb/s. The

centers of the eyes are bit-error free for 1012 bits, verified using an external TIA-to-FPGA-based bit-error-rate test. A layout error resulting in delay-imbalanced 0-to-1 and 1-to-0 paths through the thick-oxide DCVS NAND gate causes the distorted eye shape and ultimately limits the maximum achievable data-rate of the driver. Higher HVDD is required at higher data-rates solely for improving rise and fall times limited by the slow thick-oxide devices in the driver circuit. With appropriate supply scaling, the energy-cost of the driver scales gracefully from 20 fJ/bit at 2 Gb/s (VDD of 0.8V and HVDD of 0.8V) to 70 fJ/bit at 3.5 Gb/s (VDD of 1.0V and HVDD of 1.2V). The depletion modulator's reverse-bias operation and digital driver implementation result in flat energy-cost vs. data-rate, and provide more than an order of magnitude better energy-efficiency compared to the forward-bias injection modulator demonstrated previously in the same process [23].

3.4.4 Memory (Micron) Platform: Carrier-Depletion Driver Circuit

The circuit for the Micron platform, shown in Figure 3-29 does not require the thick-oxide transistors necessary in the EOS platform due to the higher native process VDD. For this reason, a voltage-limiting additional NMOS transistor is added to the final stage.

The backend interfaces with the custom TX and RX heads through 8-to-2 mux/demux tree SerDes and runs at one-fourth the data clock. The custom transmit head performs the final 2-to-1 serialization, and is designed to interface with a variety of integrated photonic modulators.

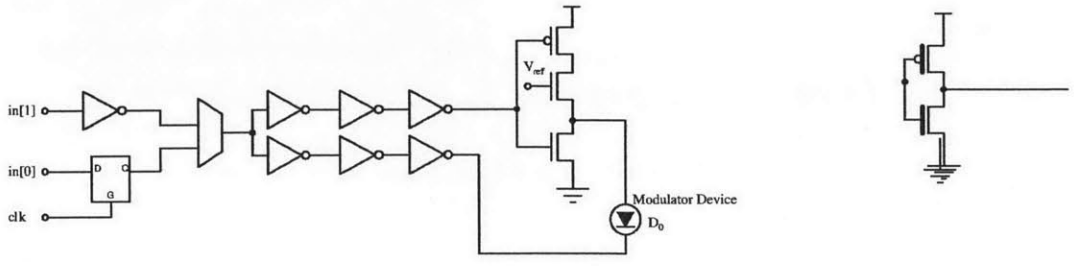


Figure 3-29: Micron modulator driver schematic.

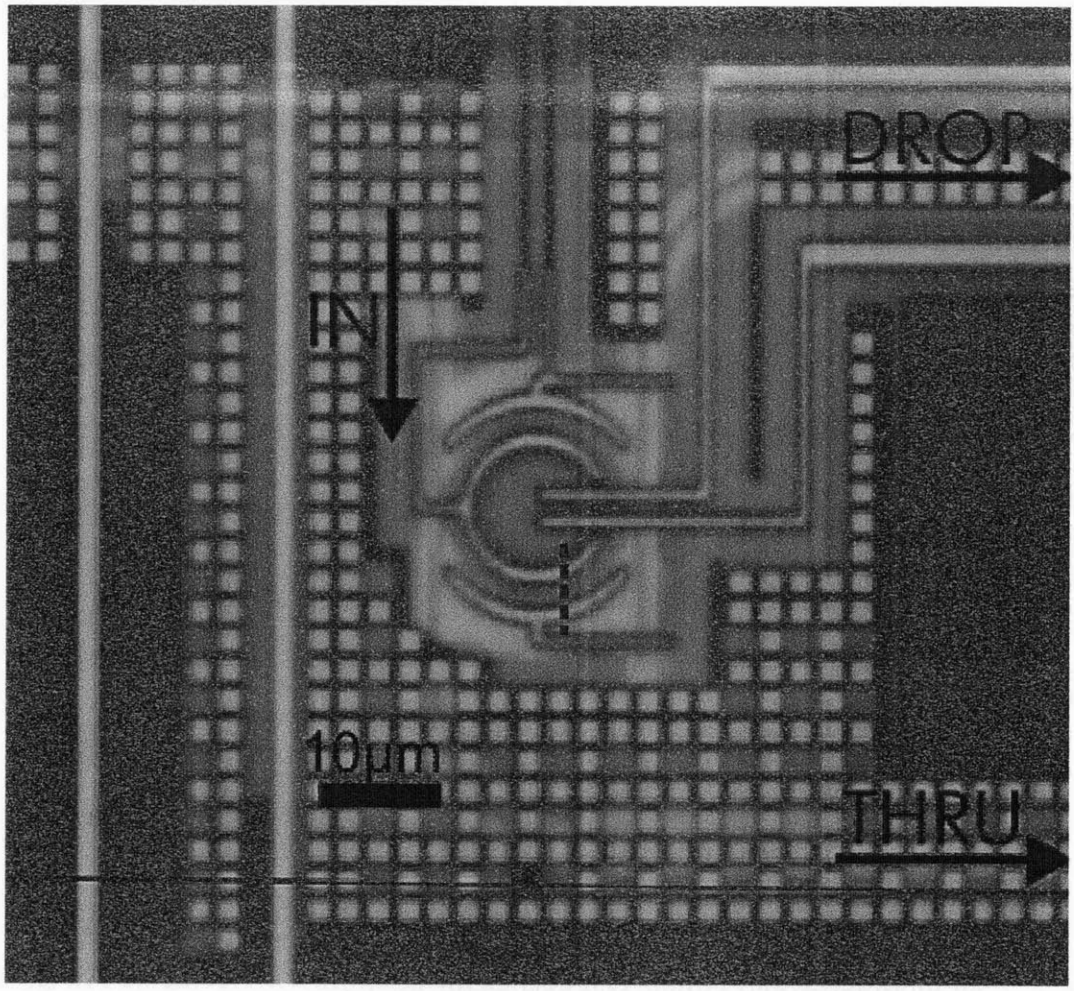


Figure 3-30: Micrograph of Micron modulator device.

The DDR modulator circuit is a push-pull diode driver with an NMOS pull-up on the device anode to limit the forward-bias voltage. The depletion-mode ridge microring modulator (shown in Figure 3-30 has simple geometry that can be electrically contacted without introducing loss to the optical mode. It is created via a partial polySi etch and doped to form a pn-junction across the ridge. The circuit drives the junction to -VDD for a full-depletion logic 1 and VREF-VT for a weak forward-bias 0, modulating the depletion region width. This creates a shift in resonance from the carrier-plasma effect and modulates the input wavelength.

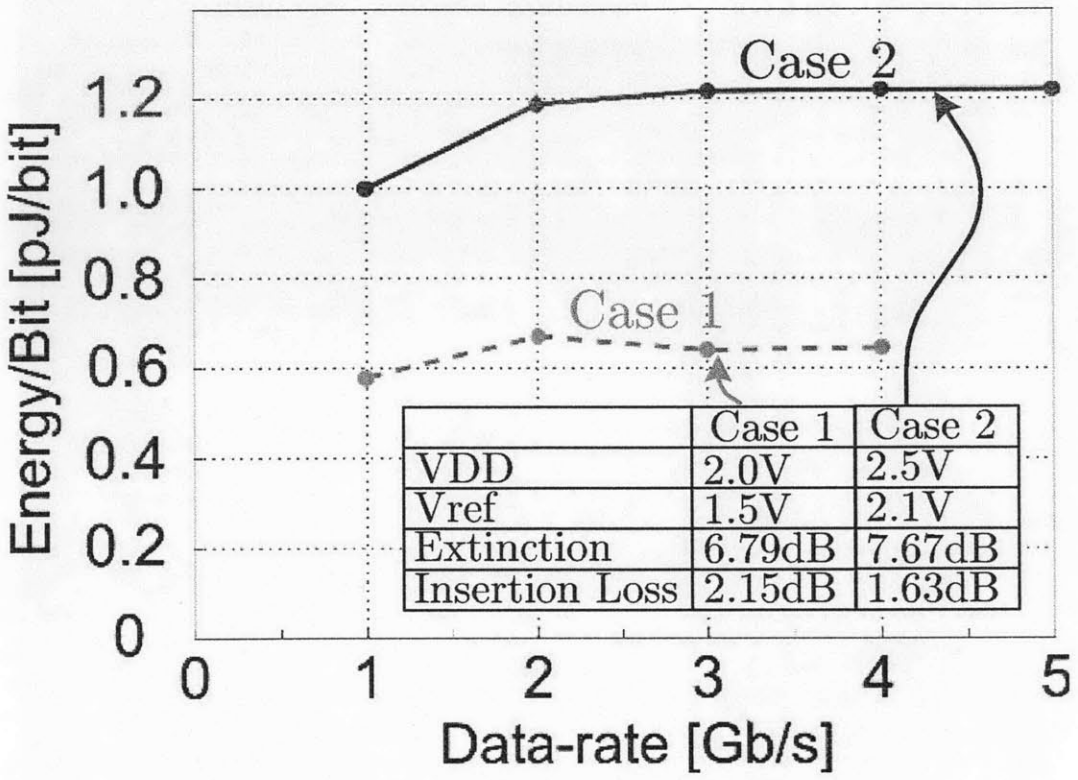


Figure 3-31: Energy-per-bit of the Micron modulator driver.

The ridge modulator and circuit energy-per-bit vs. data-rate tradeoffs are shown in Figure 3-31. To our knowledge, this chip demonstrated the first chip-to-chip monolithically-integrated photonic link in a bulk CMOS process, demonstrating that photonic interconnects need not be confined to niche, high-cost processes. Dense and energy-efficient photonic interconnects are feasible even on cost-aware mainstream bulk CMOS platforms.

3.5 EPHI

Although the focus of this thesis is on monolithically-integrated silicon photonics, in parallel, we taped out a chip in a hybrid flow where a photonics chip is wafer-level-bonded to a dedicated CMOS chip, named EPHI.

During EPHI chip design, a family of modulator drivers were created to cover a wide range of process uncertainties at design time, to provide drive functionality for a

family of modulator devices, and also to provide backup testing capability for various testing scenarios.

The primary modulator driver design is a 50μm x 50μm double-data-rate (DDR) driver circuit, shown in Figure 3-32. The driver is connected to an optical modulator device. The optical modulator device is a lateral junction modulator, leveraging the Soref electro-optical effect in silicon to electrically control the resonant frequency of the modulator rings resonance. The driver circuit moves the optical modulators resonance into and out of the laser wavelength channel to modulate the light. The driver circuit modulates the modulator diode depletion width by varying the reverse-bias voltage across the devices diode contacts; through the Soref electro-optic effect, this depletion width change causes an optical resonance shift. This driver circuit is designed to drive a wide family of optical devices. The final stage consists of a level-shifting inverter on the modulator devices cathode, where the gate of the NMOS pull-up device is connected to a voltage reference circuit (labeled VREF). The modulator diode junction is driven to VREF-Vt for a weak forward-bias, providing additional modulator extinction by further shrinking the junctions depletion region. During full depletion, the modulator diode junction is reverse-biased to VDD.

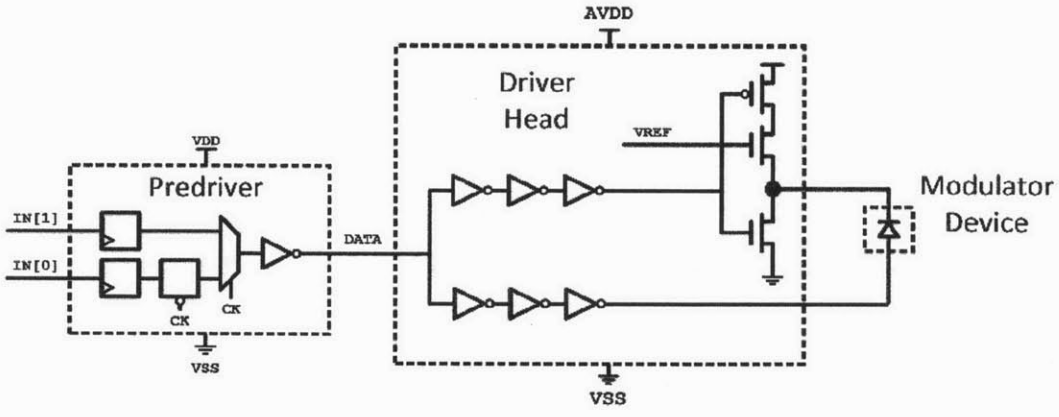


Figure 3-32: The EPHI DDR modulator driver

The modulator is fed by the digital back-end, where two PRBS generators operating at 2.5 Gb/s provide two data streams, shown in Figure 6 as IN[0] and IN[1]. A pre-driver block multiplexes these two data streams into a single 5 Gb/s data stream,

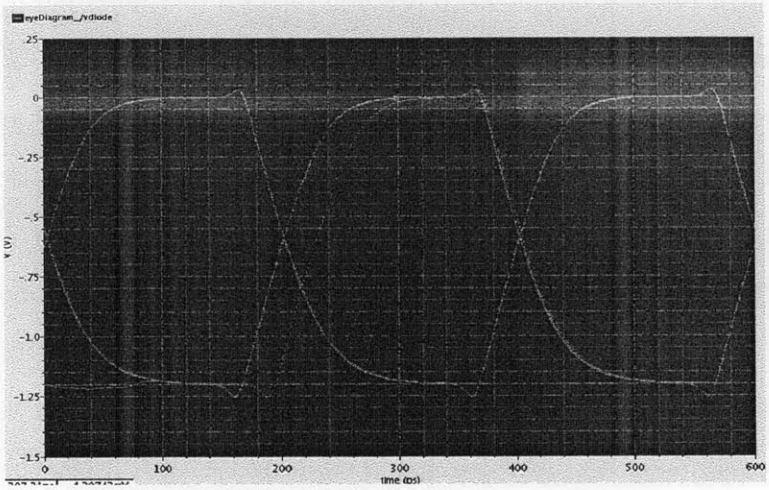


Figure 3-33: Extracted eye diagram of the modulator driver output (with 20fF load) at 5 Gbps.

carefully controlling the clock domains through the use of two flip-flops and a latch. The predriver block also buffers these signals to ensure that the fan-out is appropriate to drive the final driver head.

The layout of the modulator circuit is shown in Figure 8. The 50μm x 50μm layout consists of approximately 60% decoupling capacitors. A simplified modulator driver variant was also created, shown in Figure 3-35. This driver does not have the ability to drive the modulator junction into a weak forward bias, so the optical extinction will suffer; however, if the optical device has a high extinction, this design will be more energy-efficient and the lower extinction may not matter. In this variant, the anode of the modulator device is connected to VSS, and the cathode is connected to a buffered inverter standard-cell. This design is very simple and robust and should be able to drive modulator devices with capacitances in the range of 20-40fF at 5 Gb/s.

A third variant, an ultralite modulator driver, is shown in Figure 3-36. This design is simply a 4x-sized inverter, with an input capacitance of approximately 2.4 fF. This driver is a more extreme design because the final inverter will only be able to drive modulator devices at 5 Gb/s if the total device capacitance is under 10 fF, which is below expectations. However, if the device capacitance is low, this will be an extremely low-power design. The final inverter is connected to a separate power grid

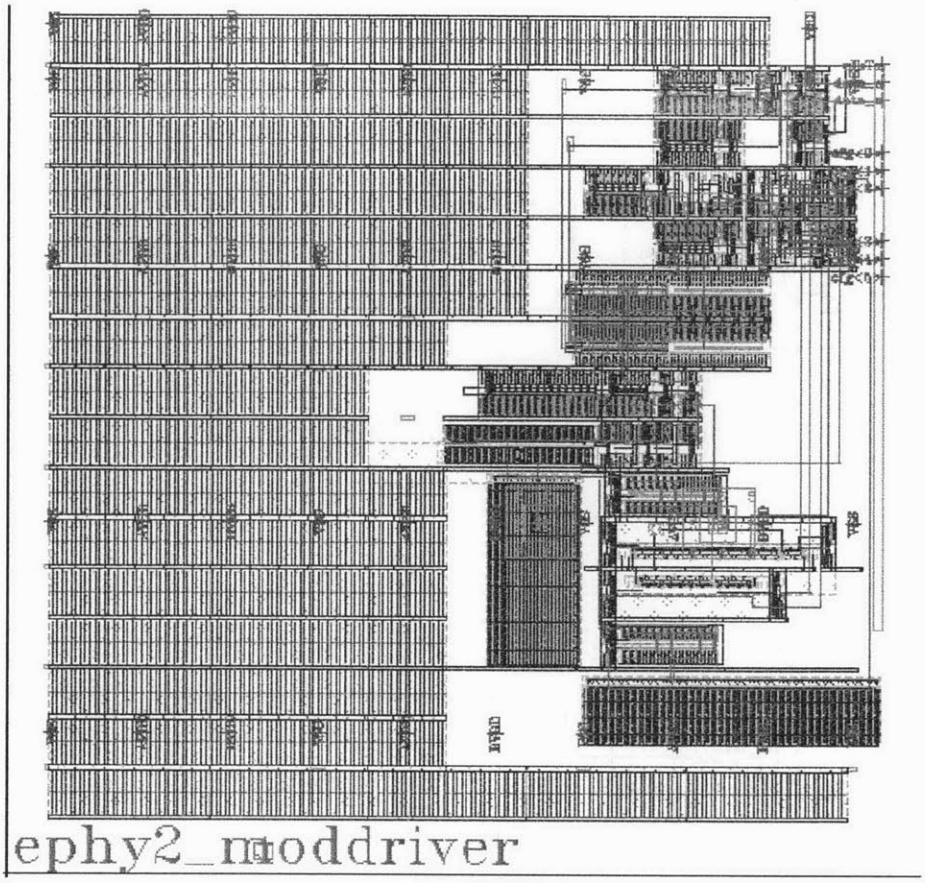


Figure 3-34: The 50 μm x 50 μm modulator driver.

(AVDD) from the pre-driver so that the final drive power can be directly measured, separate from the pre-driver and digital backend.

The digital backend is a complex place-and-routed block, and although it is verified through simulation and other verification tools, it does rely on high transistor yield to operate properly. Given that the process is under development, backup designs that can operate independently from the digital backend were created so that the modulator driver can be tested in the event of a digital backend failure. Figure 3-37 shows the schematic for the independent backup modulator driver. The final drive head, described earlier, is now fed by a CML-to-CMOS converter rather than the digital backend and predriver. The CML (current-mode-logic) inputs are fed directly from GSG pads. Low-voltage-swing CML inputs are used to boost the operating speed to 5 Gb/s. The three-stage CML-to-CMOS converter converts these 200mV

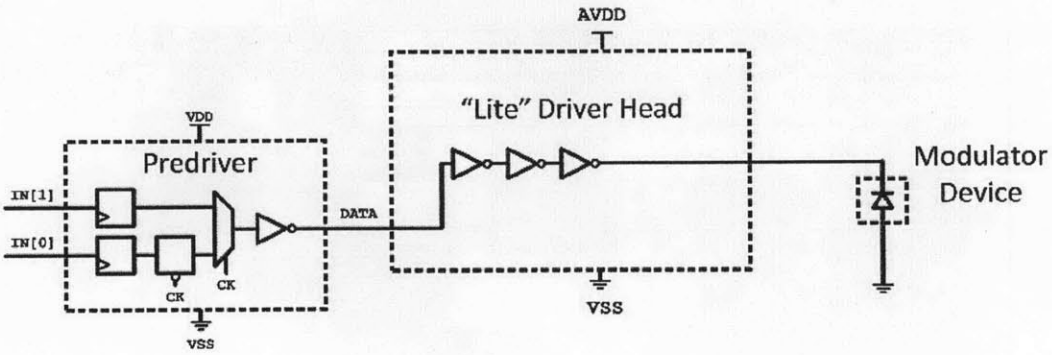


Figure 3-35: Simplified modulator driver.

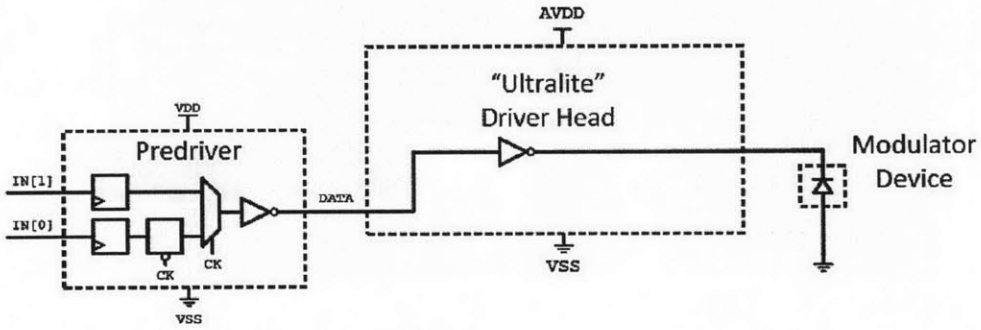


Figure 3-36: "Ultralite" modulator driver.

differential swings into a single full-rail CMOS swing at the date output. The CML-to-CMOS final stage is buffered appropriately to directly drive the final drive head without an additional pre-driver.

This same backup approach of attaching a CML-to-CMOS converter is also used for the "lite" and "ultralite" variants, as shown in Figures 3-39 and 3-40.

The Ephi modulator driver designs, from the perspective of the focus of this thesis, were a peripheral exploration into the hybrid wafer-level bonding topology presented in Chapter 1.

3.6 Conclusion

This chapter used the analysis from Chapter 2 to implement modulator driver circuits for both the carrier-injection and carrier-depletion device topologies. These designs

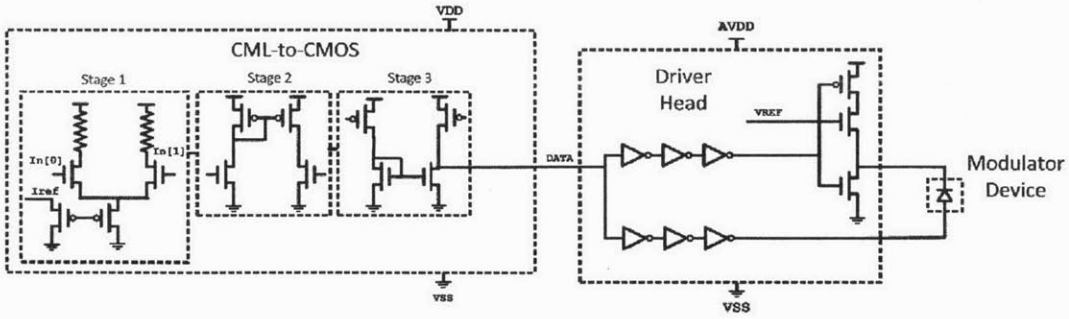


Figure 3-37: Modulator driver, fed directly from a CML-to-CMOS converter.

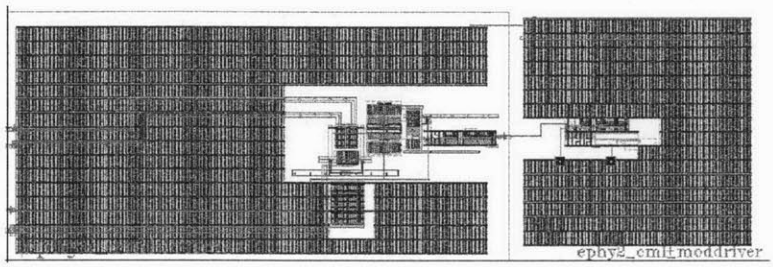


Figure 3-38: Layout of 150m x 50m CML-to-CMOS (on the left) and modulator driver (on the right).

were taped-out and tested in the lab, providing breakthrough results, and showing the importance of close interaction between device- and circuit-designers. This work provides the main contribution of this thesis, through the demonstration of the fastest, most energy-efficient monolithically-integrated modulator and driver circuit in a sub-100nm to date.

To implement these circuits in a sub-100nm thin-SOI CMOS process, some post-processing steps had to be performed on the die after they came back from the foundry. The next chapter will explain these steps in fine detail.

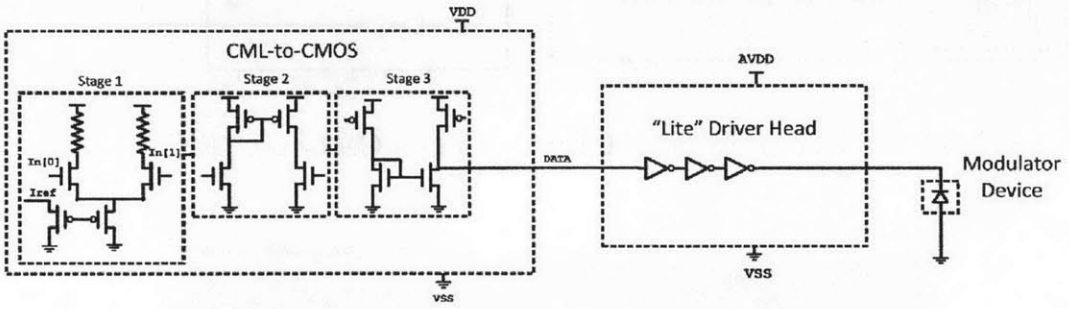


Figure 3-39: CML-to-CMOS converter attached to the lite inverter-based modulator driver.

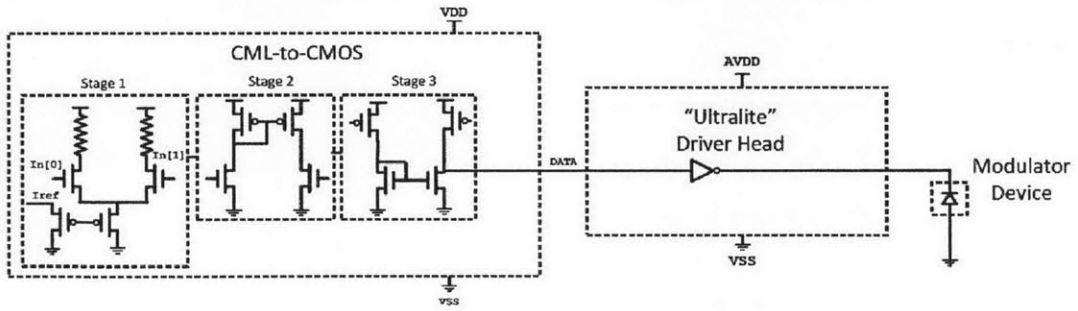


Figure 3-40: CML-to-CMOS converter attached to the ultralite single-inverter driver.

Chapter 4

Post-Processing and Bonding

To enable optics in the IBM 45nm process, a postprocessing step is necessary. After the IBM 45nm chips return from the foundry, each must be post-processed to remove the substrate directly underneath optical structures to avoid high optical losses. Xenon Diflouride (XeF_2) gas is to remove the substrate underneath the optical structures, which selectively etches silicon isotropically. Several techniques of post-processing exist, each with certain trade-offs and applicability toward testing and development. While this etch step does add cost and complication to final assembly, this complication is on par with normal packaging considerations such as wire- or flip-chip bonding. Companies such as Akustica [9] have used a similar XeF_2 localized wafer-level release process for CMOS MEMs microphones; however, with this work, we are performing the release at the die-level. The etch time must be carefully controlled so that the XeF_2 does not over-etch and destroy circuits, as happened in Figure 4-1. After etching, the resulting air gap underneath the optical devices provides the index contrast necessary to achieve optical mode confinement. The table on the following page summarizes the pros and cons of each processing technique.

Procedure	Applicability	Yield	Probing?	Wire Bonding?	Flip-Chip Bonding?	Thermal
Frontside Release (Localized)	Other procedures are superior	Low; multiple procedures (Deep-RIE + XeF_2)	Possible	Possible; weakened substrate reduces yield	Not suitable	Good
Full Release on Handle	Optical testing only (die upside-down)	High	No	No	No	Bad
Full Release and Substrate Transfer	Low-quantity testing; device chips	Low	Yes	Low yield	Not suitable	Reasonable
Flip-Chip Full Release	System chips	High	Not possible	Not possible	Yes	Bad (no substrate)
Flip-Chip Partial Release	System chips	High (presumed)	Not possbile	Not possible	Yes	Good

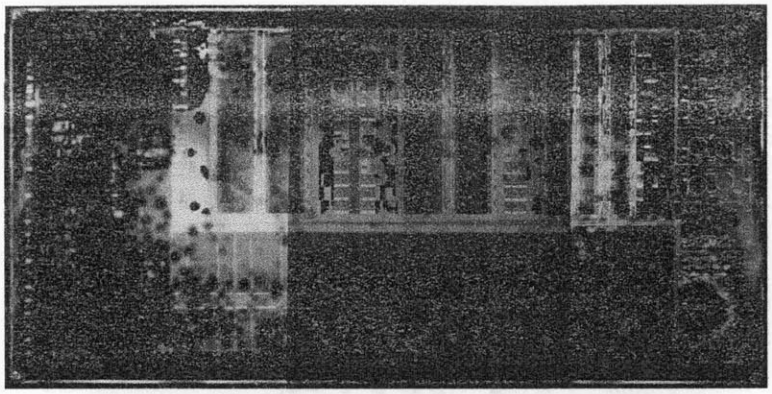


Figure 4-1: An example of XeF_2 over-etch causing catastrophic circuit damage.

4.0.1 Frontside Release

Our research group's first efforts regarding substrate release procedures were focused on frontside release techniques, because early iterations of our technology development platform were taped-out in bulk-CMOS processes. In a bulk-CMOS process, etch holes must be created extending from top-layer metal down to the substrate through the STI to expose the XeF_2 gas to the bulk substrate [12]. Rows of these etch vias run laterally across the chip in close proximity to the waveguides. One localized-removal postprocessing procedure is to reactive-ion-etch these vias away, leaving holes from the top of the chip down to the substrate. Xenon difluoride gas is released into these holes, eating away the substrate underneath the etch row and the waveguide and leaving an air gap. The end result of this process is shown in Figure 4-2.

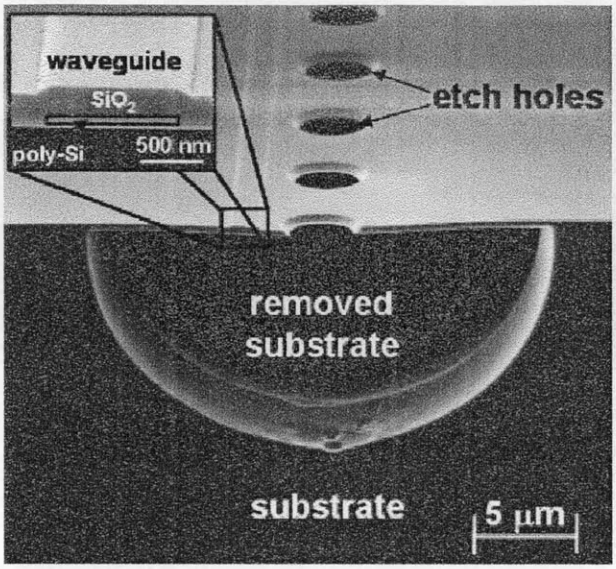


Figure 4-2: Substrate undercut using XeF_2 in a bulk-CMOS process [12].

4.0.2 Full Release on Handle

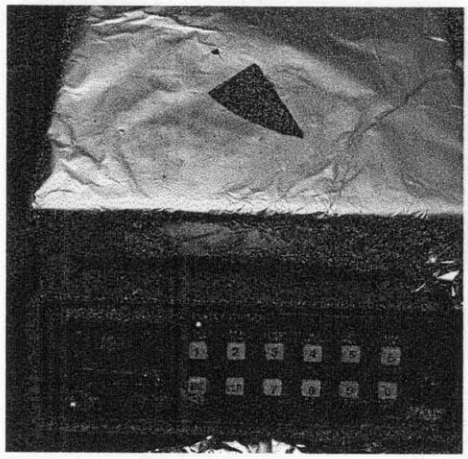
In contrast to a bulk-CMOS process, in a thin-BOX SOI process, the optical devices and waveguides made using body-Si and poly-Si are sitting on top of the thin silicon dioxide. Although this thin layer of SiO_2 does not sufficiently confine the optical mode on its own, the buried oxide does provide a natural etch-stop for the XeF_2, which etches silicon preferentially to oxide at a rate of 1000:1. A die which has been fully released on a handle wafer (or perhaps a glass slide) has its backside exposed, and passive optical structures can be easily tested. This is a simple process with a high success rate; however, it is not appropriate for any type of circuit testing.

Preparing the Die for Xenon Diflouride Release

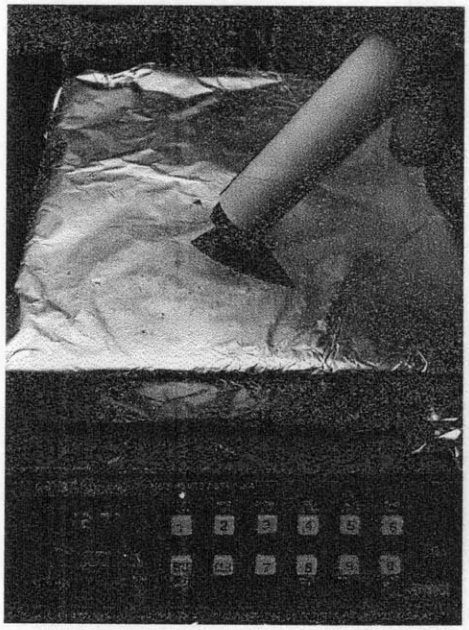
For reference, the melting point of Crystalbond 509 is around 121C.

- Place the handle wafer on a hotplate, and set the hotplate for 130C. A clean glass slide can be used in place of a handle wafer, or a silicon wafer with a thick oxide coating can be used. (In the illustrated picture, a piece of an oxide wafer piece is used, and prior to release, the sides of the wafer are coated with

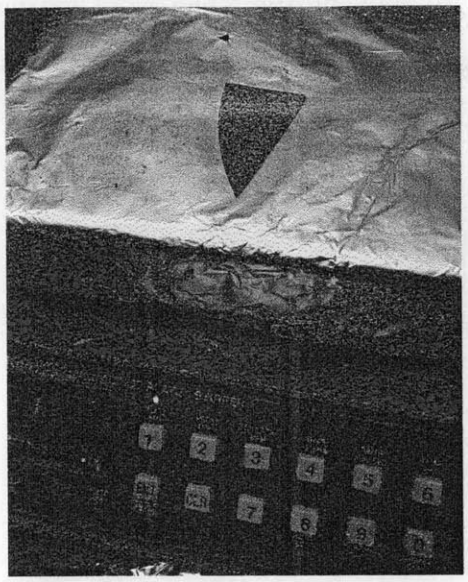
Crystalbond as well to prevent the XeF_2 from eating the handle). A silicon wafer without an oxide coating will not work, because then the handle wafer will also be attacked by the XeF_2.



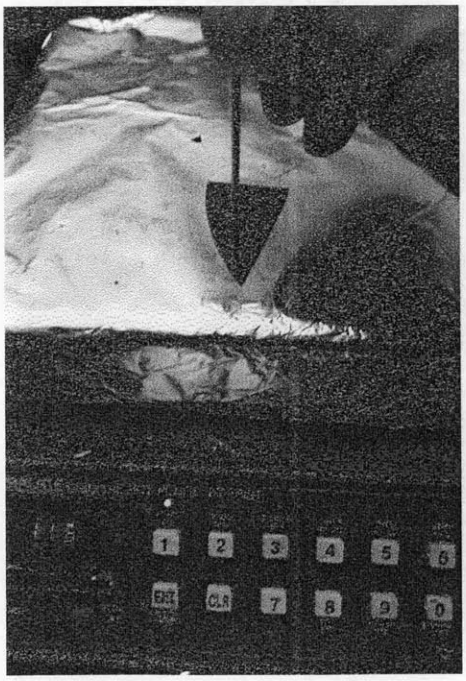
- Apply a dab of Crystalbond 509 to the handle wafer after the handle wafer has had time to fully heat up.



- The die should be placed face-down on the Crystalbond 509.



- Smear the die a few millimeters to ensure that the crystalbond has completely covered the die front surface. It is also desirable to have the crystalbond cover the sides of the die as well, but undesirable to have Crystalbond on the exposed top of the die (top, in this case, meaning backside).



- Before the first etch, estimate the die height under a microscope by focusing

on the top of the die, setting the Z-dimension to 0, and then re-focusing on the handle wafer. Typical die heights are in the 300um range.

- An SE Tech ES-2000XM etcher was used for the release. The recipe is the same for both full release chips and packaged chips; cyles of 60 second etch and 20 second pumping. The number of cycles varies due to the change in total area exposed and the remaining XeF_2 source, so the best way to control the etch precisely is do a short etch for each sample and estimate the time/cycles needed.

- Perform a short etch of perhaps 5 cycles (60 seconds each) and look at the die again under the microscope. The etch speed can then be determined. 10 μm/minute is probably the maximum etch speed that is achievable. At this point you can use a greater number of cycles safely. Try not to over-etch because the XeF_2 will also slowly eat away the oxide with a silicon-to-oxide selectivity of 1000:1.

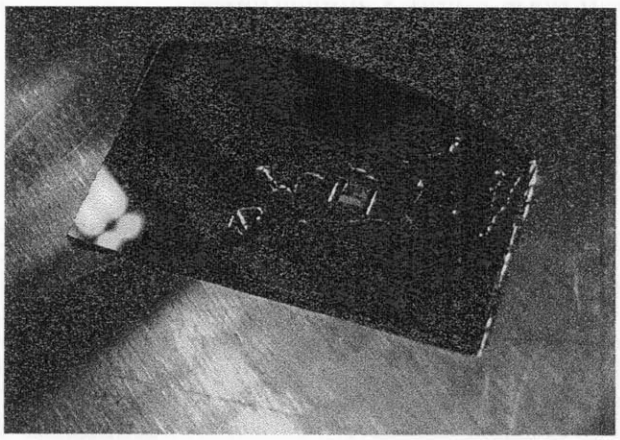


Figure 4-3: The die after release.

4.0.3 Full Substrate Release and Transfer

An extension upon the technique in the previous section is to fully release the substrate of the chip and then transfer it onto a new, optically-friendly substrate. [1] This

[1] This procedure was initially performed by Jason Orcutt.

procedure has a low yield due to the difficulty of bonding to a substrate-transferred chip, but this procedure does work nicely for device chips requiring electrical probing. It is possible to wire-bond these chips, but again, the yield is very low. First, perform the procedure of the previous section, and continue with the detailed procedure listed below.

- A replacement substrate needs to be created. A glass slide can be used to create a glass substrate, by cutting it with a die saw.

- The glass substrate should be cut to slightly smaller than the die size (say 100-200um less in both X and Y dimensions).

- It is important that the glass substrate be slightly smaller than the die. After the sample has been substrate-released, the die film will be sitting in a depression with Crystalbound surrounding it. A glass substrate that is larger than the die size will not fit inside this depression.

- The glass substrate should be cleaned with acetone, methanol and isopropanol before the substrate transfer process. When the glass is cut by the diesaw, the blade will not completely cut through the glass. As a result, on each glass square, one side will have a small ledge sticking out. This ledge should be facing up and can be used to help grip the glass substrate piece.

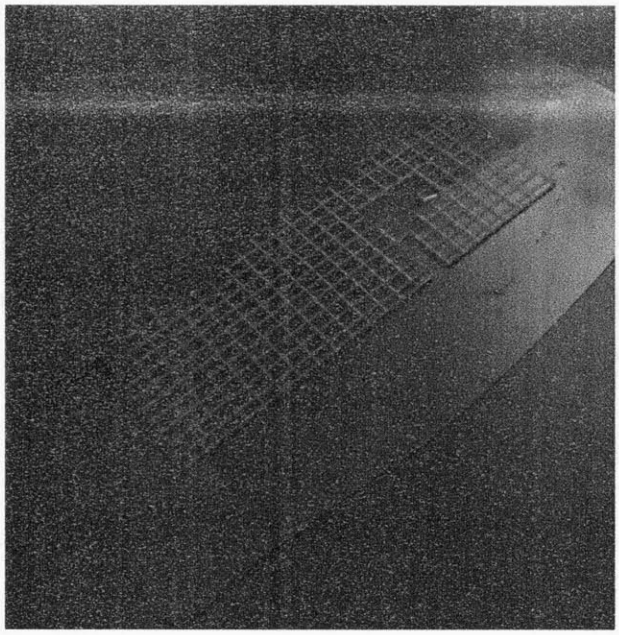
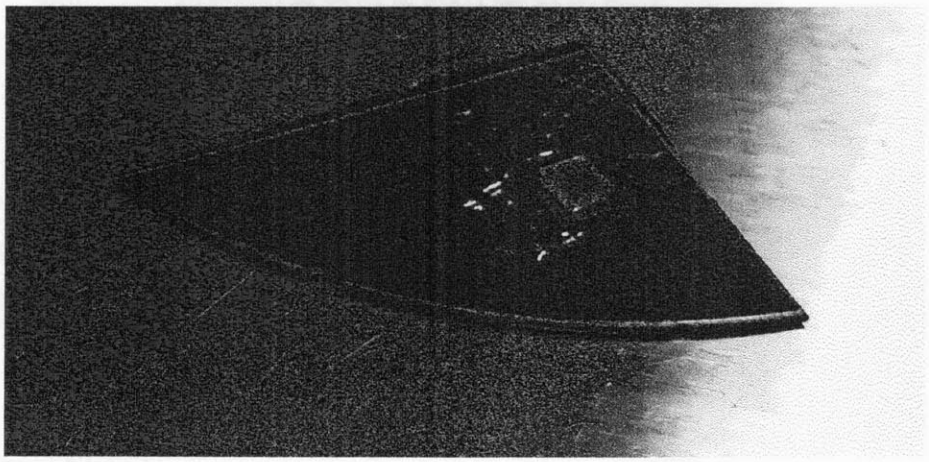


Figure 4-4: A glass slide has been diesawed into replacement substrates.

- A very small amount of Norlad Optical Adhesive 71 (NOA) [2] can be placed on the backside of the die film using the handle end of a wooden q-tip. The amount of NOA should be as small as possible.
- Orient the glass substrate such that the ledge side is facing up. Using tweezers, grab the glass substrate by the ledge, and place it inside the depression on top of the NOA.



[2] Available from `https://www.norlandprod.com/adhesiveindex2.html`

- Under a microscope, ensure that the NOA has full coverage over the glass substrate, and ensure that the glass substrate is oriented properly.
- Cure the NOA under a UV lamp for one hour or longer.
- After the NOA is cured, the die (with new glass substrate) needs to be removed from the handle wafer.
- Place the handle wafer on a hotplate at 130C.
- Once the Crystalbond has melted, carefully use tweezers grabbing the lip of the replacement substrate, and slide the die off.
- The die can be cleaned using ESD-safe Q-tips with acetone, isopropanol, and methanol to clean the die surface.

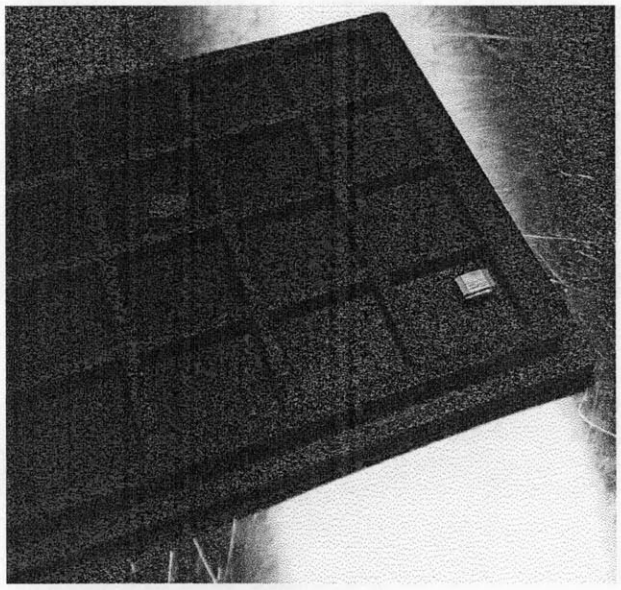


Figure 4-5: Completed substrate-transferred die.

4.0.4 Flip-Chip Full Release

The most reliable method of substrate removal, to date, is to first flip-chip bond the die to an FR4 circuit board, with underfill epoxy between the C4 solder balls. The

sides of the die can be protected using an epoxy (such as NOA). Because the bonding happens before the substrate removal, this process has a high yield.

- The die must be flip-chip bonded to the PCB with underbump epoxy.
- Clean the exposed part of the substrate of the die. Any residual underbump epoxy which is smeared on the top edge will cause the etch to occur unevenly. Use an ESD-safe Q-tip, along with a small amount of acetone, to clean the surface; follow with isopropanol and methanol. View the surface (backside) of the die under the microscope to ensure that it is clean.
- After the die is clean, apply NOA to the sides of the die. The idea is to completely protect the sides of the die from the XeF_2 gas, so that the gas can only etch from the backside of the die, and not from the sides or underneath. There is not a particular reason why NOA is used for this, other than that it has a convenient viscosity, is already used in other substrate-transfer processes (so is likely available), and is easily cured using UV light. The same NOA 71 as is used for the substrate transfer process is fine. The handle end of a Q-tip works well for applying the NOA.
- After the NOA has coated the sides of the chip, the NOA must be cured under an ultraviolet lamp for at least an hour. Leaving it longer than an hour is better.
- This substrate removal process from this point onward is now identical to the process used for full substrate release.
- Before the first etch, estimate the die height under a microscope by focusing on the top of the die, setting the Z-dimension to 0, and then re-focusing on the handle wafer. Typical die heights are in the 300um range.
- An SE Tech ES-2000XM etcher was used for the release. The recipe is the same for both full release chips and packaged chips; cyles of 60 second etch and 20 second pumping. The number of cycles varies due to the change in total

area exposed and the remaining XeF_2 source, so the best way to control the etch precisely is do a short etch for each sample and estimate the time/cycles needed.

- Perform a short etch of perhaps 5 cycles (60 seconds each) and look at the die again under the microscope. The etch speed can then be determined. 10 μm/minute is probably the maximum etch speed that is achievable. At this point you can use a greater number of cycles safely. Try not to over-etch because the XeF_2 will also slowly eat away the oxide with a silicon-to-oxide selectivity of 1000:1.

The end result of the flip-chip substrate release procedure is a nicely packaged circuit board and die, shown in Figure 4-6. This procedure is easily extended for a patterned partial etch process to help with thermal management.

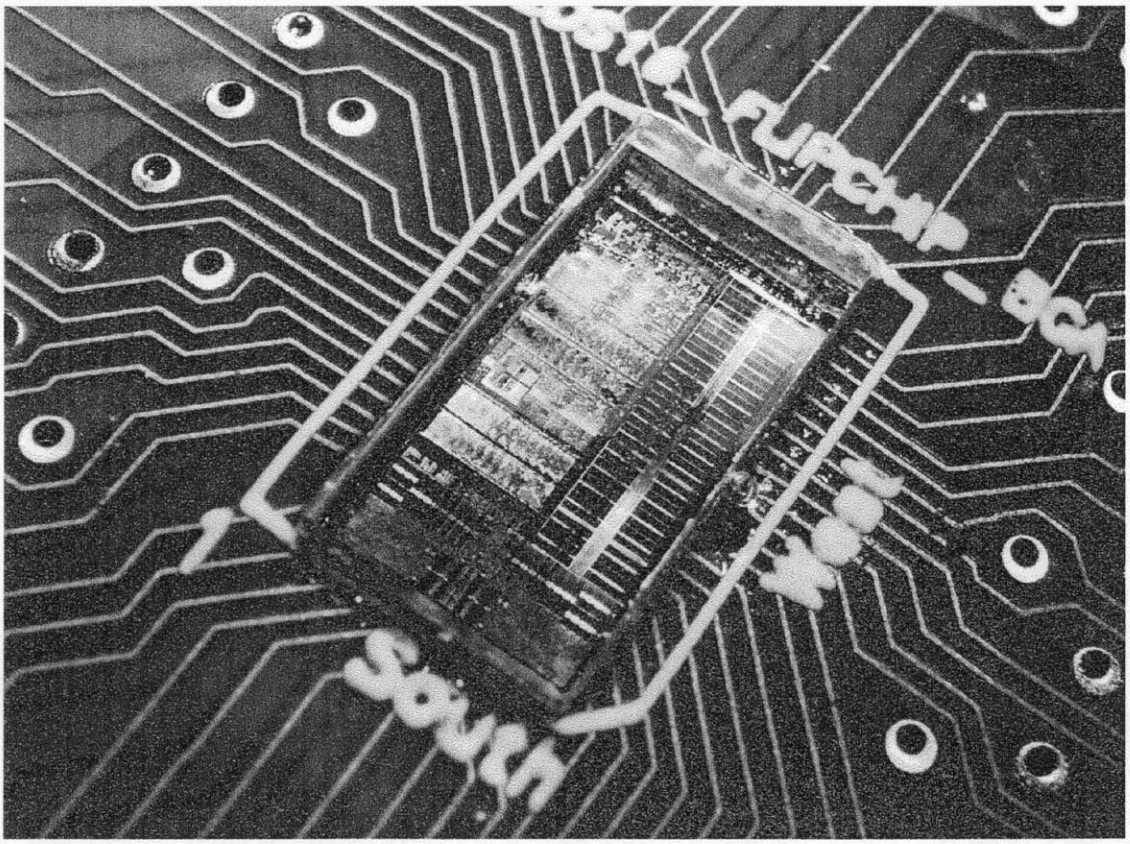


Figure 4-6: A fully released, flip-chip-bonded EOS 18 die

4.0.5 Flip-Chip Partial Release

The flip-chip partial release procedure is similar to the flip-chip full release procedure, with the addition of a coarse patterning step. This will allow more localized substrate removal of regions of the die. The main advantage of this method is that it will improve thermal dissipation for the processor on the die.

This process has not yet been completed, however, as the research project continues it will be an important procedure. The exact execution details of the patterning step have yet to be determined.

4.1 Processing Conclusion

This chapter has outlined the various substrate-release procedures that have been performed on the IBM 45nm thin-SOI family of chips. These different processes have advantages and disadvantages based on their intended application. The important takeaway point is that in a commercial application environment, the flip-chip partial release solution enables photonics, while still providing easy thermal management solutions (such as a traditional heatsink) combined with the ease and reliability of flip-chip bonding.

Chapter 5

System Modeling

5.1 Introduction

As a conclusion to the work in this thesis, the following chapter is dedicated to an analysis of the role of the modulator circuits in the overall WDM link system, as published in [10].

Chapter 1 examined in detail how future high performance computing continues to advance towards the many-core regime. It is clear that new processors will require ever-increasing bandwidth from the on-chip interconnect network and off-chip interfaces. These interconnect fabrics already occupy large portions of the chip area and consume a significant fraction of the total power in current generation processors [31]. Furthermore, projections show poor scaling of on-chip wires and I/O bandwidth density with technology [27] [4], highlighting the need for a disruptive interconnect technology to meet the throughput demands and power-efficiency requirements of many-core systems. Distance-insensitive energy-per-bit stand out as key enablers for a faster, denser, and more energy-efficient interconnect fabric.

To realize these potentials, it is essential for device, circuit and system designers to understand the relationships and trade-offs among components in this new technology, as well as the impact of the potential integration scenarios. The modulator device and driver circuits in particular play a key role in the power efficiency of these systems, especially as an enabler for dense wavelength-division-multiplexing (WDM). In this

chapter, we illustrate these design trade-offs on an example of a complete integrated WDM photonic link, by creating the component models that connect device, process and circuits parameters to link performance and power consumption. Identifying the modulator driver and circuit as one of the most-significant costs, we show that careful design-space exploration can result in an optimal set of link, circuit, and device parameters. his exploration also highlights the significance of monolithic-integration to minimize the receiver parasitics, reducing the laser power and enabling low-energy, high throughput-density interconnect fabrics.

5.2 Photonic Link Components

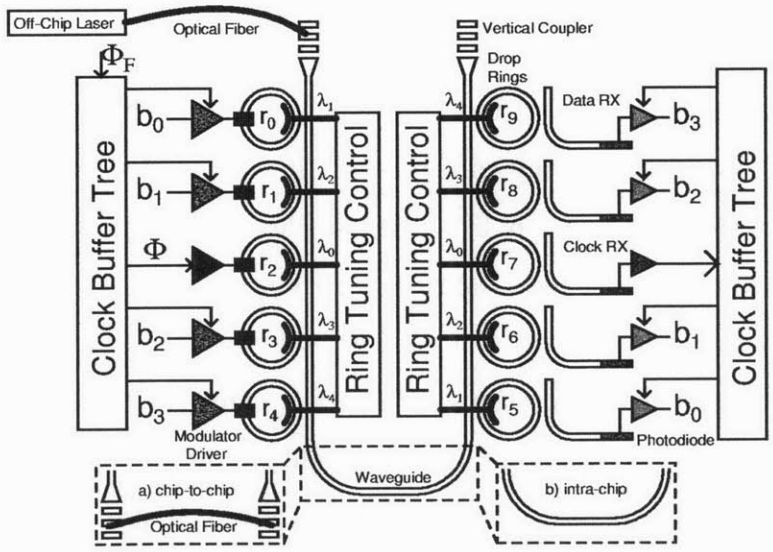
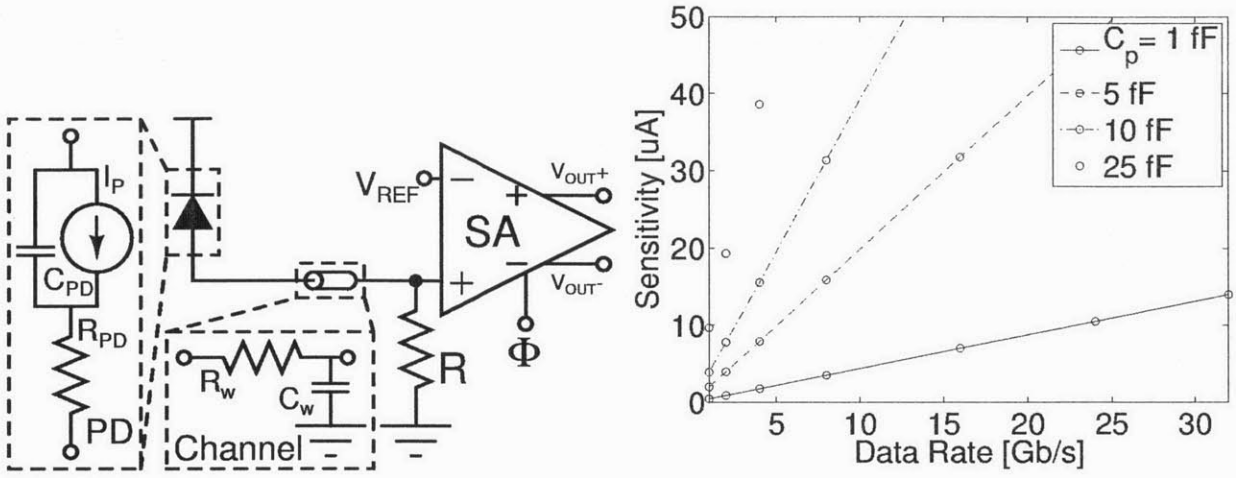


Figure 5-1: An integrated WDM photonic link. A continuous-wave (CW) multi-λ laser is coupled onto the chip through a vertical-coupling grating structure. Once on chip, frequency selective ring-resonant modulators encode digital bitstreams onto their resonant wavelengths. Each wavelength propagates along the waveguide (and possibly off-chip) until it is routed through a matching drop ring to an integrated photodiode. An optical receiver forms a bit decision based upon the photodiode photocurrent. Clock signals are routed both optically along the waveguide and electrically through local H-trees. Ring tuning circuits are used to tune the resonance of the modulator and drop rings.

To begin this analysis, a good starting point is a discussion concerning the operation of modulators and receivers, the primary data-path elements that form a WDM photonic link (Figure 5-1). As these two components depend highly on the characteristics of the devices that they use, a device-technology-driven analysis is appropriate. Two integration scenarios are considered - firstly, a monolithic integration of photonic components into the CMOS front-end (polysilicon photonics or thin-BOX SOI photonics, such as IBM 45nm SOI), and a hybrid integration scenario with an optimized SOI photonic die attached to the CMOS chip using through-silicon vias (TSVs). The monolithic integration will typically have smaller parasitic capacitances between circuits and photonic components, while potentially having higher optical losses due to fabrication process constraints.

5.2.1 Optical Data Receiver Modeling

The optical receiver converts optically-modulated data back into the electrical domain by sensing a photocurrent produced by the photodiode. In contrast to traditional optical receivers which utilize various power-hungry trans-impedance amplifiers (TIA) to combat the large photodiode parasitic capacitance, monolithic integration offers the opportunity for much-simpler, energy-efficient receiver circuits due to low photodiode parasitic capacitances. The CICC paper illustrates the relationship between the sensitivity and power consumption across ranges of data rates and parasitic capacitances, for various receiver topologies, including resistive receivers, transimpedance amplifiers (TIA), and current-integrating receivers. The receiver and photodiode performance is directly related to the modulator device and circuit; this receiver-side modeling is necessary to make modulator trade-offs at the system level. A more detailed analysis is available in the 2011 CICC paper [10], but the three topologies are briefly introduced here.



(a) Resistive optical receiver with photodiode and channel models shown.

(b) Photocurrent Sensitivity.

Figure 5-2: Resistive receiver sensitivity.

An equivalent model of the photodiode is shown in Figure 5-2a, consisting of a capacitance in parallel with a photocurrent-generating source and series resistance. The photodiode is connected to the front-end through either a Through-Silicon Via (TSV, $C_p \approx 25\,\text{fF}$) or low-level metal routing ($C_p \approx 5\,\text{fF}$). All parasitic series resistances are assumed negligible.

The photodiode in Figure 5-2a is connected to a resistive receiver with a sense amplifier. The regenerative sense-amplifier, known for energy-efficient and scalable operation, creates a full-voltage-swing interface with the digital back-end. Photocurrent is driven across the resistance, which is the front-end's gain. Figure 5-2b shows that the receiver is able to sense photocurrents of approximately $10\,\mu\text{A}$ for low C_p and data rate. The sensitivity worsens linearly with data rate as the resistance is traded for bandwith. The energy-efficiency remains constant due to the dominance of the sense amplifier's digital switching power.

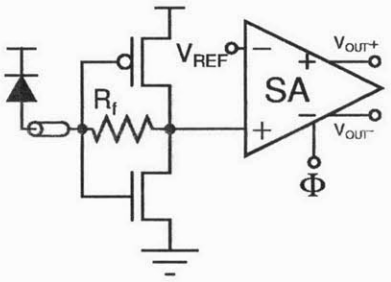


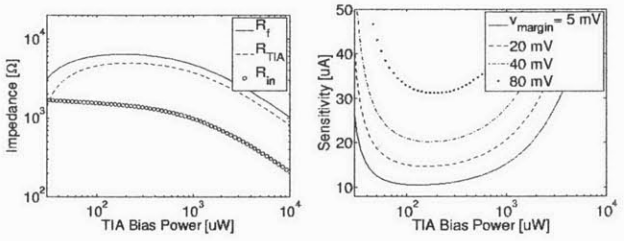
Figure 5-3: Transimpedance Amplifier.

A trans-impedance-amplifier (TIA) topology (Figure 5-3) breaks the gain-bandwidth limitation of the resistive receiver. Equation 5.2 shows the receiver's gain and use of feedback to decrease input impedance [14].

$$BW = \frac{1}{2\pi R_{in}\left(C_{PD} + k_R R_f\right)} \tag{5.1}$$

$$R_{TIA} = \frac{g_m - g_f}{g_f\left(g_m + g_{ds}\right)}, R_{in} = \frac{g_{ds} + g_f}{g_f\left(g_m + g_{ds}\right)} \tag{5.2}$$

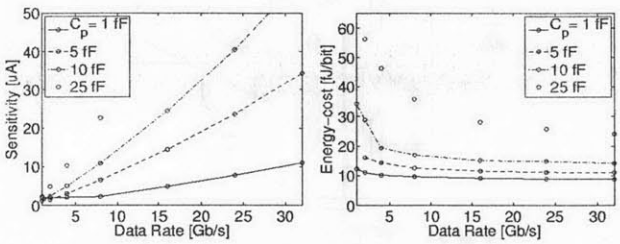
Figure 5-4a shows sensitivity-optimized designs for different TIA bias powers. In this relatively small C_p environment, large designs are penalized for their increased gate capacitance, requiring a reduction in R_{in} and therefore R_{TIA} and sensitivity. Figure 5-4b shows sensitivity optimum for various values of v_{margin}.



(a) v_{margin}=20 mV. (b) Effect of v_{margin}.

Figure 5-4: TIA design example at C_p=25 fF, DR=5 Gb/s.

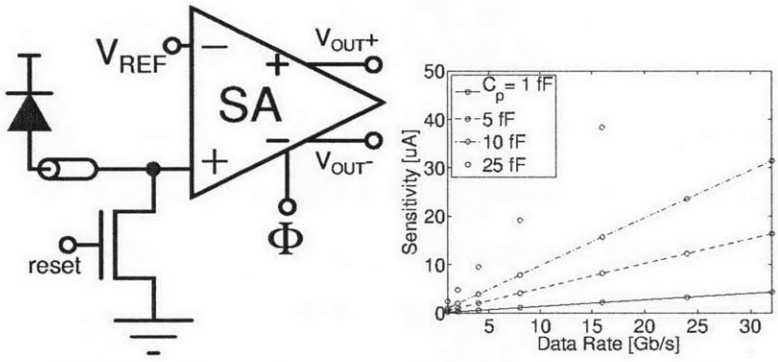
Figure 5-5 summarizes TIA performance for various values of C_p. Though the TIA achieves sensitivity superior to the resistive receiver, the power consumption is considerably worse.



(a) Photocurrent sensitivity. (b) Energy efficiency.

Figure 5-5: TIA performance.

The third topology considered is an integrating receiver (Figure 5-6a), where the photocurrent is converted to a voltage by integrating it onto a capacitor $C_{INT} = C_{PD} + C_w + C_{SA,in}$. The photocurrent is integrated over a fraction (k_{INT}=0.7) of a bit time yielding a front-end gain given by Equation 5.3.



(a) Integrating receiver. (b) Photocurrent sensitivity.

Figure 5-6: Integrating receiver design and performance.

$$R_{INT} = \frac{k_{INT} \cdot T_{bit}}{C_{INT}} \tag{5.3}$$

Figure 5-6 shows that the integrating receiver is the best performing of the three receivers considered. The energy-efficiency of the receiver is dominated by the SA as in the resistive receiver. It should be noted that this simple model has several hidden challenges remaining. The voltage on C_{INT} must be reset or at least charge-shared [7],

which is partially accounted for through k_{INT}. A small C_{INT} will also suffer from SA kickback, while increasing C_{INT} degrades sensitivity.

These receiver topologies are important to consider; from the perspective of modulator design, tradeoffs are impossible to make at either the device or circuit level without considering the other components in the system.

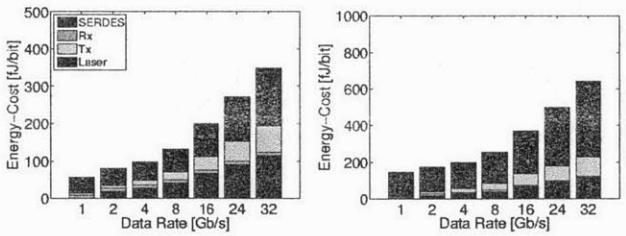
5.2.2 Single Channel Link Tradeoffs

A power optimization across modulator insertion loss, extinction ratio, and receiver topologies is performed for different link data-rates to illustrate the interactions between the modulator and receiver and the impact on wall-plug laser power. Figure 5-7 shows the energy-per-bit breakdowns for four integration scenarios.

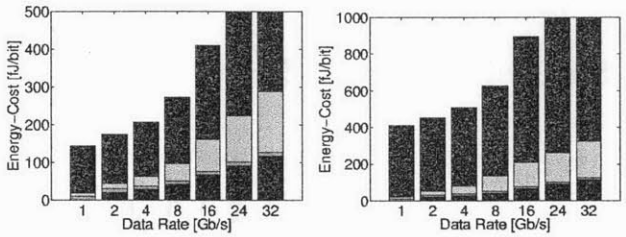
Table 5.1: Link Evaluation Parameters

Parameter	Value
Process Node	32 nm Bulk CMOS
V_{DD}	1.0 V
Device to Circuit Parasitic Cap C_P	5-25 fF
Wavelength Band λ_0	1300 nm
Photodiode Responsivity	1.1 A/W
Wall-plug Laser Efficiency P_{laser}/P_{elec}	0.3
Channel Loss	10-15 dB
Insertion Loss IL_{dB} (Optimized)	0.05-5.0 dB
Extinction Ratio ER_{dB} (Optimized)	0.01-10 dB
Bit Error Rate (BER)	10^{-15}
Core Frequency	1 GHz
SERDES Topology	Mux/Demux Tree

In all plots, the laser power is the dominant energy consumer, increasing quickly with data rate as aggressive modulation rates force a relaxation of modulator insertion loss and extinction ratio. We can see that the laser power is highly sensitive to C_P, as the laser power for C_P=25 fF is roughly 5X that of C_P=5 fF. Though the modulator tries to offset the laser cost by increasing its extinction ratio and decreasing insertion loss, it inevitably reaches a limit on its capabilities. The higher loss simply amplifies the laser power component, resulting in a 3X laser power difference between the 15 dB and 10 dB loss cases. Matching previous analysis, the optimization chose the



(a) Loss=10 dB, C_P=5 fF (b) Loss=15 dB, C_P=5 fF



(c) Loss=10 dB, C_P=25 fF (d) Loss=15 dB, C_P=25 fF

Figure 5-7: Data rate tradeoffs for a single photonic link for 4 integration scenarios. C_P=5 fF represents monolithic integration, while C_P=25 fF is expected for a TSV connection to an optical die. Channel losses of 10 dB and 15 dB correspond to on-chip and chip-to-chip links, respectively.

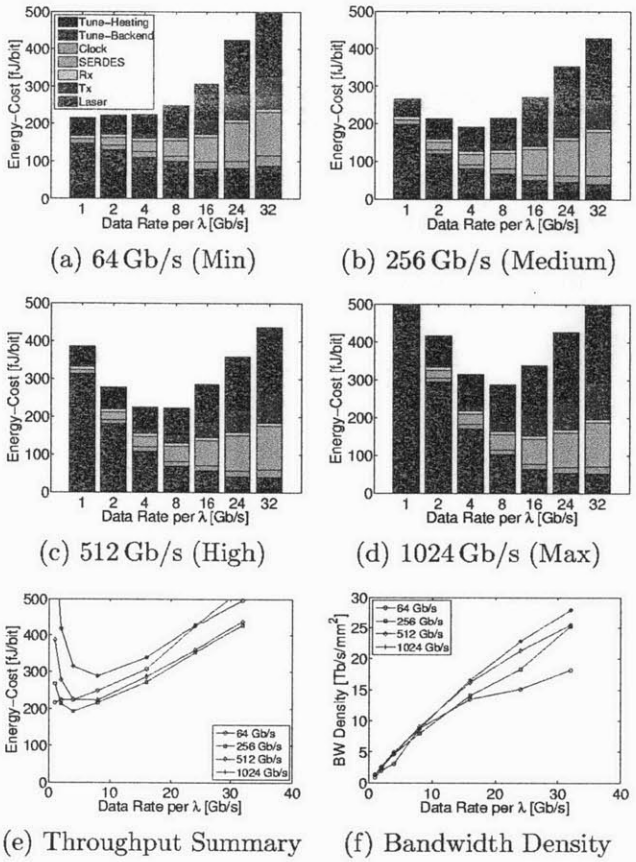
integrating receiver as the optimal receiver in all scenarios. Though our results present a grim outlook for the C_P=25 fF (optical die with TSV) scenario, we note that lower losses may be achievable with a dedicated optical die, allowing TSV integration to remain competitive.

5.3 Towards a Full WDM Link

Expanding upon the analysis for a single-channel data link, next the additional back-end components required in a high-speed multi-channel WDM link are explored.

5.4 WDM Photonic Link Evaluation

In this section, we perform a full link-level optimization and evaluation of a WDM link to quantify energy consumption tradeoffs. In our evaluation, we explore links with 4 different aggregate throughput design points, 64 Gb/s, 256 Gb/s, 512 Gb/s, 1024 Gb/s, corresponding to minimum, medium, high, and maximum bandwidth scenarios.



(a) 64 Gb/s (Min) (b) 256 Gb/s (Medium)

(c) 512 Gb/s (High) (d) 1024 Gb/s (Max)

(e) Throughput Summary (f) Bandwidth Density

Figure 5-8: Optimized power vs. data-rate for different aggregate link throughputs for Loss=10 dB, C_P=5 fF. For tuning, we assume a bit-reshuffler backend and electrically-assisted tuning with local variation σ_{rL}=40 GHz and systematic variation σ_{rS}=200 GHz. Note that the number of WDM channels changes with data rate (Channels = Throughput / Data-Rate).

Figure 5-8 shows that tuning power dominates at lower data-rates (since there are more channels given fixed throughput) and decreases with data-rate. Modulator, laser, SERDES, and receiver energies increase with data-rate and dominate at high rate-rates. At all throughput scenarios, an optimal energy balance is achieved at around 4-8 Gb/s. An overall energy-optimal point occurs at less than 200 fJ/bit for a link with 256 Gb/s of aggregate throughput and 4 Gb/s data-rate.

At the energy optimal point, we see that the energy consumption is roughly an even 3-way split between tuning, laser, and mod/rx/SERDES. As tuning power is now mostly dominated by the backend electrical components, this energy will scale favorably with technology and can be optimized using custom design. A full electrical tuning backend is also unnecessary on both modulate- and receive-side – barrel-shifts

and bit-reordering only need to be performed once – meaning backend power can be cut by another 50%. Refinement of photodetector responsivity and parasitic capacitances as well as lower-loss optical devices with improved electrical laser efficiencies can bring about further reductions in wall-plug laser power. It can be expected that energy/bit will drop to sub-100 fJ with device development, process scaling and overall link component refinement.

5.5 Conclusion

Integrated photonic interconnects are a promising solution to the throughput demands of future many-core processors. As an emerging technology, circuit, device and architecture designers require insights concerning the impact of device and circuit parameters on link-level figures of merit.

This thesis began by introducing the bottlenecks that many-core systems face today, with the growing concern that traditional copper cannot satisfy the needs of future systems. Chapter 2 described in detail two modulator device topologies; a robust but power-hungry carrier-injection design, and a sensitive but fast and energy-efficient carrier-depletion design. Chapter 2 also provided an analysis which can serve as an important meeting point between modulator circuit and device designers.

Chapter 3 used the analysis from Chapter 2 to implement modulator driver circuits for both the carrier-injection and carrier-depletion device topologies. These designs were taped-out and tested in the lab, providing breakthrough results, and showing the importance of close interaction between device- and circuit-designers. This work provides the main contribution of this thesis, through the demonstration of the fastest, most energy-efficient monolithically-integrated modulator and driver circuit in a sub-100nm to date.

Chapter 4 described the post-processing procedures that were used to enable optics in the IBM 45nm thin-SOI platform. Without this effort, the research could not have been tested and demonstrated in the lab.

The final chapter presented a design-space exploration of a WDM integrated pho-

tonic link, facilitated through a set of circuit and device models that captured the optical-electrical tradeoffs of each link component. The modulator model showed the relationship between the modulation energy and the laser power, set through the extinction ratio and insertion loss specifications. Similarly, the optical receiver models demonstrated the degradation of sensitivity with data rate, which translated directly into increased laser power requirement. Using our co-optomized models across all link components, we found that relatively low (sub 10 Gb/s) data-rates per link yielded optimal energy-efficiency accross a range of system throughputs. We showed that the photonic link is highly sensitive to parasitic capacitances present at the receiver input and that optical integration using TSVs to connect to an optical die could result in significant overhead in the laser power. This study illustrated that monolithic integration of photonic components can offer interconnect solutions with high throughput-density and energy-efficiency.

Appendix A

Tables

Component	Preliminary Design	Power Loss	Optimized Design	Power Loss
Coupler loss	1 dB/coupler	3 dB	1 dB/coupler	3dB
Splitter loss	0.2 dB/split	1 dB	0.2 dB/split	1dB
Non-linearity	1 dB	1 dB	1 dB	1 dB
Through loss	0.01 dB/ring	3.17 dB	0.01 dB/ring	3.17 dB
Modulator insertion loss	1dB	1dB	0.5 dB	0.5 dB
Crossing loss	0.2 dB/crossing	**12.8 dB**	0.05 dB/crossing	**3.2 dB**
On-chip waveguide loss	5 dB/cm	**20 dB**	1 dB/cm	**4 dB**
Off-chip waveguide loss	0.5×10^{-5} dB/cm	0 dB	0.5×10^{-5} dB/cm	0 dB
Drop loss	2.5 dB/drop	**5 dB**	1.5 dB/drop	**3 dB**
Photodetector loss	0.1 dB	0.1 dB	0.1 dB	0.1 dB
Receiver sensitivity	-20 dBm	-20 dBm	-20 dBm	-20 dBm
Power per wavelength		26.07 dBm		-1.03 dBm
		(0.40 W)		(0.78 mW)
Power required at source		**3.3 kW**		**6.38 W**

Table A.1: Optical Power Budget

Component	Latency
Serializer/Deserializer (50 ps each)	50 ps
Modulator driver latency	108 ps
Through latency (2.5 ps/adjacent channel)	7.5 ps
Drop latency (20 ps/drop)	60 ps
Waveguide latency (106.7 ps/cm)	427 ps
SM fiber latency (48.3 ps/cm)	483 ps
Photodetector + TIA latency	200 ps
Total latency	**1.385 ns**

Table A.2: Optical Data Transmission Latency

Symbol	Description and Nominal Units
L_j	Length of single diode junction for charge-injection ring in direction of optical propagation [m]
H	Waveguide height for charge-injection ring, determined by thickness of polysilicon layer [m]
W	Waveguide width for charge-injection ring, transverse to the direction of optical propogation [m]
r	Ring radius for charge-injection ring in curved section of racetrack [m]
L_{tot}	Total circumference of charge-injection ring resonator [m]
V_j	Single waveguide diode junction volume (charge-injection)[m^3]
V_{tot}	Total charge-injection ring modulator volume including waveguide and I-region volumes [m^3]
x	Position variable along charge-injection diode length, transverse to direction of optical propogation
L_d	Charge-injection diode length (similar to waveguide width W)
W_d	Charge-injection diode width (equivalent to junction length L_j)
c	Speed of light in free space; 3×10^8 m/s
c_s	Speed of light in waveguide prior to index shift [m/s]
q	Electron charge; 1.602×10^{-19} C
ϵ_0	Permittivity of free space; 8.854×10^{-12} F/m
$m_{ce}*$	Conductivity effective mass of electrons [kg]
$m_{ch}*$	Conductivity effective mass of holes [kg]
j	$\sqrt{-1}$ [unitless]
λ_s	Wavelength of ring resonance [m]
λ_n	Wavelength of ring resonance after index shift [m]
k_0	Wavenumber of light at target optical channel [m^{-1}]
k_n	Wavenumber of light after index shift [m^{-1}]

ω_0	Angular frequency of target optical channel [rad/s]
f_0	Frequency of target optical channel [Hz]
f_s	Frequency of passive ring resonance [Hz]
f_n	Frequency of shifted ring [Hz]
FSR	Ring resonator free spectral range [Hz]
k_{FSR}	Wavenumber of free spectral range [m^{-1}]
ω_{FSR}	Angular frequency of free spectral range [rad/s]
$\mathbb{Q}$	Quality factor of ring resonator [unitless]
m	Non-zero integer number of wavelengths in a round trip around the ring at resonance [unitless]
Δm	Pertubation to m due to index shift [unitless]
Δk	Pertubation to k_0 due to index shift [m^{-1}]
T_0	Transmissivity of ring filter at λ_0 prior to index shift [unitless]
T_{0dB}	Transmissivity of ring filter at λ_0 prior to index shift [dB]
T_1	Transmissivity of ring filter at λ_0 after index shift [unitless]
T_{1dB}	Transmissivity of ring filter at λ_0 after index shift [dB]
T_s	Transmissivity of ring filter at passive resonance [unitless]
T_{sdB}	Transmissivity of ring filter at passive resonance [dB]
T_N	Transmissivity of ring filter after index shift [unitless]
T_{NdB}	Transmissivity of ring filter after index shift [dB]
ER	Extinction ratio at λ_0 [unitless]
ER_{dB}	Extinction ratio at λ_0 [dB]
IL	Insertion loss at λ_0 [unitless]
IL_{dB}	Insertion loss at λ_0 [dB]
Δf	Frequency shift of ring resonance due to index shift [Hz]
$\Delta\phi_{HWHM}$	Half-width-half-maximum phase shift to light at λ_0 due to index shift
$\Delta\phi_{FWHM}$	Full-width-half-maximum phase shift to light at λ_0 due to index shift
Δf_{HWHM}	Half-width-half-maximum resonant frequency shift due to index shift [Hz]
Δf_{FWHM}	Full-width-half-maximum resonant frequency shift due to index shift [Hz]
$\overline{E_0}$	$\overline{E}$-field vector representing light in the waveguide prior to index shift [V/m]
$\overline{E_1}$	$\overline{E}$-field vector representing light in the waveguide after index shift [V/m]
$\underline{E_0}$	Complex phasor representing $\overline{E}$-field at λ_0 [V]
$\underline{E_1}$	Complex phasor representing $\overline{E}$-field after index shift [V]
n	Refractive index in unperturbed silicon [unitless]
n_g	Group index of ring waveguide [unitless]

Δn	Change in refractive index due to index shift (or temperature change)
Δn_e	Change in refractive index due to injected electrons
Δn_h	Change in refractive index due to injected holes
Δn_{eff}	Effective index change due to index shift
Γ	Waveguide mode overlap coefficient for carrier-injection modulator [unitless]
n_{fe}	Polysilicon free electron index change coefficient [cm^{-3}]
n_{fh}	Polysilicon free electron index change coefficient [cm^{-3}]
ΔN	Injected electron-hole pair concentration change in the diode I-region [cm^{-3}]
ΔN_e	Injected electron concentration change in the diode I-region [cm^{-3}]
ΔN_h	Injected hole concentration in the diode I-region [cm^{-3}]
ΔN_0	Required injected free carrier concentration change to achieve a 3 dB extinction ratio [cm^{-3}]
ΔN_{tot}	Required number of injected free carriers to achieve a 3 dB extinction ratio
ΔQ_0	Required amount of injected charge to achieve a 3 dB extinction ratio
γ	Injection ratio
$\Delta\phi'$	Phase change to light at λ_0 from a single P-I-N diode junction due to charge injection
$\Delta\phi$	Phase change to light at λ_0 from both diode junctions due to charge injection
t	Time [s]
$Q(t)$	Total injected charge in I-region [C]
Q_S	Steady-state charge in I-region [C]
τ_c	Polysilicon carrier lifetime [s]
$i(t)$	Diode current [A]
V_D	Diode bias voltage [V]
V_{DD}	Supply voltage [V]
V_{th}	Diode threshold [V]
R_{dr}	Effective driver resistance
R_A	Effective pre-emphasis driver resistance
R_B	Effective forward-bias driver resistance
R_C	Effective reverse-bias driver resistance
R_S	P-I-N diode series contact resistance

Table A.3: Symbol description

A.1 Depletion-Mode Modulator Verilog-A Model

```
`include "disciplines.vams"

/* Example Modulator Device Verilog AMS Model  */

module modulator_ridge(p, n, gnd, in, through, drop);

//-------------------------
// DEFINE PORTS
//-------------------------
inout p, n, gnd;
input in;
output through, drop;
electrical p, n, gnd, in, through, drop;
branch (p,n) res, cap;

//-------------------------
// OPTICAL PARAMETERS
//-------------------------

// Physical Constants
// Pi [unitless]
parameter real pi  = 3.1415926536;
// Boltzmann constant [J/K]
parameter real k   = 1.38e-23;
// Temperature [K]
parameter real T   = 300;
// Charge of electron [C]
```

```
parameter real q   = 1.6e-19;
// Permittivity of free space [F/m]
parameter real eo  = 8.85e-12;
// Relative permittivity of silicon [unitless]
parameter real es  = 12;
// Speed of light in vacuum [m/s]
parameter real c = 3e8;
// Refractive index change of Si for BW shift [1/m^3]
parameter real nf  = 3e-27;
// Silicon free carrier index shift polynomial coefficients
//    (From Soref 1987, computed by Milos Popovic)
parameter real nP1e     = 1.066;
parameter real nP2e     = -22.51;
parameter real nP1h     = 0.8035;
parameter real nP2h     = -17.37;
parameter real aP1e     = 1.057;
parameter real aP2e     = -18.21;
parameter real aP1h     = 1.062;
parameter real aP2h     = -18.56;

// Physical Design Parameters
// Unbiased ring resonant wavelength [m]
parameter real lambda  = 1273.5e-9;
// Silicon group index [unitless]
parameter real ng = 4;
// P-type doping [m^(-3)]
parameter real Na = 8e17 * 1e6;
// N-type doping [m^(-3)]
parameter real Nd  = 8e17 * 1e6;
// Intrinsic doping [m^(-3)]
```

```
parameter real ni  = 1e10 * 1e6;
// See diagram, un-ridged N-doped [m]
parameter real L1   = 600e-9;
// See diagram, un-ridged P-doped [m]
parameter real L2   = 1e-6;
// Ridge width [m]
parameter real wRidge  = 450e-9;
// Un-ridged poly thickness [m]
     parameter real h1 = 60e-9;
// Ridge only poly thickness [m]
     parameter real h2  = 165e-9;
// Bend radius, center to edge of ridge [m]
parameter real r_outer  = 3.5e-6;
// Confinement factor for N edge of depletion region [unitless]
parameter real GammaN = 1;
// Confinement factor for P edge of depletion region [unitless]
parameter real GammaP   = 1;
// Modulator transmissivity at resonance [unitless]
parameter real Tn  = 0.01;
// Full-width-half-max bandwidth [Hz]
parameter real fwhm  = 9.7e9;
// Electron mobility [m^2/V/s] *from Pierret p.80
parameter real ue       = 300 * 1e-4;
// Hole mobility [m^2/V/s] *from Pierret p.80
        parameter real uh       = 150 * 1e-4;
// Minority carrier lifetime - electrons (guess) [s]
        parameter real tau_e    = 150e-12;
// Minority carrier lifetime - holes (guess) [s]
        parameter real tau_h    = 150e-12;
```

```
// Derived Parameter Declarations
// Total ridge thickness [m]
real htot;
// Effective ring length [m]
real Leff;
// Modulator center frequency [Hz]
real fo_ring;
// Total ring volume [m^3]
real Vtot;
// Electron diffusion coefficient [m^2/s]
real De;
// Hole diffusion coefficient [m^2/s]
        real Dh;
// Electron decay length [m]
        real Le;
// Hole decay length [m]
        real Lh;
  // Laser frequency [Hz]
real f_laser;
// Change in refractive index [unitless]
real delta_neff;
real delta_alpha;
real delta_invQ;
// Change in FWHM [Hz]
real delta_FWHM;
// Change in resonant frequency [Hz]
real delta_fo;
// Shifted Full Width Half Max [Hz]
real fwhm_bias;
// Modulator shifted resonant frequency [Hz]
```

```
real fo_ring_bias;
real dne;
real dae;
real dnh;
real dah;

//-------------------------
// ELECTRICAL PARAMETERS
//-------------------------

// Independent Parameters
// Forward transit time [s]
        parameter real tf = 0;
// Series resistance [Ohms]
parameter real Rseries = 30;

// Derived Parameter Declarations
// Reverse-bias saturation current [A]
real Is;
// Thermal voltage [V]
real Vt;
// Built-in voltage [V]
real Vbi;
// Unbiased depletion width [m]
real xdo;
// Unbiased junction capacitance [F]
real Cjo;
// Unbiased depletion width on n side [m]
real xno;
// Unbiased depletion width on p side [m]
```

```
real xpo;
// Depletion width on n side [m]
real xn;
// Depletion width on p side [m]
real xp;
// Depletion width change on n side [m]
real dxn;
// Depletion width change on p side [m]
real dxp;
// Depletion region width [m]
real xd;
// Charge in diode [C]
real q_diode;
// Diode charge density [C/m^3]
real q_density;
// Volume of depletion region [m^3]
real Vdr;

//-------------------------
// DETERMINE TRANSMISSIVITY
//-------------------------

analog begin

// Calculate Derived Optical Parameters
// Total ridge thickness [m]
htot = h1 + h2;
// Effective ring length [m]
Leff = 2*pi*(r_outer - wRidge/2);
// Modulator center frequency [Hz]
```

```
fo_ring = c / lambda;
//fo_ring;
f_laser = 235.576e12;
// Total junction volume [m^3]
Vtot = h1*(pi*(r_outer + L2)*(r_outer + L2) - \
pi*(r_outer - wRidge - L1)*(r_outer - wRidge - L1)) \
+ h2*(pi*r_outer*r_outer - pi* \
(r_outer - wRidge)*(r_outer - wRidge));

dne = -pow(10,nP2e) * pow(Nd/1e6,nP1e);
dnh = -pow(10,nP2h) * pow(Na/1e6,nP1h);
dae = 100*pow(10,aP2e) * pow(Nd/1e6,aP1e);
dah = 100*pow(10,aP2h) * pow(Na/1e6,aP1h);

// Calculate Derived Electrical Parameters
// Electron diffusion coefficient [m^2/s]
        De        = ue * Vt;
// Hole diffusion coefficient [m^2/s]
        Dh        = uh * Vt;
// Electron decay length [m]
         Le        = sqrt(De*tau_e);
// Hole decay length [m]
        Lh        = sqrt(Dh*tau_h);

// Reverse-bias saturation current [A]
Is  = htot*Leff*q*ni*ni*(De/Na/Le + Dh/Nd/Lh);
// Thermal voltage [V]
Vt  = k*T/q;
// Built-in voltage [V]
Vbi = Vt*ln(Na*Nd/ni/ni);
```

```
// Unbiased depletion width [m]
xdo = sqrt(2*eo*es/q*(Na+Nd)/Na/Nd*Vbi);
// Unbiased junction capacitance [F]
Cjo = eo*es*Leff*htot/xdo;
xno = xdo / (1 + Nd/Na);
xpo = xdo / (1 + Na/Nd);

// Calculate diffusion lengths

if (V(cap) > Vbi) begin
xn = 0;
xp = 0;
end
else begin
xn = xno * sqrt(1 - V(cap)/Vbi);
xp = xpo * sqrt(1 - V(cap)/Vbi);
end

// Calculate change in refractive index
dxn = xno - xn;
dxp = xpo - xp;

delta_neff = GammaN*dxn/wRidge*dne + GammaP*dxp/wRidge*dnh;
delta_alpha = (GammaN*dxn/wRidge*dae + GammaP*dxp/wRidge*dah)/2;

delta_invQ = 2*delta_alpha / (2*pi/lambda*ng);
delta_fo = -delta_neff/ng*(c/lambda);
fo_ring_bias = fo_ring - delta_fo;
delta_FWHM = fo_ring_bias * delta_invQ;
fwhm_bias = fwhm+delta_FWHM;
```

```
// Calculate reverse-bias capacitance
if (V(cap) > Vbi) begin
q_diode = -2 * Cjo * Vbi;
end else begin
q_diode = - 2 * Cjo * Vbi * sqrt(1 - V(cap) / Vbi);
end

// Current through resistive branch [A]
I(res) <+ Is * (limexp(V(res) / Vt) - 1);

// Current through diode cap
I(cap) <+ ddt(q_diode);

// Calculate transmissivity of through and drop ports
V(through,gnd) <+ V(in, gnd) * (1 - (1 - Tn) /
\ (1 + pow(((f_laser-fo_ring_bias)/(fwhm_bias/2)),2)));

V(drop,gnd) <+ V(in, gnd) * \
((1 - Tn)/(1 + pow((f_laser-fo_ring_bias)/(fwhm/2),2)));
        end
endmodule
```

Bibliography

[1]

[2] 2012 ITRS report: Assembly & packaging. Technical report, International Technology Roadmap for Semiconductors, 2012. `http://www.itrs.net/`.

[3] B. Analui, D. Guckenbergerand D. Kucharski, and A. Narashima. A fully integrated 20-Gb/s optoelectronic transceiver implemented in a standard 0.13-μm CMOS SOI technology. IEEE Journal of Solid State Circuits, 41(12), December 2006.

[4] C. Batten et al. Building manycore processor-to-dram networks with monolithic silicon photonics. HOTI '08, pages 21–30, Aug. 2008.

[5] J.F. Buckwalter, Xuezhe Zheng, Guoliang Li, K. Raj, and A.V. Krishnamoorthy. A monolithic 25-gb/s transceiver with photonic ring modulators and ge detectors in a 130-nm cmos soi process. Solid-State Circuits, IEEE Journal of, 47(6):1309–1322, June 2012.

[6] Intel Corporation. From sand to circuits: How intel makes integrated circuit chips, 2008.

[7] A. Emami-Neyestanak et al. A 1.6 gb/s, 3 mw cmos receiver for optical communication. VLSI '02, pages 84–87, 2002.

[8] J. Foresi, M. Black, A. Agarwal, and L. Kimerling. Losses in polycristalline silicon waveguides. American Institute of Physics Applied Physics Letters, 68(15), April 1996.

[9] K. Gabriel. "akustica", 2005.

[10] M. Georgas, J. Leu, B. Moss, C. Sun, and V. Stojanović. Addressing link-level design tradeoffs for integrated photonic interconnects. Custom Integrated Circuits Conference (CICC), September 2011.

[11] M. Georgas, J. Orcutt, R. Ram, and V. Stojanović. A monolithically-integrated optical receiver in standard 45-nm soi. European Solid-State Circuits Conference (ESSCIRC), September 2011.

[12] C. Holzwarth, J. Orcutt, H. Li, M. Popović, V. Stojanović, J. Hoyt, R. Ram, and H. Smith. Localized substrate removal technique enabling strong-confinement microphotonics in bulk si CMOS processes. Conference on Quantum Electronics and Laser Science, 2008.

[13] A. Kern, A. Chandrakasan, and I. Young. 18Gb/s optical IO: VCSEL driver and TIA in 90nm CMOS. Symposium on VLSI Circuits, October 2007.

[14] C. Kromer, G. Sialm, T. Morf, M. Schmatz, F. Ellinger, D. Erni, and H. Jäckel. A low-power 20-GHz 52-dbΩ transimpedance amplifier in 80-nm CMOS. IEEE Journal of Solid-State Circuits, 39(6), June 2004.

[15] B. Lee, B. Small, J. Foster, K. Bergman, Q. Xu, and M. Lipson. Demonstrated 4x4 Gbps silicon photonic integrated parallel electronic to WDM interface. Optical Fiber Communication Conference, 2007.

[16] G. Li, X. Zheng, J. Lexau, Y. Luo, H. Thacker, T. Pinguet, P. Dong, D. Feng, S. Liao, R. Shafiiha, M. Asghari, J. Yao, J. Shi, I. Shubin, D. Patil, F. Liu, K. Raj, R. Ho, J. Cunningham, and A. Krishnamoorthy. Ultralow-power silicon photonic interconnect for high-performance computing systems. Proceedings SPIE, 7607, January 2010.

[17] M. Lipson. Guiding, modulating, and emitting light on silicon–challenges and opportunities. Journal of Lightwave Technology, 23(12), December 2005.

[18] C. Manolatou, S. Johnson, S. Fan, P. Villeneuve, and H. Hausand J. Joannopoulos. High-density integrated optics. Journal of Lightwave Technology, 17(9), September 1999.

[19] R. Meade, O. Tehar-Zahav, Z. Sternberg, E. Megged, G. Sandhu, J.S. Orcutt, R. Ram, V. Stojanovic, M.R. Watts, E. Timurdogan, J. Shainline, and M. Popovic. Integration of silicon photonics in a bulk cmos memory flow. In Optical Interconnects Conference, 2013 IEEE, pages 114–115, May 2013.

[20] A. Mekis, S. Gloeckner, G. Masini, A. Narasimha, T. Pinguet, S. Sahni, and P. De Dobbelaecre. A grating-coupler-enabled CMOS photonics platform. IEEE Journal of Selected Topics in Quantum Electronics, 17(3), May/June 2011.

[21] G. Moore. Cramming more components onto integrated circuits. Electronics, pages 114–117, April 1965.

[22] G. Moore. Progress in digital integrated electronics. Electron Devices Meeting, 21:11–13, 1975.

[23] B. Moss, C. Sun, M. Georgas, J. Shainline, J. Orcutt, J. Leu, M. Wade, Y. Chen, K. Nammari, X. Wang, H. Li, R. Ram, M. Popović, and V. Stojanović. A 1.23 pj/b 2.5 gb/s monolithically integrated optical carrier-injection ring modulator and all-digital driver circuit in commercial 45nm SOI. IEEE International Solid-State Circuits Conference, February 2013.

[24] A. Narasimha, A. Behnam, Y. Liang, T. Sleboda, and C. Gunn. A fully integrated 4x10 Gb/s DWDM optoelectronic transceiver in a standard 0.13μm CMOS SOI. IEEE International Solid State Circuits Conference, February 2007.

[25] J. Orcutt, B. Moss, C. Sun, J. Leu, M. Georgas, J. Shainline, E. Zgraggen, H. Li, J. Sun, M. Weaver, S. Urosevic, M. Popović, R. Ram, and V. Stojanović. An open foundry platform for high-performance electronic-photonic integration. (Submitted - Preprint), 2012.

[26] J. Orcutt, A. Khiloand M. Popović, C. Holzwarth, H. Li, M. Dahlem, B. Moss, F. Kärtner, E. Ippen, J. Hoyt, V. Stojanović, and R. Ram. Demonstration of an electronic photonic integrated circuit in a commercial scaled bulk CMOS process. Conference on Lasers and Electro-Optics, 2008.

[27] S. Palermo et al. A 90 nm CMOS 16 Gb/s transceiver for optical interconnects. JSSCC, 43(5):1235–1246, May 2008.

[28] D. Pham et al. Overview of the architecture, circuit design, and physical implementation of a first-generation Cell processor. JSSC, 41(1):179–196, 2005.

[29] R. Pierret. Semiconductor Device Fundamentals. Addison-Wesley Publishing Company, Reading, Massachusetts, 1996.

[30] S. Pradhan, B. Schmidt, L. Martinez, Q. Xu, V. Almeida, C. Barrios, and M. Lipson. Electro-optic modulator on silicon-on-insulator substrates using ring resonators. CLEO, 2005.

[31] Reid J. Riedlinger et al. A 32nm 3.1 billion transistor 12-wide-issue Itanium; processor for mission-critical servers. In ISSCC '11, pages 84 –86, feb. 2011.

[32] J.M. Shainline et al. Depletion-mode carrier-plasma optical modulator in zero-change advanced CMOS. Optics Letters, 24(14), August 2013.

[33] R. Soref and B. Bennett. Electrooptical effects in silicon. IEEE Journal of Quantum Electronics, 23(1), January 1987.

[34] D. Staelin. Electromagnetics and Applications. Massachusetts Institute of Technology, Cambridge, MA, 2008.

[35] V. Stojanović, K. Asanović, J. Hoyt, R. Ram, F. Kärtner, K. Berggren, M. Schmidt, H. Smith, and E. Ippen. Integrated photonic networks for compact, energy-efficient supercomputers. Project review, Massachusetts Institute of Technology, February 2007.

[36] A. Strollo. A new SPICE subcircuit model of power P-I-N diode. IEEE Transactions on Power Electronics, 9(6), November 1994.

[37] T. Tamir. Integrated Optics, volume 7 of Topics Appl. Phys. Springer, Berlin, second edition, 1979.

[38] M. Watts, D. Trotter, R. Young, A. Lentine, and W. Zortman. Limits to silicon modulator bandwidth and power consumption. Proc. of SPIE, 7221, 2009.

[39] Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya, and M. Lipson. 12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators. Optical Society of America, 15(2), January 2007.

[40] I. Young, E. Mohammed, J. Liao, A. Kern, S. Palermo, B. Block, M. Reshotko, and P. Chang. Optical i/o technology for tera-scale computing. IEEE Solid-State Circuits Conference, 2009.

[41] I. Young, E. Mohammed, J. Liao, A. Kern, S. Palermo, B. Block, M. Reshotko, and P. Chang. Optical i/o technology for tera-scale computing. IEEE Journal of Solid-State Circuits, 45(1), January 2010.

[42] X. Zheng, J. Lexau, Y. Luo, H. Thacker, T. Pinguet, A. Mekis, G. Li, J. Shi, P. Amberg, N. Pinckney, K. Raj, R. Ho, J. Cunningham, and A. Krishnamoorthy. Ultra-low-energy all-CMOS modulator integrated with driver. Optics Express, 18(3), February 2010.

[43] Xuezhe Zheng, Ying Luo, J. Lexau, F. Liu, Guoliang Li, H.D. Thacker, I. Shubin, Jin Yao, R. Ho, J.E. Cunningham, and A.V. Krishnamoorthy. 2-pj/bit (on-chip) 10-gb/s digital cmos silicon photonic link. Photonics Technology Letters, IEEE, 24(14):1260–1262, July 2012.