To my father, who taught me the pleasure of finding things out; to my mother for her unconditional love and support; and to all my teachers, mentors, and advisors for their continued support in my more than 20 years of study.
Acknowledgments

First of all, I would like to express my deepest gratitude towards my advisor, Prof. Azita Emami, for her patience and excellent guidance over the years. The supportive and collaborative environments of MICS and Caltech were crucially important factors in helping me realize my potential. I would like to acknowledge the assistance and mentorship of my fellow senior labmates, Dr. Meisam Honarvar, Dr. Mayank Raj, and Manuel Monge. I would like to thank the rest of my defense committee, including Prof. Ali Hajimiri, Prof. Kerry Vahala, Prof. David Rutledge, and Prof. Hyuck Choo, for their inspirational interactions, scholarly advice, and timely suggestions.

I am grateful that I have had the opportunity to work with some of the finest researchers of our time. Guidance from Prof. Ali Hajimiri and help from his students Behrooz Abiri and Aroutin Khachaturian played a key role in my achievements during my Ph.D studies. I had the privilege to work with Dr. Sylvie Menezo and her team at CEA Leti, France. Their vision and fantastic discussions enabled my key contributions and cutting-edge research. During my last year collaborations with Rockley Photonics, I was very lucky to meet and learn from many topnotch scientists and researchers. I would like to thank Dr. Andrew Rickman, one of the most successful entrepreneurs of our time, for his mentorship and trust in me. Immeasurable appreciation for the help and support I received from Dr. Aaron Zilkie, Dr. Guomin Yu, Hooman Abediasl, David Nelson, and many other smart and kind researchers and engineers at Rockley Photonics.

I also thank the following current and past members of MICS and CHIC groups at Caltech: Dr. Matthew Loh, Dr. Juhwan Yoo, Krishna Settaluri, Abhinav Agarwal, Kuan-chang Chen, Dr. Mahsa Shoaran, Amir Safaripour, Dr. Alex Pai, Prof. Kaushik Sengupta, Dr. Kaushik Dasgupta, Prof. Firooz Aflatooni, Prof. Steven Bowers, Dr. Florian Bohn, Constantine Sideris, Brian Hong, Reza Fatemi, Matan Gal-Katziri, and Milad Taghavi. They all have been amazing companions and friends over the years. I would like to acknowledge Michelle Chen, Tanya Owen, and Carol Sosnowski for their administrative support and specially thank Laura Flower Kim and Daniel Yoder for making Caltech such a welcoming environment for international students.

I wish to thank the National Science Foundation, IFC, and Rockley Photonics for their financial support. Also, donated resources from ST Microelectronics, Leti, OpSIS, CMP, and Cosemi Technologies enabled this research.
Finally, I want to thank my parents for working their best to provide me the opportunities to pursue my dreams and interests, and also my sister for her support and encouragement throughout my PhD studies.
Abstract

Integrated circuit scaling has enabled a huge growth in processing capability, which necessitates a corresponding increase in inter-chip communication bandwidth. As bandwidth requirements for chip-to-chip interconnection scale, deficiencies of electrical channels become more apparent. Optical links present a viable alternative due to their low frequency-dependent loss and higher bandwidth density in the form of wavelength division multiplexing. As integrated photonics and bonding technologies are maturing, commercialization of hybrid-integrated optical links are becoming a reality. Increasing silicon integration leads to better performance in optical links but necessitates a corresponding co-design strategy in both electronics and photonics. In this light, holistic design of high-speed optical links with an in-depth understanding of photonics and state-of-the-art electronics brings their performance to unprecedented levels. This thesis presents developments in high-speed optical links by co-designing and co-integrating the primary elements of an optical link: receiver, transmitter, and clocking.

In the first part of this thesis a 3D-integrated CMOS/Silicon-photonic receiver will be presented. The electronic chip features a novel design that employs a low-bandwidth TIA front-end, double-sampling and equalization through dynamic offset modulation. Measured results show -14.9dBm of sensitivity and energy efficiency of 170fJ/b at 25Gb/s. The same receiver front-end is also used to implement source-synchronous 4-channel WDM-based parallel optical receiver. Quadrature ILO-based clocking is employed for synchronization and a novel frequency-tracking method that exploits the dynamics of IL in a quadrature ring oscillator to increase the effective locking range. An adaptive body-biasing circuit is designed to maintain the per-bit-energy consumption constant across wide data-rates. The prototype measurements indicate a record-low power consumption of 153fJ/b at 32Gb/s. The receiver sensitivity is measured to be -8.8dBm at 32Gb/s.

Next, on the optical transmitter side, three new techniques will be presented. First one is a differential ring modulator that breaks the optical bandwidth/quality factor trade-off known to limit the speed of high-Q ring modulators. This structure maintains a constant energy in the ring to avoid pattern-dependent power droop. As a first proof of concept, a prototype has been fabricated and measured up to 10Gb/s. The second technique is thermal stabilization of micro-ring resonator modulators through direct measurement of temperature using a monolithic PTAT temperature sensor.
The measured temperature is used in a feedback loop to adjust the thermal tuner of the ring. A prototype is fabricated and a closed-loop feedback system is demonstrated to operate at 20Gb/s in the presence of temperature fluctuations. The third technique is a switched-capacitor based pre-emphasis technique designed to extend the inherently low bandwidth of carrier injection micro-ring modulators. A measured prototype of the optical transmitter achieves energy efficiency of 342fJ/bit at 10Gb/s and the wavelength stabilization circuit based on the monolithic PTAT sensor consumes 0.29mW.

Lastly, a first-order frequency synthesizer that is suitable for high-speed on-chip clock generation will be discussed. The proposed design features an architecture combining an LC quadrature VCO, two sample-and-holds, a PI, digital coarse-tuning, and rotational frequency detection for fine-tuning. In addition to an electrical reference clock, as an extra feature, the prototype chip is capable of receiving a low jitter optical reference clock generated by a high-repetition-rate mode-locked laser. The output clock at 8GHz has an integrated RMS jitter of 490fs, peak-to-peak periodic jitter of 2.06ps, and total RMS jitter of 680fs. The reference spurs are measured to be 64.3dB below the carrier frequency. At 8GHz the system consumes 2.49mW from a 1V supply.
## Contents

Acknowledgments .......................................................... iv

Abstract ........................................................................... vi

1 Introduction ................................................................... 1
  1.1 Organization ........................................................... 5

2 Background ..................................................................... 8
  2.1 Optical Interconnect Basics ........................................... 8
    2.1.1 Basic Definitions ................................................ 9
    2.1.2 Clocking ............................................................ 11
      2.1.2.1 Jitter ........................................................... 11
      2.1.2.2 Subrate Clocking ........................................... 12
    2.1.3 Fundamental Integrated Photonic Building Blocks ....... 12
      2.1.3.1 Passive Components and Waveguides ............... 13
      2.1.3.2 Modulators ................................................ 14
      2.1.3.3 Photodetectors .............................................. 14
  2.2 Opportunities In Silicon Photonics ............................... 16

2.3 Hybrid Integration Of Electronics and Photonics ............ 19

3 Optical Receivers ........................................................... 21
  3.1 High-speed Optical Receiver Overview ......................... 21
    3.1.1 Photodetector Model ........................................... 22
  3.2 Prior Art in Design of Optical Receivers ....................... 23
    3.2.1 Transimpedance Amplifiers ................................ 23
    3.2.2 Integrating Front-ends ....................................... 25
  3.3 3D-Integrated CMOS/SiP Optical Receiver With a LBW TIA Integrating Front-end 28
    3.3.1 3D-Integration Platform ...................................... 31
    3.3.2 Receiver Architecture Overview ........................... 32
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.3</td>
<td>Circuit Implementation</td>
<td>35</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Experimental Results</td>
<td>36</td>
</tr>
<tr>
<td>3.3.5</td>
<td>Performance Analysis And Design Guidelines</td>
<td>40</td>
</tr>
<tr>
<td>3.4</td>
<td>An Adaptive 3D-Integrated CMOS/SiP Parallel Optical Receiver with Injection-Locked QuarterRate Clocking</td>
<td>45</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Body-biasing</td>
<td>45</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Quadrature-locked loop</td>
<td>45</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Receiver Architecture Overview</td>
<td>48</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Circuit Implementation</td>
<td>49</td>
</tr>
<tr>
<td>3.4.5</td>
<td>Measurement Results</td>
<td>51</td>
</tr>
<tr>
<td>3.4.5.1</td>
<td>QLL Measurements</td>
<td>52</td>
</tr>
<tr>
<td>3.4.5.2</td>
<td>Optical Receiver Measurements</td>
<td>52</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary</td>
<td>57</td>
</tr>
</tbody>
</table>

4 Optical Transmitters

| 4.1     | High-speed Optical Transmitter Overview                                                         | 58   |
| 4.1.1   | Direct Modulation vs External Modulation                                                       | 58   |
| 4.2     | Mico-ring Modulators: Opportunities And Challenges                                               | 60   |
| 4.3     | Monolithic Silicon-photonic PTAT temperature sensor for micro-ring resonator thermal stabilization | 63   |
| 4.3.1   | Structure Overview                                                                             | 63   |
| 4.3.2   | Supporting Chip                                                                                | 66   |
| 4.3.3   | Measurement Results                                                                            | 66   |
| 4.4     | Differential Optical Ring Modulator:                                                           | 71   |
| 4.4.1   | DRM Structure Overview                                                                         | 72   |
| 4.4.2   | Supporting Chip                                                                                | 74   |
| 4.4.3   | Measurement Results                                                                            | 74   |
| 4.5     | Carrier-injection Micro-ring Modulator Transmitter with Switched-capacitor Pre-emphasis        | 76   |
| 4.5.1   | Transmitter Architecture Overview                                                              | 77   |
| 4.5.2   | Supporting Circuits                                                                            | 79   |
| 4.5.3   | Measurement Results                                                                            | 81   |
| 4.6     | Summary                                                                                        | 84   |

5 Clock Generation

| 5.1     | Overview of Clock Generation for Optical Links                                                 | 85   |
| 5.2     | Prior Art In Low-jitter Clock Generation                                                       | 86   |
List of Figures

1.1 Original graph by Gordon Moore in the 1965 Electronics Magazine article in which he made a prediction about the semiconductor industry [1]......................................................... 1
1.2 Microprocessor core count scaling (left) and microprocessor clock frequency scaling (right) [3] (ISSCC trends, 2012). ................................................................. 2
1.3 (a) Per-pin data-rate vs. year for a variety of common IO standards. (b) Data-rate vs. process node and year. [4] ................................................................. 3
1.4 Examples of complex electronic systems. (a) A high-performance multi-core processor (Intel 80 Core Terraflop). (b) A server rack in a data center. ......................... 4
1.5 3D-integrated CMOS/Silicon photonic optical receiver developed by researchers at Caltech and CEA Leti [6]. ................................................................. 5

2.1 Basic elements of an optical interconnect. ....................................................... 8
2.2 Eye diagram, showing fundamental definitions and sources of noise. ............... 9
2.3 Diagram for BER vs SNR due to noise with Gaussian distribution. ............... 10
2.4 (a) Strip waveguide. (b) Rib waveguide. ....................................................... 13
2.5 Techniques for coupling light to a photonic integrated circuit: (a) Grating coupler, (b) Edge-coupled optical fiber to an on-chip waveguide. (c) V-groove coupling of optical fiber to an on-chip waveguide ....................................................... 13
2.6 (a) Packaged passive silicon-photonic chip (Courtesy Imec). (b) Silicon-photonic chip with active devices (Courtesy OpSIS). (c) Monolithic CMOS/silicon-photonic chip (Courtesy IBM). ....................................................... 16
2.7 Die micrograph of a complex SOI photonic chip fabricated at IME A*STAR (Courtesy IME). ....................................................... 17
2.8 Hybrid integration schemes. (a) Wirebonding. (b) Flip-chip bonding. (c) Through-silicon vias (TSV). ....................................................... 19

3.1 Optical receivers are one of the primary elements of optical interconnects. ........ 21
3.2 Equivalent circuit model of a reverse-biased photodiode. ............................ 22
3.3 Different TIA front-end topologies: (a) Common-gate TIA. (b) Regulated cascode TIA. (c) Common-source shunt-shunt TIA. (d) Common-source shunt-shunt TIA with direct feedback. (e) Inverter-based TIA.

3.4 Data resolution in integrating front-ends.

3.5 A shunt-shunt resistive feedback TIA model.

3.6 In double sampling receivers with resistive front-end sensitivity is limited by charge sharing and kT/C noise.

3.7 Recently published optical receivers.

3.8 (a) Layout of the silicon photonic chip. (b) Die micrograph of the fabricated silicon photonic chip.

3.9 Measured photo-detector capacitance as a function of bias voltage.

3.10 (a) Scanning electron microscope top view of the CuP grown on the electronic wafer. (b) Cross section of the micro-pillar.

3.11 Top-level architecture of (a) integrating receiver with low-bandwidth TIA (b) TIA-based receiver.

3.12 Z-domain block diagram of the integrating optical receiver with low-bandwidth TIA.

3.13 Circuit-level implementation of individual building blocks.

3.14 (a) 3D-integrated CMOS/silicon-photonic optical receiver. (b) CMOS chip die micrograph. (c) Grating coupler and polarization splitter on the silicon photonic chip. (d) Top-view of the 3D-integrated optical receiver.

3.15 Optical measurement setup.

3.16 (a) Input optical eye diagram at 25Gb/s for integrating receiver. (b) Output recovered and de-multiplexed eye diagram at 6.25Gb/s for the integrating receiver. (c) Input optical eye diagram at 21.2Gb/s for the conventional receiver. (d) Output recovered and de-multiplexed eye diagram at 5.3Gb/s for the conventional receiver.

3.17 Optical receiver measurement results. (a) Sensitivity of integrating and conventional receivers bonded with CuP and wire-bond. (b) Energy efficiency and power consumption of the optical receiver. (c) Power consumption breakdown for integrating and conventional receiver.

3.18 Sensitivity curves of the integrating optical receiver.

3.19 Measured bathtub curves of the integrating receiver and conventional TIA-based receiver.

3.20 Noise sources for (a) integrating optical receiver and (b) conventional TIA-based optical receiver.

3.21 Simulated sensitivity vs. speed for integrating and conventional optical receivers.
3.22 (a) Simulated sensitivity vs. PD capacitor for RC and low-bandwidth TIA front-end at 25Gb/s. (b) FoM vs. PD capacitor for RC and low-bandwidth TIA front-end at 25Gb/s. ................................. 44
3.23 (a) FD SOI MOS structure. (b) Threshold voltage ($V_{th}$) variation with back bias ($V_b$). 46
3.24 Quadrature phase error of a two-stage ILO in the unlocked case (a) close to lock and (b) far from lock. (c) Mean quadrature phase error vs $f_0$ (d) Effect of injection strength on MQPE. ........................................ 46
3.25 A block diagram of the quadrature-locked loop. ........................................ 47
3.26 Architecture of a single channel of the optical receiver (single channel) and body-biasing circuit ................................................................. 48
3.27 Receiver clock distribution architecture. ............................................................ 49
3.28 Schematic circuit implementation of the proposed Quadrature Locked Loop (QLL). . 50
3.29 Simulated ring oscillator characteristics. .............................................................. 51
3.30 Body-bias generator circuit. ............................................................................. 51
3.31 Test setup for parallel optical receiver. .............................................................. 52
3.32 Measured jitter transfer function and VDD noise transfer function (left). Measured phase noise of the reference, and oscillator in unlocked and locked states (right) at 8GHz for 32GB/s data. .................................................. 53
3.33 Measured quadrature waveforms. ...................................................................... 53
3.34 Chip micrograph and layout details. ................................................................. 54
3.35 (a) Measured input eye diagram at 32Gb/s. (b) Output eye diagram and recovered clock for PRBS-15 optical data at 32Gb/s. (c) Bathtub curves for 20Gb/s and 32Gb/s. 54
3.36 BER versus optical power (receiver sensitivity) at different data-rates (top). Optical sensitivity vs. data rate (bottom). ......................................................... 55
3.37 Measured power consumption breakdown and energy-per-bit. ......................... 56
4.1 Optical transmitters are one of the primary elements of optical interconnects. .... 58
4.2 Conceptual block diagrams of optical transmitters. (a) Direct modulation of the laser (b) External modulation of the laser. ................................................. 59
4.3 Index-modulation optical ring modulators and different ways of changing the index of refraction. ................................................................. 60
4.4 Operation of Micro-ring modulator. ................................................................. 61
4.5 Diagram of micro-ring modulator with electric field in various locations. .......... 62
4.6 Proposed structure of ring resonator with integrated heater and PTAT temperature sensor. ................................................................. 64
4.7 Concept of a feedback loop to stabilize micro-ring’s temperature. ..................... 65
4.8 COMSOL heat transfer simulations. (a) Temperature uniformity across distributed PTAT sensor. (b) Effect of temperature gradient due to heater on PTAT sensor.

4.9 Die micrograph of the fabricated micro-ring modulator with integrated heater and PTAT temperature sensor.

4.10 (a) DC static transmission of the micro-ring. (b) Measured integrated heater tunability of the micro-ring. (c) Measured PTAT voltage versus temperature. (d) Measured micro-ring resonance wavelength versus temperature.

4.11 High-speed optical response of the micro-ring (a) at 10Gb/s (b) at 20Gb/s.

4.12 High-speed measurement setup with induced thermal fluctuations and temperature stabilization feedback loop.

4.13 (a) Peltier heater/cooler supplied current over time. (b) Closed loop integrated heater voltage and PTAT voltage. (c) Output optical eye diagram without thermal tuning feedback. (d) Output eye diagram with thermal tuning feedback.

4.14 Index-modulated micro-ring’s (a) static transmission, (b) optical frequency response, (c) simulated Q versus -3dB bandwidth.

4.15 (a) Index-modulated ring. (b) Coupling-modulated ring. (c) Proposed differential ring modulator.

4.16 Block diagram of the differential ring modulator.

4.17 Die micrograph of the fabricated prototype. (a) Optical input grating coupler. (b) Y-junction. (c) Heater for phase shift controller. (d) Photodiode connected to one output for testing purposes. (e) Optical output grating coupler.

4.18 Measurement setup.

4.19 Output optical eye diagram of the differential ring modulator operating (a) At 5Gb/s (b) At 10Gb/s.

4.20 (a) Measured steady-state transmission of the micro-ring near a resonance wavelength (b) Tunability of the micro-ring using a heater.

4.21 Carrier-injection micro-ring modulator structure.

4.22 Top-level block diagram of the transmitter and the proposed switched-capacitor-based pre-emphasis technique.

4.23 Schematic circuit details of the proposed micro-ring modulator driver with switched-capacitor pre-emphasis.

4.24 Layout of the silicon photonic micro-ring modulator with on-chip PTAT temperature sensor. Circuit implementation details of the current bandgap and measured micro-ring modulator resonance when bias is varied. Schematic block diagram of the feed-forward bias-based wavelength stabilization technique.
4.25 CMOS chip micrograph (top left) and silicon photonic chip micrograph (bottom). The layouts of the driver (top right) and layout for tuning circuitry (bottom right). 80

4.26 Measured characteristics of the micro-ring modulator. Measured output optical eye diagram of the optical transmitter with and without pre-emphasis. 81

4.27 Characteristics of the PTAT sensor and temperature-dependency of the micro-ring modulator’s resonance wavelength. Output optical eye diagram of the micro-ring modulator with and without wavelength tuning stabilization in presence of emulated ambient temperature noise. 82

5.1 Clock generators are one of the primary elements of optical interconnects. 85

5.2 Example of prior art. (a) Conventional multiplying PLL (MPLL). (b) Edge-combining DLL-based frequency synthesizer. (c) MDLL. (d) Injection-locked-MPLL (IL-MPLL). 86

5.3 Proposed first-order frequency synthesizer. 89

5.4 Top-level architecture of the first-order frequency synthesizer. 90

5.5 Principle of phase-interpolation based reference injection. 91

5.6 Digital coarse-tuning flow chart. 92

5.7 Principle of beat frequency in fine-tuning frequency detection. 93

5.8 Transistor-level schematics of main building blocks. (a) LC quadrature VCO. (b) Master-slave S/H. (c) Differential phase-interpolator. 95

5.9 Details of fine-tuning frequency-detection. 96

5.10 Electrical/optical reference generator. 96

5.11 Real-time acquisition for a reference clock of 400MHz and output clock of 8GHz. 97

5.12 (a) Measurement showing reference spurs at 8GHz with 400MHz reference. (b) Output clock jitter histograms. (c) Total jitter measurement. (d) Phase noise measurement. 97

5.13 Total RMS jitter of output clock versus frequency. 98

5.14 Quadrature LC oscillator. (a) Simulated Q Versus frequency. (b) Measured phase noise of open-loop LC VCO. (c) Quadrature phase mismatch. (d) Coarse-tuning code-word versus frequency. 99

5.15 Measured nonlinear characteristics of the S/H and PI. 101

5.16 (a) Power consumption and efficiency at different frequencies. (b) Power breakdown at 8 GHz. 102

5.17 Die micrograph and layout details of the implemented prototype. (a) Design of inductor. (b) Digital coarse-tuning block. (c) Placement of the rest of the circuitry between the inductors. 103

5.18 Block diagram of the system and filtering of VCO phase noise in the presence of a clean reference. 104
5.19  Loop response measured by applying a step frequency to the reference clock. . . . . . 105
List of Tables

3.1 Performance Summary and Comparison of Optical Receivers ...................... 40
3.2 Performance Summary of the Parallel Optical Receiver ............................ 56
4.1 Optical Transmitter Performance Summary ............................................. 83
5.1 Frequency Synthesizer Performance Summary ........................................ 102
Chapter 1

Introduction

In his 1965 Electronics Magazine article, Grodon Moore observed that the economically-efficient number of transistors per integrated circuit doubled approximately every year from 1962 to 1965 and predicted that this trend could be sustained [1] (Fig. 1.1). Fifty years later, perhaps to his own surprise, Moore’s law continues\(^1\) to propel the greatest economic engine of the modern technological age [2]. A higher level of integration and further miniaturization of transistors drove (at

Figure 1.1: Original graph by Gordon Moore in the 1965 Electronics Magazine article in which he made a prediction about the semiconductor industry [1].

\(^1\)Moore’s law was later revised to doubling transistors every 18-24 months.
least initially) a rise in clock speeds and power efficiency, resulting in more powerful processors and higher capacity communication units. Even a slowdown in voltage scaling and the advent of power consumption as a major obstacle in clock speed scaling have not prevented continued advancement in computational power; increases in transistor density have enabled designers to work around these limits by adopting new approaches such as multi-core and heterogeneous computing (Fig. 1.2, 1.4). For more than a decade, these new approaches and architectures have increased the input output (IO) aggregate bandwidth requirements at a rate of about $2 \times$ to $3 \times$ every two years. From mobile devices to supercomputers, wireline IO has been instrumental in enabling the incredible scaling of computers. This demand is driven by a variety of applications in computing and networking, including chip-to-chip interconnects, memory-processor interface, graphics, backplane, rack-to-rack, and LAN. The increase in bandwidth is enabled partly by increasing the number of IO pins per component. As a result, IO circuitry consumes an increasing amount of area and power on today’s chips. Increasing bandwidth has also been enabled by rapidly accelerating the per-pin data-rate. Fig. 1.2(a) shows that data-rate per-pin has doubled approximately every four years across a number of IO standards [4]. Fig. 1.2(b) shows data-rates for published transceivers enabled in part by process-technology scaling [4]. The continued CMOS scaling and rise of parallel computing has led to unprecedented strain on electrical channels. As Moore’s law continues to improve speed of on-chip data processing and computation, demand for high-bandwidth IO grows. The required bandwidth is achieved by both increasing the data-rate of each IO pin and increasing the number of IO pins per chip.

Technology scaling helps with increasing the speed and minaturization of IO transceiver circuitry and on-chip clocking. However, the number of pins does not scale as fast due to physical and packaging limitations, and the performance of electrical links remains limited by channel characteristics. The dielectric and resistive losses of printed circuit board (PCB) traces increase at higher frequen-
cies. Such frequency dependent attenuation causes inter-symbol interference (ISI) and ultimately degradation of signal-to-noise-ratio (SNR). In addition, reflections from discontinuities in the signal path due to transitions from chip-to-package and package-to-board generate more ISI and further reduce the SNR. These problems are exacerbated as the data rate increases. Beside, as the number of interconnects increases, the spacing between channels decreases to allow for accommodating more channels and therefore achieving higher aggregate data-rate. Higher interconnect density results in excessive capacitive and inductive coupling between adjacent channels, which results in crosstalk noise and signal degradation.

The frequency-dependent loss in the copper wires has long motivated the use of optical communication for transmission of data over long distances. Higher bandwidth density, improved SNR, fewer amplifier stages along the signal path, and lower cost have made optical fiber communication the ideal choice for longer links. Although optical interconnects have been around for sometime, multi-Gb/s short links are still dominated by electrical interconnects. In order to continue scaling data rates, equalization techniques can be employed to compensate for the ISI. However, the power and area overhead associated with equalization make it difficult to achieve target data-rates with a realistic power budget. As a result, rather than being technology limited, current high-speed IO link designs are becoming channel and power limited. Integrated photonics and particularly silicon photonics, leveraging technologies developed in CMOS industry over decades, have the potential to enable a breakthrough in optical interconnects over short distances [5]. The negligible frequency dependent loss of optical channels enables higher data-rate links that can be scaled with on-chip CMOS transistor enhancements without excessive equalization complexity and power consumption. Also, high-density hybrid integration and wavelength division multiplexing (WDM) in optical domain allows very high information density.

Low-parasitic, dense hybrid integration is a key component of successful implementation of high-speed optical interconnects (Fig. 1.5). The advantage of hybrid integration over the monolithic

![Figure 1.3](image)

Figure 1.3: (a) Per-pin data-rate vs. year for a variety of common IO standards. (b) Data-rate vs. process node and year. [4]
approach is that the material and process developed for electronics and photonics are optimized differently. As Moore’s law continues to shrink transistor sizes, in high-speed applications it is imperative to choose the most advanced CMOS technology and keep in mind that there is no Moore’s law (in the form of scaling the devices) in photonics. Size of on-chip waveguides are limited to the wavelength of light and scaling photonic components such as detectors and modulators often deteriorates their performance.

Designing a high-performance optical interconnect entails optimizing metrics with trade-offs entangled between electronics and photonics. For example, total power consumption of an optical link is the sum of laser power and the power consumed in electronics. Higher sensitivity receiver front-ends that consume higher power reduce the required laser power. Depending on photonic device performance and losses in optical couplings, the overall optical link power lies somewhere in the trade-off between receiver’s sensitivity and power. Accordingly, a hollistic co-design of electronics and photonics is necessary to avoid sub-optimal performance in high-speed optical interconnects. In doing so, it is crucial to design the electronic circuitry with a deep understanding of photonics. On the other hand, it is important to design the photonic devices with an eye on available state-of-the-art electronics.
1.1 Organization

This dissertation covers three fundamental elements of optical interconnect with an emphasis on opportunities for holistic design. Chapter 2 provides a background for design of optical interconnects. In this chapter, first the metrics used for gauging the performance of an optical interconnect are reviewed. An overview of fundamental building blocks of a high-speed optical interconnect is presented and recent advancements and opportunities for silicon photonics are discussed. Finally, motivations for hybrid-integration of electronics and photonics and holistic design are elaborated on. Chapter 3 provides an introduction to optical receivers and a review of prior art in designing optical receivers. We show how the choice of sensitivity and electronic power through design topology and optimization determines the overall link power consumption. Considering this trade-off, a 3D-integrated CMOS/Silicon photonics optical receiver [6,7] is presented. The 3D integration is based on Copper Pillar flipchip technology, enabling low parasitic capacitance and 40µm pitch for interconnection [8]. The receiver architecture features a novel integrating architecture combining a low-bandwidth TIA front-end double-sampling and dynamic offset modulation [9]. A measured prototype achieves -15dBm of sensitivity and 170fJ/bit of energy efficiency at 25Gb/s. We study different trade-offs in designing an optical receiver and how to choose between a full-bandwidth TIA front-end and integrating architecture using a resistive front-end or a low-bandwidth TIA front-end [10]. The design methodology is supported by measurements of two 3D-integrated prototypes based on a conventional TIA and a double-sampling integrating receiver. In a follow-up work, a 3D-integrated source-synchronous 4-channel parallel optical receiver is presented [11]. A quadrature-locked loop (QLL) based clocking is used to generate accurate clock phases for a 4-channel optical receiver using a forwarded clock at quarter-rate [12]. An adaptive body biasing circuit is designed to maintain the per-bit energy consumption across wide data-rates. The prototype measurements indicate a record-

Figure 1.5: 3D-integrated CMOS/Silicon photonic optical receiver developed by researchers at Caltech and CEA Leti [6].
The low power consumption of 4.87mW per-channel at 32Gb/s (153fJ/bit). The receiver sensitivity is measured to be -8.8dBm at 32Gb/s.

In Chapter 4, first an overview of optical transmitters is presented. We review the fundamental operation and theory behind ring-resonator structures and the difficulties of using them as modulators. We introduce three solutions for primary challenges associated with implementation of micro-ring modulator CMOS/silicon photonic optical transmitters. A differential ring modulator (DRM) is presented [13] that breaks the trade-off between bandwidth and quality factor of micro-rings, which is known to limit the speed of high-Q ring modulators. The DRM maintains the energy stored in the ring constant, unlike coupling modulation schemes, and hence does not suffer from power droop when long sequences of identical bits are transmitted. As a proof of concept a prototype has been fabricated and measured to operate up to 10Gb/s. In this chapter, we also present a scheme for thermal stabilization of micro-ring resonator modulators through direct measurement of ring temperature using a monolithic PTAT temperature sensor [14,15]. The measured temperature is used in a feedback loop to adjust the thermal tuner of the ring. This scheme obviates the need for constantly tapping a portion of output optical power for monitoring or complex circuitry. The closed-loop feedback system is demonstrated to operate in the presence of thermal perturbations at 20Gb/s. Lastly, a CMOS-silicon photonic optical transmitter based on carrier-injection ring modulators is presented. It features a novel low-power switched-capacitor-based pre-emphasis that effectively compensates for the modulator bandwidth limitation. A feedforward wavelength stabilization technique via direct measurement of ring temperature using a monolithic PTAT sensor is also presented. The optical transmitter achieves energy efficiency of 342fJ/bit at 10Gb/s and the wavelength stabilization circuit consumes 0.29mW.

In Chapter 5 of this dissertation we take a look at clock generation for optical and electrical links. An overview of prior art in clock generation schemes and their shortcomings is presented. A novel low-power first-order frequency synthesizer architecture is presented as being suitable for high-speed on-chip clock generation [16,17]. The proposed design features an architecture combining an LC quadrature voltage controlled oscillator (VCO), two sample-and-holds, a phase interpolator, digital coarse-tuning, and rotational frequency detection for fine-tuning. Similar to multiplying delay-locked loops (MDLL), this architecture limits jitter accumulation to one reference cycle, as jitter during one reference cycle does not contribute to the next reference cycles. Also, instead of using multiplexer switches commonly employed in multiplying delay-locked loops (MDLL), the reference clock edge is injected by phase interpolation to support higher frequencies and lower jitter. Functionality of the frequency synthesizer is validated between 8-9.5GHz, LC VCOs range of operation. First order dynamic of the acquisition has been analyzed and demonstrated through measurement. The output clock at 8GHz has an integrated RMS jitter of 490fs, peak-to-peak periodic jitter of 2.06ps, and total RMS jitter of 680fs. Different components of jitter have been analyzed and separate measurements
have been done to support the analysis. Finally, in Chapter 6 we summarize the conclusions of this work.
Chapter 2

Background

2.1 Optical Interconnect Basics

Figure 2.1: Basic elements of an optical interconnect.

Figure 2.1 shows the components and configuration of the most basic clocked optical link: transmitter, free-space or guided-wave interconnection medium as channel and a receiver, with clocks at each end to time the data. Optical transmitters are based on direct modulation of the laser or separate modulators and a continuous wave (CW) laser. The optical signal sent through the channel exists in the continuous time, analog optica power domain. Therefore, both transmitter and receiver have an analog nature. The transmitter converts the data from electrical domain to optical domain with the goal of maintaining the highest signal integrity. The purpose of optical receiver is to convert the optical power to electrical signal and determine the optimum decision point, in time and amplitude, to estimate the original bit-stream and minimize errors.
2.1.1 Basic Definitions

In the simplest form of amplitude modulation in an optical link, the binary levels of “1” and “0” are defined as optical power amplitudes, $P_1$ and $P_0$. The extinction ratio, expressed as a fraction in dB, is defined as $r_e = \frac{P_1}{P_0}$. The optical modulation amplitude (OMA) is the difference between two optical power levels associated with “1” and “0”: $OMA = P_1 - P_0$. Given the average optical power, $P_{av} = (P_1 + P_0)/2$, the relation between extinction ratio and OMA can be derived as

$$OMA = 2P_{av} \frac{r_e - 1}{r_e + 1}.$$  \hspace{1cm} (2.1)

When the extinction ratio is high, $OMA \approx 2 \times P_{av}$, but as the extinction ratio drops the OMA, or that portion of optical power that is usable, decreases. Therefore, higher extinction ratio is desirable to reduce the overall power of an optical link. A helpful and common tool for visualizing the effects of noise and jitter on an optical link is the optical eye diagram, generated by superimposing many bit-time intervals (also known as unit interval (UI)) of the data on one another (Fig. 2.2).

![Eye diagram](image)

Figure 2.2: Eye diagram, showing fundamental definitions and sources of noise.

In an optical link, the received signal is the sum of the transmitted values and noise which appears as an added signal with random value. At the sampling point, there is a small but finite probability for the noise amplitude to be greater than the signal amplitude. This probability determines the probability of a wrong decision or the bit-error rate (BER). The BER indicates how many errors are
likely to occur for a certain number of resolved bits. For example, the probability of error due to additive white Gaussian noise (AWGN) can be expressed as a function of the signal-to-noise ratio (SNR) in case of an equiprobable one or zero [18]:

\[ BER = P_{\text{error}} = \int_{A}^{\infty} \frac{1}{2\pi \sigma_n^2} \exp\left(\frac{-x^2}{2\sigma_n^2}\right) dx = 1 - Q\left(\frac{A}{\sigma_n}\right) = 1 - Q(SNR), \] 

(2.2)

where \( A \) is the signal amplitude, \( \sigma_n \) is the standard deviation of the noise, and \( Q(x) \) is the \( Q \)-function and represents the tail probability of the standard normal distribution. Other than the white noise, there are other sources of noise that can degrade the overall SNR, such as device shot noise, supply, and substrate noise. These noise sources, unlike the white noise, are bounded in amplitude and usually scale with the signal amplitude as well as signal activity. As a result, the absolute BER cannot be solely related to the total noise power, as shown in Equation 2.2. Nevertheless, the SNR versus BER analysis serves as a useful tool. This method can be used to illustrate how an offset in the decision level can degrade performance. The offset can be considered as a reduction of the signal amplitude for one of the two binary levels. The sensitivity of an optical receiver is the minimum magnitude of OMA required to achieve a specified BER or signal-to-noise ratio (SNR). Sensitivity of an optical receiver is a crucial determining factor in overall link power, as it sets the laser power at the transmitter.

Besides amplitude noise, the second major contributor to BER is uncertainty in timing of the receiver. Like amplitude noise, this uncertainty is a random process, and it is characterised by the jitter of the receiver and transmitter clocks. Both sources of jitter shift the sampling point away from its optimum, and have the effect of reducing the amplitude margin and degrading the BER. This effect
is of particular concern as data rates increase, since jitter can become a substantial portion of a UI. As a result, timing margin is more concerning than voltage margin in high-speed links [19] and clocking high-speed optical links requires careful consideration.

2.1.2 Clocking

As the bit-time interval decreases, the receiver has a smaller timing margin and clocking becomes more difficult. In high-speed optical interconnects, understanding clocking schemes and designing a clocking structure that results in best signal integrity is of utmost importance. There are several common clocking schemes:

- **Synchronous**: In a synchronous link, the transmitter and receiver clocks are assumed to have the same frequency and phase. This is generally only a viable clocking scheme at low data rates.

- **Mesochronous**: In a mesochronous link, the transmitter and receiver clocks are assumed to have the same frequency, but may have different phases. A commonly used subset of this category is the source-synchronous link, where the clock is generated at the transmitter and forwarded along with the data.

- **Plesiochronous**: In a plesiochronous link, the transmitter and receiver clocks may have slight differences in frequency. The receiver is required to align its clock by extracting timing information from the incoming data stream.

- **Asynchronous**: an asynchronous link is not really clocked at all. Rather, it uses either control symbols inserted in the data stream itself or handshaking signals to transfer timing information.

The mesochronous/source-synchronous and plesiochronous styles are most routinely used for parallel electrical interconnect design. Since they require relatively straightforward timing recovery at the receiver (when compared to plesiochronous links), source-synchronous links are frequently used in computer systems, particularly where the link is composed of many data pins and the relative cost of adding a clock pin is small. For the same reason, source-synchronous clocking schemes are also attractive for highly parallel optical interconnects.

2.1.2.1 Jitter

Jitter is broadly defined as short-term variations of a signal with respect to its ideal position in time. As clock frequencies increase, IOs become more susceptible to deviations in clocks output transition from its ideal position. Excessive jitter can increase the BER of a communications signal. Therefore, accurate understanding of jitter and designing low-jitter clocking schemes are necessary for ensuring the reliability of a system. The two major components of jitter are random jitter and deterministic jitter.

- **Random Jitter**

Random jitter (RJ) is timing noise that cannot be anticipated because it has no predictable pattern
(or by definition it is random). Its randomness comes from inherent random noise in electronic and optical devices and typically exhibits a Gaussian distribution. This noise gets superimposed on the clock signal during its transition period and results in timing errors at the switching points. According to central limit theorem, RJ is Gaussian because it results from the composite effects of many uncorrelated noise sources. Given RJ’s Gaussian distribution, its instantaneous value is not mathematically bounded and therefore it is characterized by its standard deviation or root mean square (RMS) deviation from ideal clock timing.

- Deterministic Jitter

Deterministic jitter (DJ) is timing jitter that is repeatable and predictable. It is not intrinsic or random and has a specific source. DJ is effectively caused by imperfections in a device or transmission media, signal modulation, power supply noise, or cross-talk. DJ can be further sub-classified into periodic jitter and data-dependent jitter. The example of an interfering noise coming from a switching power supply is periodic because the noise will have the same frequency as the switching power supply. An example of data-dependent jitter is ISI due to isochronous coded serial data stream. Both types of DJ are linearly additive and always have a known source, i.e., they are correlated to (or caused by) something. This jitter component has a non-Gaussian probability density function and is always bounded in amplitude. Given its predictable nature, DJ is characterized by its bounded, peak-to-peak value.

2.1.2.2 Subrate Clocking

In a classical full-rate optical link, the period of the clock is the same as the length of a UI and, for example, a 25 Gb/s link will operate with a 25 GHz clock. At multi-Gb/s data rates, however, the high-frequency clocks required for this approach consume large amounts of power and complicate the process of timing recovery. Therefore, designers use sub-rate clocking, multiplexing/demultiplexing schemes, where the clock operates at some integer fraction of the data-rate and the data is transmitted and/or received using multiple clock phases. Although it is, in principle, possible to generate as many phases of the clock as desired and lower the clock rate arbitrarily, practical concerns (typically physical distribution of clock phases and layout considerations) limit optical link implementations to half-rate and quarter-rate; in a half-rate link, the positive (0) and negative (180) edges of the clock can be used directly, and it is fairly straightforward to generate in-phase (0) and quadrature (90) clocks and their negations (180 and 270) for quarter-rate systems.

2.1.3 Fundamental Integrated Photonic Building Blocks

Similar to micro electronics, miniaturization and more integration of more functions on a single chip radically drive down the cost of photonic devices and generally moving data in the optics domain. During the past decade, several integrated photonic devices have been introduced, making it possible
to design fundamentally new integrated photonic systems. Here we review those integrated photonic devices that are necessary to realize a high-speed optical link.

### 2.1.3.1 Passive Components and Waveguides

![Waveguide Figure](image)

**Figure 2.4:** (a) Strip waveguide. (b) Rib waveguide.

The most common on-chip interconnection scheme is through waveguides, physical structures that guide electromagnetic waves in optical spectrum. Depending on their mode structure, waveguides are classified into two general categories: single-mode and multi-mode. Modes are the possible solutions of the Helmholtz equation for waves, derived by applying boundary conditions to the Maxwell’s equations. These modes determine how the wave is distributed in space. A single-mode waveguide is one that allows the existence of only one mode while a multi-mode waveguide allows many. Modal dispersion is more severe in multi-mode waveguides due to multiple spatial modes. Therefore, single-mode waveguides tend to have higher bandwidth than multi-mode fibers and are often preferred in integrated photonics. On-chip waveguides can be categorized according to their geometry into strip or rib waveguides. A strip waveguide is a strip of the guiding material confined between cladding layers. In a rib waveguide, the guiding layer consists of strip on top of a slab (Fig. 2.4).

One of the major challenges in integrated photonics is cost-effective coupling of light between the chip and outside world. Coupling of light between on-chip waveguides and off-chip fibers is typically

![Coupling Techniques Figure](image)

**Figure 2.5:** Techniques for coupling light to a photonic integrated circuit: (a) Grating coupler, (b) Edge-coupled optical fiber to an on-chip waveguide. (c) V-groove coupling of optical fiber to an on-chip waveguide.
done using grating couplers, edge couplers, or V-grove coupling (Fig. 2.5). These approaches have demonstrated loss of less than 1dB per interface [20–22].

2.1.3.2 Modulators

Amplitude modulation of light can be achieved through plasma dispersion effect, FranzKeldysh effect, and quantum-confined Stark effect. Plasma dispersion effect is the one mechanism that is most widely used in integrated photonic modulators. It works based on changing the free carrier density of a guiding medium, which induces changes in refractive index and modulates the light. Several different mechanisms of manipulating free carrier density have been investigated. Among those, carrier-depletion-mode and carrier-injection mode are the most promising candidates for high-speed data-communication applications. Carrier-depletion devices are based on a reverse-biased pn junction and carrier-injection devices are based on a forward-biasd p-i-n diode. Carrier-depletion devices are widely used for high-speed operation and carrier-injection devices are used for low-voltage operation. Both mechanisms have demonstrated speeds up to 25Gb/s with hybrid-integrated CMOS drivers [23,24]. Modulators can be implemented through resonant or non-resonant structures. Mach-Zehnder interferometer (MZI) structures are typically as non-resonant structures used for amplitude modulation. With traveling-wave designs, data rates of more than 40 Gb/s have been demonstrated [25]. Resonant structures can be used to dramatically reduce area and power consumptions. This comes at the cost of a dramatically narrower operating wavelength and susceptibility to temperature fluctuations. High-speed ring modulators have been demonstrated to operate at up to 40 Gb/s [26].

2.1.3.3 Photodetectors

There are two types of commonly used devices for optical/electrical conversion: p-i-n diodes and metal-semiconductor-metal (MSM) diodes. In both of these devices an electric field pulls the carriers generated by the incident photons to the electrodes. The resulting current, i.e., photocurrent, is proportional to the number of photons absorbed per unit time. In a p-i-n diode, a reverse-bias across the diode ensures a strong electric field in the intrinsic region and negligible reverse bias current (dark current) in absence of light. Germanium and or SiGe are mostly suitable for integrated systems that use silicon as guiding medium. They can be both integrated close to or directly connected to a silicon waveguide, so that the guided light can be evanescently coupled or buttcoupled into the photodetector and the photodetector can have a small cross section to reduce device capacitance and improve speed.

Avalanche photodiode (APD) is another detecting device that is used to improve responsivity of photodiodes. APD is a photodetector that provides a built-in gain stage through avalanche multiplication. APDs used in high-speed optical links need to achieve high gain-bandwidth products without sacrificing noise or responsivity. Currently, separate absorption and charge multiplication
(SACM) structure has proven to be the most successful category for high-speed applications [27].
2.2 Opportunities In Silicon Photonics

Historically, photonic devices have been developed using different materials and in highly specialized fabrication facilities. These individual devices are separately packaged and connected using fibers. For example, in an ethernet switch rack it is common to see a system that is comprised of RF CMOS or Bipolar chips for high-bandwidth electronics, FPGA’s, or highly scaled CMOS chips for digital control, diffused waveguides on glass for optical multiplexers and passives, Germanium photodetectors, lithium niobate modulators, indium phosphide lasers, and MEMS-based optical switches. The processes used to make each of these devices are fundamentally different and incompatible with one another. As a result, often each device is made in a specialty fabrication facility with very low volume. Therefore, most photonic components are high-cost components compared with truly high-volume electronic chips. Additionally, a large fraction of the final system cost and yield loss is due to photonic packaging processes. Photonic packaging technologies require alignments with submicron accuracy and the packages are often hermetically sealed and sometimes gold-coated.

Nowadays, many computing and networking applications require a system-in-package with optical interconnects. Silicon photonic technology shows great assets for integrating multiple functions into a single package, and manufacturing most or all of them using the same fabrication facilities used to build microelectronics. This will significantly drive down the cost of integrated systems with a variety of components.

Silicon photonic foundries and processes are to a large extent the same foundries that were built to develop transistors. This seems to be counter-intuitive as the electronics industry has spent billions of dollars developing the most compact and fastest transistors without knowing that someday the same process will be used to generate, detect, modulate, and in general manipulate light. In reality, every effort to directly integrate photonic functionality in the same CMOS (or Bipolar) silicon wafer has yielded devices with poor performance. Even if the electronic processes could yield competitive photonic devices, they wouldn’t make economic sense; advanced microelectronic chips are now

Figure 2.6: (a) Packaged passive silicon-photonic chip (Courtesy Imec). (b) Silicon-photonic chip with active devices (Courtesy OpSIS). (c) Monolithic CMOS/silicon-photonic chip (Courtesy IBM).
at 14nm node whereas silicon photonic chips require primitive (or at most modest) lithography technologies like 90nm process. Over the last few years we have learned that silicon allows us to perform all of the key optical functions at a reasonably competitive performance level, with the exception of a laser. Particularly, photonic components required to establish an optical link have been successfully demonstrated. This includes silicon-based modulators, multiplexers and demultiplexers, Germanium photodetectors, and finally heterogeneous integration of III-V/Si lasers with silicon wafers. The silicon photonics community has developed process flows that permit re-use of CMOS fabrication infrastructure to build complex systems. In doing so, several organizations have shown the possibility of rearranging and reusing modular process steps to develop silicon photonic chips. Fig. 2.6 shows recent successful development of passive, active, and monolithic structures in silicon photonic platforms, using the existing CMOS foundry infrastructure.

It is particularly compelling that materials, techniques, and technologies developed over the past 50 years in microelectronic industry is being repurposed to build photonic integrated circuits. We can now leverage billions of dollars of investment that went into building our modern CMOS facilities to build system-in-packages comprised of complex microelectronic and photonic circuits. Hence, there is an immediate path to large-scale production and commercialization using existing infrastructure of CMOS fabs that use 200 and 300nm silicon on insulator (SOI) wafers.

Silicon photonics appears to be an industry at its infancy with similarities to CMOS industry in

Figure 2.7: Die micrograph of a complex SOI photonic chip fabricated at IME A*STAR (Courtesy IME).
the early 1970s. As shown in [28], the number of components in silicon photonic system design appears to double every 12-18 months. Photonic device designers appear to invest some of their effort into integrating those devices at larger scale and create new systems. The tools and support infrastructure of silicon photonic design appears to be developing at a quite remarkable pace. This is in part due to lessons learned from electronics industry, which painstakingly developed design and verification software and tools that were crucial for successful tape-out of large integrated circuits. Existing tools for RF and thermal simulation of electronics are routinely used for simulation of the same aspects of photonic devices. Development of design kits, PDKs, has almost separated designers from fabrication, resurfacing the semiconductor fabless industry. In the early ages of the electronics industry in 1970s, Lynn Conway at Xerox PARC and Carver Mead of Caltech developed the concept of multi-project wafer (MPW), where multiple designs shared the same wafer in a single manufacturing run [29]. These efforts consolidated in foundation of MOSIS, an organization that introduced public access to semiconductor manufacturing through MPWs. Revisiting MPW’s success, we have seen a number of emerging photonic silicon fabrication facilities offer MPW services, such as OpSIS, ePIXfab, IME, and CMC Microsystems. Complex silicon photonic systems are enabling a huge number of new applications beyond data communication. Some of the particularly promising applications are biosensing [30], LIDAR systems [31], optical gyroscopes [32], radio-frequency integrated optoelectronics [33], coherent communications [34], and laser noise reduction [35].
2.3 Hybrid Integration of Electronics and Photonics

Intimate integration of electronics and photonics can significantly improve system performance, reduce cost, and open up huge opportunities for transforming existing system and design of new systems. There are two general categories for integration: monolithic integration and hybrid integration. Monolithic electronic/photonic integration, such as Luxtera’s 130nm CMOS process [36], requires significant process modification, e.g., addition of silicon etching steps for waveguide and germanium for photodetectors. Additional layers have to be inserted such that existing CMOS transistors continue to function. Another approach is to use unmodified advanced technologies such as 45nm SOI CMOS process and integrate the required photonic devices within design rule limits [37]. Because of these constraints, operational wavelengths end up being near silicon absorption edge in silicon germanium. Accordingly, there have been limited reports of competitive high-speed optical communication systems that use such technologies.

Unlike monolithic integration approach, hybrid integration of electronics and photonics provides a twofold gain. First, the photonic technology can be independently optimized without need to sacrifice photonic device performance for CMOS design rules. Second, the CMOS technology continues to improve with Moore’s law, it is possible to quickly switch to newer generations of CMOS or Bi-CMOS and integrate them with inexpensive silicon photonics. At the end we end up with a state-of-the-art electronic/photonic system. In a high-speed optical interconnect, the choice of hybrid integration scheme can affect the overall footprint and density of the front-end, which is

![Figure 2.8: Hybrid integration schemes. (a) Wire-bonding. (b) Flip-chip bonding. (c) Through-silicon vias (TSV).](image)
particularly important for parallel interconnects. Silicon photonic chips can be hybrid-integrated to CMOS chips via a number of techniques, which are reviewed by Krishnamoorthy et al. in [38] and Miller et al. in [39]. The most mature hybridization techniques are wire bonding and flip-chip bonding. Wire-bonding has greater parasitic inductance and capacitance, reducing performance at high bit-rates as compared to flip-chip bonding. Hence flip-chip bonding is more suitable for high-performance applications requiring minimum front-end capacitance. Other 3D-integration schemes including through-silicon vias (TSV) [40] and low-capacitance copper pillar interconnects [41] offer even higher density and lower parasitics.

Similar to monolithic integration, electronic/photonic hybrid integration provides an opportunity for hollistic co-design of electronics photonics. This approach will open up new design opportunities for systems that were simply not feasible in a modular design. A hollistic design approach ensures performance metrics that have trade-offs tangled between electronics and photonics that are truly optimized.
Chapter 3

Optical Receivers

3.1 High-speed Optical Receiver Overview

Optical receivers are one of the most important building blocks of an optical interconnect. The impact of their well rounded design goes beyond the receiver’s characteristics and affects the required laser power and required OMA by transmitter. In this chapter we review prior art in optical receiver design and propose a design methodology for co-designed and co-integrated optical receivers for best overall link performance. In the next two sections we review some of the properties of photodetectors used in high-speed optical interconnects and follow with a review of conventional TIA design and its challenges.

Figure 3.1: Optical receivers are one of the primary elements of optical interconnects.
3.1.1 Photodetector Model

Fig. 3.2 shows a basic circuit model for a p-i-n photodiode. The photocurrent \((I_{op})\) is proportional to the input optical power \((P_{op})\). The diode responsivity \(R\) is defined as

\[ R = I_{op}/P_{op}. \] (3.1)

Responsivity of photodiodes depend on wavelength and can be expressed as \(^1\)

\[ R = q\eta\lambda/hc, \] (3.2)

where \(\eta\) is the diode quantum efficiency. A p-i-n diode that is designed optimally has \(\eta \approx 1\), which at 1550nm yields \(R = 1.2\text{A/W}\). Well-designed Germanium p-i-n detectors have responsivities close to this value. The noise current \(I_n\) in Fig. 3.2 is primarily due to shot noise. Background illuminations and thermal noise of the series resistor are other sources of noise. In most cases the photodiode noise is negligible compared with input referred noise of the receiver front-end circuitry. The photodiode capacitance combined with bonding parasitic capacitance is usually the dominant input load for a receiver and depending on receiver topology may impact the bandwidth, power consumption, and sensitivity.

\(^1\)\(hc/\lambda\) is the photon energy, where \(h\) is the Planck constant and \(c\) is the speed of light.
3.2 Prior Art in Design of Optical Receivers

3.2.1 Transimpedance Amplifiers

The trade-off between the bandwidth and gain of a front-end with a simple resistor makes it impractical for many applications. The effective input resistance of the front-end can be reduced significantly by adding an active gain stage to the design to make a transimpedance architecture. A transimpedance amplifier (TIA) is an analog front-end that has a high current-to-voltage gain with reduced input impedance. The addition of the active stage, with transistors or other active components, increases the noise. Nevertheless, with a careful design, very high SNRs are possible at the output of an optimized transimpedance amplifier. In this section, we briefly discuss the performance and trade-offs of a number of different TIAs. From electronics perspective, the most important specs of receiver front-ends are bandwidth, sensitivity, power consumption, area, and dynamic range. Some of the most commonly used TIA topologies are common-gate, common-source with shunt-shunt resistive feedback and regulated cascode TIA stage. Common-gate (CG) topology, shown in Fig. 3.3(a), is used to achieve a low input impedance and at the same time a high gain [42].
This topology isolates the photodiode and bonding capacitance from the gain resistor and therefore has a wide bandwidth. The effective input impedance is $1/g_m$ of the input transistor $M_1$, and the transimpedance is $R_D$. The direct additive noise of the resistor $R_D$ lowers the sensitivity. The equivalent input-referred current noise power spectral density is given by

$$i_n^2 = i_{nd}^2 + i_{nr}^2 = 4kT \left( \frac{1}{R_D} + \gamma g_m \right), \quad (3.3)$$

where $i_{nd}$ is the RMS of the current noise of $M_2$, $i_{nr}$ is the RMS of the current noise of $R_D$, $\gamma$ is the excess noise coefficient of $M_2$. The common-gate front-end relaxes the effect of large input capacitance on the receiver’s bandwidth. However, due to small $g_m$ of the transistors in more advanced technologies, they cannot totally isolate the parasitic capacitance. Also, a small $g_m$ exacerbates the noise and stability performance of the amplifier stage.

A regulated cascode (RGC) configuration (Fig. 3.3(a)) addresses these issues and is used in a variety of receiver designs [43,44]. The RGC input mechanism enhances the effective transconductance significantly; therefore higher bandwidths are feasible. Using small-signal analysis, the input resistance of the RGC front-end can be approximated as

$$R_{in} \approx \frac{1}{g_{m1}(1 + g_{m2}R_2)} \quad (3.4)$$

The input impedance is now smaller than the CG topology by a factor of $(1 + g_{m2}R_2)$. On the other hand, the feedback stage produces a second pole, resulting in a peaking in the frequency response. In order to avoid peaking and ensure stability, the gate width of $M_1$ or the resistance $R_2$ have to be reduced. Reducing $R_2$ decreases the input transconductance. In this case, in order to obtain the same $(1 + g_{m2}R_2)$, the drain bias current of $M_2$ needs to be increased, thus resulting in larger power consumption. Decreasing the width of transistor $M_1$ lowers the transconductance but not as much compared with reducing $R_2$. Nonetheless, it may lead to the increase of the channel thermal noise contribution from $M_1$ due to smaller $g_m$. The noise performance of RGC stage is thoroughly analyzed in [45]. Similar to the CG front-end the noise of bias resistor and gain resistor directly impact the total noise of the TIA. However, the enhanced input transconductance reduces the high-frequency noise contribution of transistor $M_1$ due to the large input parasitic capacitance [43].

The third TIA front-end topology that we will look at is a shunt-shunt feedback TIA. This alternative design relays the noise-headroom trade off in TIAs [46–48]. While the transimpedance gain is approximately $R_f$, the input impedance at DC is reduced to about $R_f/A$, where $A$ is the gain of the amplifier. In most designs the overall TIA bandwidth, BW, is limited by the pole at the input
node, not by the voltage amplifier, and is given by

\[ BW \approx \frac{A}{R_f(C_{in} + C_p)}, \]  

(3.5)

where \( C_{in} \) here is the input capacitance of the amplifier. The total input-referred noise of the TIA and the noise of the photodiode determine the sensitivity of the receiver. The TIA noise is mostly due to the thermal noise of the feedback resistor \( R_f \) and the input-referred noise of the amplifier. Optimization for lowest input-referred noise design yields \( C_{gs} \) to be \( \frac{1}{2} \) to 1 times \( C_p \) [49].

Three simple designs of common-source shunt-shunt feedback TIA are given in Fig. 3.3(c)-(e). In the first design, with enough open-loop gain the noise contributions from resistance \( R_1 \) and transistor \( M_2 \) can be very small and the noise of the amplifier is dominated by the input-referred noise of transistor \( M_1 \). In this design because of limited voltage headroom, the voltage drop across \( R_1 \) is limited to \( V_{dd} - (V_{gs1} + V_{gs2}) \), which results in higher noise, lower bandwidth, and reduced open-loop gain. In order to ensure stability in case the output node is highly capacitive, the load should be isolated by an extra source follower added in parallel to the existing one, driving the output capacitor. The second design can afford a larger voltage drop across \( R_1 \). In this topology the bandwidth is maximized when the system is under-damped. Hence, the pole at the gate of \( M_2 \) can be chosen to enhance the overall bandwidth by up to 40

\[ Z_i(s) = -\left( \frac{A}{A+1} R_f - \frac{r_0}{A+1} \right)(1 + s \frac{r_0 C_L + (R_f + r_0)C_{in}}{A+1} + s^2 \frac{R_f C_{in} r_0 C_L}{A+1})^{-1}, \]  

(3.6)

where \( A \) is the DC gain of the inverter stage at the operational bias, \( C_L \) is the load of the next stage, \( r_0 \) is the output resistance of the inverter stage, and \( C_{in} \) is the total input capacitance. Also the bandwidth of the stage can be written as

\[ BW \approx \frac{1 + A}{r_0 C_L + (R_f + r_0)C_{in}}. \]  

(3.7)

This expression is derived with the assumption that the two poles of the closed-loop transfer function are far apart. A complete analysis for the sensitivity of this topology can be found in [119]. In order to overcome the gain-bandwidth trade-off for this topology, multi-stage inverter-based TIAs [50]. This comes at the expense of higher power and added noise due to subsequent stages. Note that even in the case of a multi-stage TIA, the first stage is the most important contributor of input-referred noise and SNR.

### 3.2.2 Integrating Front-ends

In an integrating front-end, the input impedance of the receiver is designed to be purely capacitive within the frequency range of the input data. The photocurrent is integrated on the capacitor seen
Figure 3.4: Data resolution in integrating front-ends.

at the input node. This capacitor, $C_{tot}$ is the sum of the diode, bonding and front-end circuitry capacitors. If the average input current during a bit “zero” is $I_0$ and during a bit “one” is $I_1$, the voltage swing at the input node is going to be $\Delta V_0 = (I_0 \times T_b)/C_{tot}$ for a zero bit and $\Delta V_1 = (I_1 \times T_b)/C_{tot}$ for a one, $T_b$ here is the bit-time interval. This voltage swing is critical in determining the BER of the receiver and has to be large enough to have large enough SNR at the input of the comparator. The sensitivity directly depends on the size of the capacitor $C_{tot}$. If a single photodiode is simply connected to a capacitive node, $I_0$ and $I_1$ will be both positive or negative. If the signal has a high extinction ratios, $I_0$ will be small and close to zero. However, when receiving a stream of continuous data, the voltage at the capacitive input node will be saturated to high or low values. A simple solution for this problem is to use differential optical beams coming to two different photodiodes, creating a “totem pole”. Upon shining the light beams to each photodiode, one charges the input node, and the other one discharges it. If the incoming data is a “one”, the voltage of input node goes to high values and if it is “zero” the voltage goes to low values. If the input optical power is high enough to charge and discharge the input node all the way to $V_{dd}$ and $Gnd$ in less than a bit-time, then a simple inverter can recover a full swing voltage across the load $C_L$. This front-end, sometimes called “receiver-less front-end” [51], is very simple to design and consumes minimal electrical power. In this design, the required input OMA for a full voltage swing is $OMA = C_{tot}V_{dd}/RT_b$, where $R$ is the diode responsivity and $T_b$ is the bit-time interval. The OMA is proportional to $C_{tot}$, requiring very small photodiode and bonding capacitances and a small voltage buffer that follows it. If short pulses are not used and optical energy is uniform over one UI, for a short rise-time and fall-time at the input node, the OMA needs to be higher than the minimum value needed for full swing operation. In that case, the voltage of the input node may saturate to $V_{in} = V_{dd} + V_{int}$ or $V_{in} = -V_{int}$, where $V_{int}$ is the photodiode intrinsic or built-in voltage. A clamping diode is needed at the input node to limit the voltage swing and aviode stressing the input
transistor of the next stage. The receiver-less approach is particularly attractive for clock generation when low-jitter short-pulse lasers are available. Depending on the design of the detector, the rise times can be less than 10ps, which is faster than the FO4 delay of today’s CMOS technologies. Short rise-time and large signal swings reduce supply noise sensitivity, which makes this approach suitable for multi-GHz clock generation in noisy environments. A special case for integrating receivers is the double-sampling integrating front-end. In this front-end, as shown in Fig. 3.4, the voltage of the input node is sampled at bit transition moments. The two consecutive samples are compared using strongARM or sense-amp to resolve the incoming bit. If $V_n > V_{n-1}$, the decision will be a “1” and if $V_n < V_{n-1}$, the bit will be resolved as “0”. We will look at this front-end more carefully in the following sections.
3.3 3D-Integrated CMOS/SiP Optical Receiver With a LBW TIA Integrating Front-end

Advancement in heterogeneous integration technologies has been focused on higher density and lower parasitics, i.e., lower capacitance, resistance, and inductance. The motivation behind this trend is that the extra capacitor due to bonding added to photo-detectors (PD) capacitor creates the dominant pole (\( \omega_{p1} \)) of the TIA (Fig. 3.5). This is the limiting factor in the speed of conventional optical receivers based on TIAs. Also, in parallel optical receivers bond-wire can be a source of crosstalk between channels.

For a given integration technology, the input impedance of the TIA has to be reduced to push the dominant pole further. This is achieved by higher amplifier gain. However, higher gain comes at the cost of increased power consumption or by increasing the gain-bandwidth by area-consuming techniques such as inductive peaking. Integrating receivers have been introduced with the aim of removing these limitations [52]. In an integrating optical receiver based on double-sampling, the photocurrent is integrated on the front-end capacitor and is sampled at the beginning and end of the bit-time interval. A comparator, e.g., a strongARM sense amplifier, is used to distinguish ones and zeros by determining which sample is larger. Due to the integrating nature of the optical receiver, the pole associated with this front-end cap is not the bottleneck for achieving higher data-rates. In integrating front-ends with a resistive termination, the charge integrated on the front-end capacitor gets shared with sampling capacitors (Fig. 3.6), causing sensitivity degradation. This issue has been addressed by adding a low-bandwidth TIA to the front-end in this work [6], [7]. The low-bandwidth TIA provides isolation between PDs capacitor and sampling capacitors, which reduces charge-sharing effect and enables use of ultra-low capacitance PDs in advanced silicon photonic technologies.

In this chapter, we aim at developing a design guideline to choose among low-bandwidth TIA, resistive, and full-bandwidth TIA front-ends. In order to evaluate different receiver designs and
target a specific performance, various parameters can be considered. Power consumption, energy efficiency, sensitivity, and operational data-rate with a specific bit error-rate (BER) are some of these parameters. However, there are inherent trade-offs between these parameters and they cannot be independently enhanced. The question is how to assess performance of different designs and decide which one is superior. Different figures of merit \((FoM)\) have been proposed to quantify overall performance of optical receivers and compare them. For example, \([53]\) uses

\[
FoM_{sensitivity} = \frac{\text{sensitivity (Watts}_{pp}) \times \text{Responsivity}_{PD}}{C_{in} \times \text{DataRate}} \tag{3.8}
\]

to quantify design performance based on receivers sensitivity. In many other cases such as \([54]\), the following \(FoM\) is used:

\[
FoM_{TIA} = \frac{\text{Gain} \times \text{Bandwidth}}{\text{Power}} \tag{3.9}
\]

From an optical link system perspective, one of the most important metrics is the total power consumption, which includes the power of the laser at the transmitter side \([55]\). From a receiver circuit design perspective, the two main factors that determine the overall power consumption are sensitivity and electrical power consumption. The following equation defines an \(FoM\) to capture the impact of receiver on total power consumption of the optical link due to both electrical and optical components:

\[
FoM_{Power} = \frac{\text{Power}_{electrical} + K \times \text{Sensitivity (Current}_{pp})}{\text{DataRate}} \tag{3.10}
\]
where $K$ captures the laser efficiency, optical coupling losses, and responsivity of the photodiode (3.11):

$$K = \frac{\text{Efficiency}_{\text{laser}} \times \text{Couplinglosses}}{\text{Responsivity}_{\text{PD}}}.$$  (3.11)

Note that while the electrical power of transmitter circuitry is not included in the FoM presented in 3.10, the laser power, modulators insertion loss, and its coupling losses are taken into account. This is because the receiver sensitivity determines the laser power and the coupling losses are proportional to initial laser power. It is instructive to plot 3.10 for some recently reported optical receivers based on different receiver architectures [6, 7, 53–64] (Fig. 3.7). The portion of power consumption associated with laser (blue) and electronics (red) is separated. K is calculated for a responsivity of 0.8A/W, laser wall-plug efficiency of 15%, and coupling loss of 3dB for a transmitter architecture based on CW laser and MZI-based modulator. As can be seen, the laser power is important and often the dominating factor. Advancement in integrated photonics, CMOS scaling, and packaging technology such as the 3D-integrated optical receiver presented in [65] help reduce the overall power consumption of an optical link. However, the choice of circuit topology and its sensitivity can play an important role in optimal overall link design and reduction of total power consumption. In light of the discussion above, we present a compact 3D integrated optical receiver, which is designed specifically to take advantage of advanced silicon-photonics and low-parasitic integration technology to achieve high sensitivity and low power consumption. The low-bandwidth TIA front-end enhances sensitivity of the double-sampling receiver architecture and enables realization of a high-sensitivity optical receiver operating at 25Gb/s.

Figure 3.7: Recently published optical receivers.
3.3.1 3D-Integration Platform

The silicon photonic chip in this receiver is designed in Letis advanced silicon photonic platform and comprises a 2D surface grating coupler, which has two outputs coupled to a waveguide photodetector. This technology enables low parasitic capacitance and dense interconnections of CMOS and Silicon Photonics. Fig. 3.8 shows the layout and die micrograph of the silicon photonic chip. The chip is designed such that electrical connections to CMOS chip are routed through silicon photonic die. Fig. 3.9 shows PD capacitance as a function of reverse bias voltage. For a reverse bias of

![PD Capacitance](image)

Figure 3.9: Measured photo-detector capacitance as a function of bias voltage.
-0.5V to -2V, the PD capacitance is measured to be less than 8fF. The -3dB bandwidth of the PD is measured to be larger than 18GHz when terminated with a 50Ω resistor. The 3D-integration involves Under Bump Metallization (UBM) of the silicon photonic chip, and growth of copper micro-pillars on the Electronic wafer. The EIC was then flip-chip mounted on the silicon photonic IC. Fig. 3.8(b) shows the die micrograph of the Silicon photonic chip with UBM processed on the photonic wafer. Fig. 3.10(a) shows the scanning electron microscope (SEM) top view of the CuP grown on the electronic wafer and Fig. 3.10(b) shows the cross section of the micro-pillar. The minimum pitch for adjacent copper pillar bonds is 40µm, allowing realization of dense optical links [41]. The parasitic capacitance and resistance associated with each bonding is measured to be less than 25fF and 1Ω. The parasitic inductance is negligible. This low-parasitic integration is one of the key elements for achieving high bandwidth and high sensitivity optical links.

3.3.2 Receiver Architecture Overview

Fig. 3.11 shows the proposed integrating receiver and a conventional TIA-based receiver designed for comparison. In conventional receivers, TIAs are used to reduce the input impedance and increase the bandwidth. In this work a scalable 3-stage TIA, based on inverters with a resistive feedback, is used. This architecture is particularly suitable for highly scaled CMOS technologies and consumes relatively low power and area. The proposed integrating receiver uses a low-power, low-bandwidth TIA as front-end with a bandwidth that is much lower than the bit-time interval. Note that the integrating nature of the receiver comes from the sampling node (output of the low-bandwidth TIA).

Figure 3.10: (a) Scanning electron microscope top view of the CuP grown on the electronic wafer. (b) Cross section of the micro-pillar.
At this node, the impedance of the sampling capacitors at frequency of operation is much lower than the output impedance of the low-bandwidth TIA. Therefore, most of the charge is integrated on the sampling capacitors. The low bandwidth TIAs output is sampled at the beginning and end of the bit time ($T_b$). These samples, ($V[n]$, $V[n+1]$) are then compared to resolve each bit ($\Delta V[n] < 0$ results in “1” and $\Delta V[n] > 0$ results in “0”). Note that the voltage difference, $\Delta V[n]$, is input pattern dependent due to low-bandwidth nature of the front-end. For example, a “1” followed by a long sequence of “0” generates a stronger $\Delta V[n]$ compared to a “1” followed by many “1”s. Dynamic
offset modulation (DOM) is utilized to provide a constant voltage at sense-amps input irrespective of the stream of data [66]. DOM essentially increases the voltage difference for weak ones/zeros and decreases it for strong ones/zeros. The underlying principle used for DOM is that identical consecutive bits shift the sampling node voltage away from its DC average. So, the introduced offset is proportional to the value of the voltage at the sampling node compared with its DC average. The DOM will be investigated using z-domain analysis later in this section.

Double sampling technique allows de-multiplexing by using multiple clock phases and samplers. The poles associated with input and output nodes of the LBW TIA are as follows:

\[
\omega_{p,\text{in}} = \frac{1 + A}{R_f(C_{in} + C_{PD})} \tag{3.12}
\]

\[
\omega_{p,\text{out}} = \frac{1}{4R_fC_{S/H}}, \tag{3.13}
\]

where \(R_f\) is the feedback resistance, \(C_{in}\) is the capacitance looking into the TIA, and \(C_{PD}\) is the photodiode and parasitic capacitance combined. Fig. 3.12 shows the simple model for the receiver in the z-domain. The voltage of the sampler can be written in the z-domain as

\[
V(z) = \left(\frac{R_fI_{PD}}{1 + 1/A}\right)\left[\frac{1}{1 - e^{-\frac{T_b}{\tau_{p1}}}} z^{-1} + \frac{1}{1 - e^{-\frac{T_b}{\tau_{p2}}}} z^{-1}\right], \tag{3.14}
\]

where \(T_b\) is the bit time interval and \(\tau_{p1}\) and \(\tau_{p2}\) are time constants associated with low-bandwidth TIA poles. By subtracting the previous sample, \(V[n-1]\), from \(V[n]\) the result, \(\Delta V\), can be written in the z-domain as

\[
\Delta V'(z) = V(z)(1 - z^{-1}) - \beta z^{-1}V(z), \tag{3.15}
\]

where \(\beta\) is the DOM coefficient and has to be chosen such that \(\Delta V'\) becomes independent of \(z\). \(\beta\) can be found only if the low-bandwidth TIA is designed with its output pole at much lower frequency compared to \(1/T_b\) and the input node is at a higher frequency compared to \(1/T_b\). In that case by choosing \(\beta\) as

\[
\beta = 1 - \frac{T_b}{\tau_{p1}}, \tag{3.16}
\]

the DOM provides a constant voltage at sense-amps input regardless of the data sequence. This constant voltage is given by the following equation:

\[
\Delta V' \approx \frac{R_fI_{PD}(1 - e^{-T_b/\tau_{p1}})}{2(1 + 1/A)}. \tag{3.17}
\]
3.3.3 Circuit Implementation

Fig. 3.13 shows details of circuit level implementation for the optical receivers building blocks. The high-speed sense-amp has digital offset cancellation using a bank of five NMOS capacitors in accumulation mode. The sense-amp is followed by an SR-latch to retrieve the NRZ data. A CMOS quadrature divider is used to generate the four phases required for operation of the optical receiver.
Sampling capacitors are followed by an amplifier with a gain of 4.5dB, which also provides isolation between sampling nodes and sense-amp to minimize kickback. Dynamic offset modulation that is employed at the output of the amplifier is also implemented as another differential pair, sharing the same load. The photodiode emulator is a high-speed open-drain PMOS pair that steers a pre-set current between replica low-bandwidth TIA and the optical receiver under test. A separate current mirror is used to set the extinction ratio. The low-bandwidth TIA has a transimpedance of 3kΩ and has a digitally controlled 7-bit current DAC at its input to set the DC point at the input of the low-bandwidth TIA. The S/H capacitor is chosen to be 12fF to minimize noise and current sensitivity while achieve 25Gb/s operational data-rate.

Two prototypes were fabricated in a 28nm CMOS technology to compare performance of the integrating architecture with a conventional TIA-based receiver. Receivers occupy an active area of 0.0018mm². The 3-stage TIA architecture design is optimized to have maximum bandwidth for the given technology (28nm CMOS). Fig. 3.14(a)-(d) show the 3D integration of electronics/photonics as well as the top view of CMOS and SiPh chips. Each prototype is composed of two receivers, one with a photodiode emulator (Fig. 3.13) and one for optical testing with 3D-integrated silicon photonics.

### 3.3.4 Experimental Results

Fig 3.15 shows the optical measurement setup used for testing the chips. Initial verifications were done using the on-chip emulator, which mimics the photodiode current with an on-chip switchable current source and a bank of capacitors, to emulate the parasitic capacitances due to PD and bonding. An on-chip CML-to-CMOS converter generates the full swing clocks from an off-chip clock source and the four phases of clock are generated using an on-chip quadrature divider. The on-chip clock was measured to have about 8-ps peak-to-peak jitter. The functionality of the receiver was first validated using the on-chip emulator and PRBS-7, PRBS-9, and PRBS-15 sequences. \( R_f \) and \( C_{in} \) were chosen to be 3kΩ and 50fF, respectively (\( 4R_fC_{S/H} \approx 150\text{ps} \)). Functionality of the DOM for long sequences of ones or zeros was validated using a 100 MHz square-wave current applied to the input to the receiver while the front-end sampled the input at 25 Gb/s. In this case, 250 consecutive zeros will be followed by 250 ones. For an input time constant of about 0.105ns, these 250 consecutive bits push the input to the saturation limits. Fig. 3.16 shows the optical input eye diagram and output recovered and de-multiplexed eye diagram for both receivers at their maximum speed. The silicon photonic chip uses a grating coupler to couple light from an off-chip source. The capacitance due to the CuP bonds and pad is estimated to be less than 25fF and the photodiode capacitance is measured to be less than 8fF. The optical beam from a 1550nm DFB laser diode is modulated by a high-speed Mach-Zehnder modulator and coupled to the photodiode through a single-mode fiber. The PD responsivity including grating coupler losses was measured to be 0.2A/W.
The receivers were tested using a PRBS-15 sequence. The maximum achievable data-rates for the integrating receiver and conventional receiver are measured to be 25Gb/s and 21.2Gb/s, respectively.
Figure 3.16: (a) Input optical eye diagram at 25Gb/s for integrating receiver. (b) Output recovered and de-multiplexed eye diagram at 6.25Gb/s for the integrating receiver. (c) Input optical eye diagram at 21.2Gb/s for the conventional receiver. (d) Output recovered and de-multiplexed eye diagram at 5.3Gb/s for the conventional receiver.

Fig. 3.17(a) shows measured sensitivity vs. data-rate for both receivers, bonded with wire-bond and CuP integration. For bit-error rate (BER) of $10^{-12}$, the conventional receiver requires -10.4dBm of optical modulated amplitude (OMA) at 21.2Gb/s (its maximum speed), while the integrating architecture requires OMA of -16.1dBm at 21.2Gb/s and -14.9dBm at 25Gb/s. The coupling loss, measured to be 6dB, is included in these measurements. Fig. 3.17(b) shows power consumption and energy efficiency of the receivers at different data rates. In both designs, the power of digital elements increases linearly with speed. The 3-stage TIA design offers per-bit energy consumption of 226fJ/b at 21.2Gb/s compared with the integrating architecture that has per-bit energy consumption of 171fJ/b at 21.2Gb/s. Energy efficiency of the integrating design reaches its peak of 170fJ/b at 25Gb/s. Fig. 3.17(c) shows power consumption break-down for both receivers. Fig. 3.18 shows the sensitivity curves of the integrating optical receiver at different data-rates. The coupling loss...
Figure 3.17: Optical receiver measurement results. (a) Sensitivity of integrating and conventional receivers bonded with CuP and wire-bond. (b) Energy efficiency and power consumption of the optical receiver. (c) Power consumption breakdown for integrating and conventional receiver.

Figure 3.18: Sensitivity curves of the integrating optical receiver. (measured to be 6dB) is included in this plot. Fig. 3.19 shows bathtub curves of receiver designs at their respective maximum operational speeds. Table 3.1 summarizes the performance of the designed optical receivers.
3.3.5 Performance Analysis And Design Guidelines

Input referred noise is the critical parameter determining the sensitivity of an optical receiver. Multiplying it by the signal to noise ratio (SNR), which is calculated for a target BER, yields receivers current sensitivity. The gain of the TIAs first stage has a critical role as noise of all subsequent stages gets divided by this gain. When designing an inverter-based TIA for a given CMOS technology, there is a maximum gain-bandwidth product a single stage can achieve. Thus, the gain has to be reduced to achieve sufficient bandwidth at high data-rates. Lower gain at the first
stage worsens the sensitivity of the receiver. Fig. 3.20 shows noise sources for integrating receiver and the conventional TIA-based receiver. The minimum required current for the integrating receiver operation is

\[ I_{b,\text{integrating}} = SNR \times \sigma_{n,\text{prop}} + \frac{1}{R_{LBW,TIA}} \times \frac{V_{\text{offset}}}{A_B}, \]

where \( \sigma_{n,\text{prop}} \) is the total input referred noise for the integrating receiver, \( V_{\text{offset}} \) is the residual sense-amp offset after calibration, \( A_B \) is the buffer gain, and \( R_{LBW,TIA} \) is the low-bandwidth TIA transimpedance. In the case of conventional TIA-based receiver, we have

\[ I_{b,\text{conventional}} = SNR \times \sigma_{n,\text{conv}} + \frac{1}{R_{TIA}} \times \frac{V_{\text{offset}}}{A_{\text{AMP}}}, \]

where \( \sigma_{n,\text{conv}} \) is the total input referred noise for the conventional receiver, \( A_{\text{AMP}} \) is the amplifier gain, and \( R_{TIA} \) is the TIA transimpedance. Referring all noise sources to input, the total input

![Figure 3.20: Noise sources for (a) integrating optical receiver and (b) conventional TIA-based optical receiver.](image-url)
referred noise can be written as

$$\sigma^2_{n,prop} = \frac{1}{R^2_{LBW,TIA}} \times (\sigma^2_{S/H} + \sigma^2_B + \frac{\sigma^2_{SA}}{A^2} + \sigma^2_{J} + \sigma^2_{LBW,TIA}).$$

(3.20)

Similarly, the input referred noise of the conventional TIA-based receiver could be written as

$$\sigma^2_{n,conv} = \frac{1}{R^2_{TIA}} \times \left( \sigma^2_{AMP} + \frac{\sigma^2_{SA}}{A^2} + \sigma^2_{TIA} \right).$$

(3.21)

For a given power consumption, the gain-bandwidth product of the TIA stage is constant. Therefore, by choosing larger $R_{LBW,TIA}$ compared with $R_{TIA}$ the overall sensitivity improves. The sense-amp noise contribution could be modeled as a sampler with gain, which has an input referred noise of [41]

$$\sigma^2_{SA} = \frac{1}{R^2_{TIA}} \times \frac{2kT}{A^2_{SA}C_A},$$

(3.22)

where $C_A$ is the sense-amp decision node capacitance, and $A_{SA}$ is the sense-amp gain. This capacitance is set to be about 15fF to cover the expected offset range. The sense-amp has a gain of 1.1, which results in a sense-amp input referred noise of about $(0.2\mu A)_{rms}$. The buffer noise is calculated as

$$\sigma^2_B = \frac{1}{R^2_{TIA}} \times \frac{8kT}{g_m} (\gamma + \frac{1}{R_B}),$$

(3.23)

where $\gamma$ is the transistor noise coefficient, $g_m$ is the transconductance of the stage, and $R_B$ is the load resistance. This noise contribution of this buffer stage with 4.5dB of gain is simulated to be around $(0.22\mu A)_{rms}$. The sampling capacitors noise contribution is equal to

$$\sigma^2_{S/H} = \frac{1}{R^2_{TIA}} \times \frac{2kT}{C_{S/H}}.$$  

(3.24)

The factor of two is due to the fact that we have two differential sampling capacitors connected to the buffer. Clock jitter is also an important factor when calculating the receiver sensitivity. Deviations from ideal sampling time translate to voltage level uncertainties in the sampling voltages that could be modeled as a noise source with variance of

$$\sigma^2_J = \frac{1}{R^2_{TIA}} \times \left( \frac{\sigma_{CLK}}{T_b} \right)^2 \Delta V^2_b,$$

(3.25)

where $\sigma_{CLK}$ is the clock RMS jitter [18]. Given the measured clock RMS jitter of 0.9ps, $\sigma_J$ is calculated to be around $(0.17\mu A)_{rms}$. The total noise due to dynamic offset modulation is measured to be

$$\sigma^2_{DOM} = \frac{1}{R^2_{TIA}} \times \frac{\beta^2}{2A^2} (\sigma^2_J + \sigma^2_S + 2\sigma^2_B).$$

(3.26)
This contribution is negligible as the dynamic offset modulation coefficient, $\beta$, is much smaller than buffer gain, $A$. Finally, the low-bandwidth TIA noise is simulated to be $(0.18 \mu A)_{rms}$. It is instructive to plot $I_b$ as a function of operational data-rate. To do so, the required bandwidth in GHz is set to be 0.7 times data-rate in GHz and the gain and bandwidth are simulated for each data-rate for the given technology, 28nm bulk CMOS. The resulting sensitivity versus data-rate trade-off is shown in Fig. 3.21. Note that the crossing point where conventional TIA-based receiver achieves better sensitivity is 8Gb/s. Above this data-rate, the gain of the TIA stage has to be reduced to achieve the required bandwidth and 3.25 becomes larger than 3.26. At low data-rates, the additional terms present in 3.25 make the current sensitivity of the integrating receiver larger. Another design consideration in integrating receivers is the choice between resistive and low-bandwidth front-ends. To investigate this, we simulated sensitivity of RC and low-bandwidth TIA front-ends at 25Gb/s using different PD capacitances. A ratio of $\times 10$ between the photo-detector capacitor and the sampling capacitor is assumed to avoid excessive charge sharing. The minimum controllable capacitance for sample-and-hold is around 3.5fF, limited by next stages parasitic capacitance. Fig. 3.22(a) shows simulation of sensitivity vs. photodiode capacitance. As can be seen, for a PD capacitance below 135fF, the low-bandwidth TIA front-end achieves better sensitivity. Note that this is true even for a different data-rate, as in both LBW TIA and RC front-end cases the sensitivity is dominated by integration time and linearly increases with data-rate. In order to account the power overhead consumed by the low-bandwidth TIA, the FoM defined in (3) is plotted for a data-rate of 25Gb/s in Fig. 3.22(b). In this case, the low-bandwidth TIA front-end becomes superior for a PD capacitance below 115fF. If we run this simulation for higher data-rates TIA front-end becomes superior for a higher PD capacitance. That is because the power consumption of digital elements and sensitivity linearly scales with data-rate but the power consumption of LBW TIA remains relatively unchanged.

Figure 3.21: Simulated sensitivity vs. speed for integrating and conventional optical receivers.
Figure 3.22: (a) Simulated sensitivity vs. PD capacitor for RC and low-bandwidth TIA front-end at 25Gb/s. (b) FoM vs. PD capacitor for RC and low-bandwidth TIA front-end at 25Gb/s.
3.4 An Adaptive 3D-Integrated CMOS/SiP Parallel Optical Receiver with Injection-Locked QuarterRate Clocking

Optics allows very high density interconnection through wavelength division multiplexing (WDM). Recent developments of high-density hybrid-integration combined with dense electronics and silicon photonics pave the way for massive parallel optical interconnects. Here we present a source-synchronous 4-channel parallel optical receiver that leverages the receiver architecture presented in the previous section. Copper pillars are used to hybrid-integrate the electronics chip with a silicon photonic chip that comprises a dense array of detectors, WDM and two grating couplers for optical input data, and the forwarded clock.

A quadrature-locked loop (QLL) based clocking is used to generate accurate clock phases for a 4-channel optical receiver using a forwarded clock at quarter-rate. QLL is a frequency tracking technique used to increase the locking range of the ring based quadrature injection locked oscillator. This technique is used to generate the accurate quadrature phase from a single phase of electrical/optical clock without any frequency division. In this receiver architecture, the QLL drives an injection-locked oscillator (ILO) at each channel, without any repeaters for local quadrature clock generation, ensuring low power clocking. Each local ILO has deskew capability for phase alignment. The wide locking range of the QLL ensures reliable operation across wide data rates. The clocking of the optical receiver is designed by my fellow lab-mate, Mayank Raj.

3.4.1 Body-biasing

Because of building blocks with bias currents, energy per-bit of optical receiver degrades at lower data-rates. In fully depleted silicon-on-insulator (FD SOI) CMOS process body biasing (BB) effect is significantly enhanced compared to bulk CMOS. In this process, the channel forms in an ultrathin (7nm) layer of intrinsic silicon over a layer of buried oxide (BOX). Given how thin the buried oxide layer (25nm) is and the presence of the conducting layer under the BOX, the effect of body biasing (BB) is improved. Therefore, $V_t$ of the transistors can be tuned around 80-150mV per 1V modulation of $V_{BB}$ (transistor’s body bias voltage), depending on device type (Fig. 3.23). We present an adaptive body biasing circuit that is designed to maintain the per-bit energy consumption of the receiver across wide data-rates.

3.4.2 Quadrature-locked loop

The novel frequency tracking method used in this receiver exploits the dynamics of injection locking in a quadrature ring oscillator to increase the effective locking range and produce accurate quadrature phases. When a ring oscillator with natural frequency $f_0$ is injected with an external signal with
Figure 3.23: (a) FD SOI MOS structure. (b) Threshold voltage ($V_{th}$) variation with back bias ($V_b$).

Figure 3.24: Quadrature phase error of a two-stage ILO in the unlocked case (a) close to lock and (b) far from lock. (c) Mean quadrature phase error vs $f_0$ (d) Effect of injection strength on MQPE.

frequency $f_{inj}$, the outputs of the ring oscillator incur a phase mismatch error if $f_0$ is not equal to $f_{inj}$ [67]. It is shown in [12] that the mean quadrature phase error (MQPE) contains information about the difference between the natural frequency of the oscillator and injected frequency, i.e., $|f_{inj} - f_0|$, in both locked and unlocked states. In locked state, where $f(t) = f_{inj}$, this relationship
is expressed as

\[ MPQE = \frac{\pi}{2} \left( \frac{f_{inj}}{f_0} - 1 \right) = \frac{\pi}{2} \left( \frac{\omega_{inj}}{\omega_0} - 1 \right) \].

(3.27)

Variation of instantaneous frequency of the oscillator in unlocked state yields the quadrature phase error in unlocked state. Differentiating the transient response of an ILO, we get

\[ f = f_{inj} + \frac{f_b^2}{f_0 - f_{inj}} \times \frac{\sec(\pi f_b t)^2}{1 + \left( \frac{f_{inj}}{f_0 - f_{inj}} + \frac{f_b}{f_0 - f_{inj}} \tan(\pi f_b t) \right)^2}, \]

(3.28)

where \( f_l \) is the locking range and \( f_b = \sqrt{(f_0 - f_{inj})^2 - f_l^2} \). The quadrature phase error exhibit beats with frequency a \( f_b \) as shown in Fig. 3.24(a),(b). This periodicity allows us to calculate the MQPE in the unlocked state by integrating the quadrature error from 0 to \( 1/f_b \). Fig. 3.24(c) shows the variation of MQPE with change in \( f_0 \) for a fixed \( f_{inj} \) of 7GHz and injection strength (k) of 0.05. It has two regions, locked and unlocked. As expected, the MQPE is 0 for \( f_{inj} = f_0 \). In the locked state the MQPE increases (almost linearly) as \( |f_{inj} - f_0| \) increases. MQPE goes to zero asymptotically as \( |f_{inj} - f_0| \) increases in the unlocked state. This suggests that the MQPE is a measure of the sign of \( f_{inj} - f_0 \) in both locked and unlocked states, and the quadrature phase error detector can be used as a phase frequency detector (PFD) in an injection locking environment. Therefore, the quadrature error can indeed be used in a feedback system to set the natural frequency \( (f_0) \) of the oscillator such that \( f_0 = f_{inj} \), thereby boosting the effective locking range. As a feature of this technique, the MQPE can be controlled by changing the injection strength. As shown in Fig. 3.24(d), increasing the injection strength (K) increases the intrinsic locking range of the injection locked oscillator, widening the linear region. This feature is useful, as injection strength can be controlled externally and used as means of controlling the MQPE. Fig. 3.25 shows how the quadrature-locked loop concept can be
used to implement an injection locked two stage differential ring oscillator. Instantaneous quadrature error is measured by using a phase detector (PD), which takes the $I$ and $Q$ phases of the clock from an ILO as inputs. The error is averaged using a charge pump and a loop filter, and fed back to the oscillators $V_{ctrl}$. The loop tracks the changes in the injected frequency and natural frequency of the oscillator until their difference $f_{inj} = f_0$ is minimized, assuring a wide locking range.

### 3.4.3 Receiver Architecture Overview

Figure 3.26 shows the top-level architecture of the adaptive receiver (single channel) with dynamic BB using $V_{ctrl}$ of the QLL. The first stage of the receiver is a low-power TIA with $3k\Omega$ feedback resistor. The receiver architecture is similar to that of single-channel design described in section 3.3.2. Similarly a de-multiplexing factor of four is used immediately after the TIA using quarter-rate clocked samplers. The quarter-rate architecture of the receiver, necessitates accurate quadrature clocks. In addition, due to the multiple channels there is a need for per channel deskewing to align the clock to the data. We explain the details of accurate quadrature phase generation and deskewing in next sections. The clocking structure is shown in Fig. 3.27. The optical receiver has four optical data inputs and one forwarded clock (electrical/optical) input. The optical clock is converted to an electrical clock using a TIA. The electrical clock is then sent to a global QLL circuit. The QLL generates four quadrature phases. The four phases are distributed without any repeaters and sent to local ring oscillators, which are placed near the clocked optical receivers. The local ring oscillators are injection locked to the global clock and frequency of oscillation is varied to control the phase of the local ring oscillators output (deskew). The data receivers have a quarter rate architecture and

![Figure 3.26: Architecture of a single channel of the optical receiver (single channel) and body-biasing circuit](image)
hence require accurate quadrature phases. Symmetric injection with four clock phases ensure that quadrature accuracy is maintained even with deskew. The QLL and clocking circuitry were designed by my fellow labmate, Mayank Raj.

![Clock Distribution Architecture](image)

Figure 3.27: Receiver clock distribution architecture.

3.4.4 Circuit Implementation

The optical receiver has four optical data inputs and one forwarded clock (electrical/optical) input. The optical clock is converted to an electrical clock using a simple TIA that has relaxed sensitivity requirements (as it is carrying a single tone rather than data). The prototype is implemented such that the reference clock can be injected both electrically and optically. The TIAs output voltage amplitude (150 mV) is sufficient for the IL architecture because of its high voltage gain. The electrical clock is then sent to a global QLL circuit. The QLL generates four quadrature phases. The four phases are distributed without any repeaters and sent to local ring oscillators, which are placed near the clocked optical receivers. The local ring oscillators are injection locked to the global clock and frequency of oscillation is varied to control the phase of the local ring oscillators output (deskew). The data receivers have a quarter rate architecture and hence require accurate quadrature phases. Symmetric injection with four clock phases ensure that quadrature accuracy is maintained in presence of deskew. Each local injection-locked oscillator consists of a V-to-I converter and a two-stage, cross-coupled, pseudo-differential current-starved ring oscillator. A two-stage ring oscillator is chosen to minimize power consumption. Fig 3.28 shows the circuit implementation of the QLL.

Circuit implementation of the receiver’s building blocks is similar to the implementation of the
single-channel receiver presented in 3.3.3. Most optical receivers have analog building blocks with bias currents. These are biased to provide the maximum bandwidth and gain for operation at the highest data rates, thus consuming maximum power. For operation at lower data rates a high bandwidth is not required. However, since the bandwidth of the analog components do not change with data-rates, power is wasted at lower data-rates. This leads to degradation of the power efficiency (the energy per-bit) of the optical receiver at lower data-rates. It is advantageous to bias the circuits adaptively so as to reduce the bias current (and hence power) of the analog components at lower data rates. This requires information about the data rate and a method to use this information to change the bias currents of the analog components. The former is provided by the QLL as it generates the $V_{ctrl}$ (Fig. 3.29), which is dependent on the input clock frequency, and hence the data rate. The latter is achieved by taking advantage of the FD SOI technology as described in section 3.4.1. The $V_{ctrl}$ generated by the QLL follows the ring oscillators characteristics as shown in Fig. 3.29, i.e., as the reference frequency increases the $V_{ctrl}$ decreases from 1 to 0. The body bias generator is designed so that the transfer function from $V_{ctrl}$ of the QLL to VBB generator outputs is such that receivers building blocks optimally work at any given data-rate. By fitting the transfer function of the body bias generator from $V_{ctrl}$ of QLL to body bias of respective blocks, the gain-bandwidth product of the TIA and Amps gain are adaptively set to be proportional to the data-rate. This task is achieved by the circuit shown in Fig. 3.30. This circuit is a simple level shifter that is designed to fit the transfer function of $V_{ctrl}$ to $V_{BB}$ required for the optimal gain-bandwidth product of analog blocks.
3.4.5 Measurement Results

The test chip is fabricated in a 28nm FD SOI CMOS process. The die micrograph and core detail are presented in Fig. 3.34. The core area is 300µm×60µm, in a 5mm×1.1mm die. The top metal layers are designed to be compatible with copper-pillar flip-chip bonding as well as bond-wire. The clock output from the QLL is symmetrically distributed to all four local ILOs with a total trace length 260m (Fig. 3.34).
3.4.5.1 QLL Measurements

To demonstrate the increase in locking range we disable the loop and set the Vctrl (Fig. 1) of the ILO at VDD/2. Without the quadrature phase error tracking, a locking range of 77.4GHz is observed at an injection strength (K) of 0.05. With the loop activated the locking range improves to 411GHz. The quadrature correction loop needs to run slower than the injection locked loop to assure stability. Under such conditions, the effective bandwidth of the system (when locked) is dictated by the injection locking process. To demonstrate this property, jitter transfer function (JTF) of the system is measured (Fig. 3.32). The JTF has a Lowpass characteristic with a bandwidth of 250MHz and a 20dB/dec decay, suggestive of a first-order system. Ring oscillators are susceptible to power supply variations. Injection locking helps in suppressing low frequency VDD noise, as shown in Fig. 3.32. The measurement is made by adding sinusoidal noise on the VDD and then measuring the relative frequency sidebands on the output in unlocked and locked cases. Integrated output jitter (100KHz-GHz) of 558fs and 577fs are measured at 8GHz (32Gb/s operation) for electrical and optical inputs, respectively. At the highest locking frequency (11GHz) the integrated output jitter is 642fs. Quadrature error measurements are done by directly measuring deviation of I and Q phases (Fig. 3.33), and an average of 1.5° was recorded across the locking range. The QLL consumes 2.77mW at 11GHz.

3.4.5.2 Optical Receiver Measurements

The optical test setup is shown in Fig. 3.31. For optical testing, the receiver is bonded to a photodiode with responsivity of 0.9A/W (Fig. 3.34). The total capacitance at the input node was estimated to be 120fF. The optical beam from a 1550nm distributed feedback (DFB) laser is modulated by a high speed Mach-Zender modulator (MZM) and coupled to the photodiode with a single-mode fiber. The optical fiber is placed close to the photodiode aperture using a micro-positioner. As the beam
has a Gaussian profile, the gap between the fiber tip and the photodetector causes optical intensity loss. Combined optical loss due to the optical coupling and optical connector is measured to be 2.8dB. Quarter-rate clock generated by the pattern generator was used as (electrical) reference for the QLL.

The functionality of the receiver is validated using the PRBS-7, 9 and 15 sequences generated by the pattern generator (Fig. 3.31). Each of the four channels are tested separately. Fig. 3.35(a) shows the optical input eye diagram at 32Gb/s. Fig. 3.35(b) shows the recovered quarter-rate data eye diagram for 32Gb/s optical data, for one of the channels. Fig. 3.35(c) shows the bath curves
for 32Gb/s and 20Gb/s. Error free (BER=10^{−12}) operation is shown for 0.16UI and 0.33UI for 32 and 16Gb/s respectively. The maximum achievable data-rate (32Gb/s) is limited by the maximum data-rate of the external pseudo random bit sequence (PRBS) generator. Fig. 3.36 shows the measured BER as the optical power is varied for different data rates. From this information we derive the optical sensitivity as a function of data-rate, shown in Fig. 3.36. The receiver achieves more than -12dBm of sensitivity at 16Gb/s, which reduces to -10dBm at 28Gb/s and -8.8dBm at 32Gb/s. Sensitivity degradation with increased data rate is mainly due to reduced bit interval and integration time.

The receivers power breakdown and power efficiency (energy per-bit) are shown in Fig. 3.37. Total power consumption per channel at the highest data rate (32Gb/s) is 4.87mW. The QLL and local ILOs consume a third of the total power. To show the efficacy of the adaptive body biasing scheme,
two sets of measurements are done with the adaptive VBB generator on and off (Fig. 3.37(a)). When adaptive VBB generator is active, the per-bit energy efficiency improves from 103fJ/b at 32Gb/s to 94fJ/b at 16Gb/s. Without the body bias the per-bit energy efficiency at 16Gb/s is 160fJ/b. Table 3.2 summarizes the performance of the parallel optical receiver and QLL-based clocking. Low power QLL based clocking and body biasing helps achieve the highest efficiency compared to the state-of-the art. Ring oscillator based clock distribution helps achieve a very compact design that is the smallest compared to other works.
Table 3.2: Performance Summary of the Parallel Optical Receiver

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>28nm FD SOI</td>
</tr>
<tr>
<td>Data-rate</td>
<td>32Gb/s</td>
</tr>
<tr>
<td>Energy-per-bit</td>
<td>103fJ/b data and 50fJ/b clock</td>
</tr>
<tr>
<td>Active area</td>
<td>$0.3 \times 0.06 \text{mm}^2$ (4 channels)</td>
</tr>
<tr>
<td>Optical Sensitivity</td>
<td>-8.8dBm at 32Gb/s</td>
</tr>
<tr>
<td>QLL Locking range</td>
<td>4-11GHz</td>
</tr>
<tr>
<td>Output integrated jitter (RMS)</td>
<td>577fs at 8GHz, 642fs at 11GHz</td>
</tr>
<tr>
<td>I/Q error</td>
<td>$1.5^\circ$</td>
</tr>
</tbody>
</table>
3.5 Summary

In this chapter, we presented a receiver architecture that is suitable for 3D integration and advanced silicon photonic technologies. The receiver architecture features a double-sampling receiver with dynamic offset modulation and low-bandwidth TIA. Experimental results validated the feasibility of an ultra-low power 25Gb/s receiver and its superior performance over a conventional TIA-based architecture in terms of power consumption, sensitivity, and speed. It was shown that for the proposed architectures, the integrating receiver with a low-bandwidth TIA achieves higher sensitivity at data-rates higher than 8Gb/s compared with conventional TIA-based receiver. Also, we demonstrated that when the front-end input capacitance is smaller than 135fF, higher sensitivity is achieved by employing a low-bandwidth TIA instead of simple resistive front-end. So, for the presented particular topologies the target data-rate determines the choice between conventional and integrating receivers. If an integrating receiver is chosen, the input node capacitance determines the choice between a LBW TIA front-end and a resistive front-end. Two prototypes were fabricated in 28nm bulk CMOS and fully tested to validate our analysis. The double-sampling optical receiver with low-bandwidth TIA and dynamic offset modulation consumes 170µW/Gb/s while operating at 25Gb/s. It has an optical sensitivity of -16.1dBm at 21.2Gb/s, which drops to -14.9dBm at 25Gb/s. The conventional TIA-based receiver consumes 226µW/Gb/s while operating at 21.2Gb/s and has an optical sensitivity of -10.4dBm.

The same receiver architecture was implemented in a source-synchronous parallel optical receiver. This design features quadrature-locked loop (QLL), a frequency tracking technique used to increase the locking range of the ring based quadrature injection locked oscillator. This technique was used to generate accurate quadrature phases from a single phase of electrical/optical clock without any frequency division. In order to maintain the energy efficiency of the receiver at lower data-rates, body biasing was used to adaptively reduce power consumption at lower data-rates. The system was implemented in 28nm FD SOI CMOS and supports up to 32Gb/s of data-rate. An adaptive body biasing scheme was used to achieve an ultra-low power consumption of 153fJ/bit in a wide range of data-rates. The sensitivity of the receiver was measured to be -8.8dBm at 32Gb/s. The QLL measurements show a locking range of 4-11GHz, output integrated jitter of 577fs when using optical clock, and a maximum quadrature error of 1.5°.
Chapter 4

Optical Transmitters

4.1 High-speed Optical Transmitter Overview

![Image of optical transmitter diagram]

Figure 4.1: Optical transmitters are one of the primary elements of optical interconnects.

4.1.1 Direct Modulation vs External Modulation

Two strategies, illustrated in Fig. 4.2, can be employed to perform the operation of transferring electrical data to the optical domain. In the first option, called direct modulation, a beam of light turns on and off with input electrical data of “one” and “zero”. In the second option, external modulation, a continuous wave (CW) laser is used to emit light which has constant power. A modulator is then used to code the beam of light with “one” and “zero”. The simplest form of modulation is on-off keying (OOK) and is simply modulating the optical power of the laser to mark it as “one” or “zero”.

Until now, direct modulation and external modulation have both demonstrated promising performance. Direct modulation of lasers through vertical-cavity surface-emitting lasers (VCSEL) have
demonstrated speeds upward of 56Gb/s citekuchta201356, westbergh2013high. Nowadays, VCSELs have versatile applications in short-range fiberoptic communication such as Gigabit Ethernet. Silicon cannot be used as an active material for the laser, and therefore wafer bonding of III-V materials has been used for integration of lasers in integrated photonic systems [68].

External modulation can be implemented in several ways, among them are electro-absorption and electro-optic modulation. The former relies on the modification of the absorption of a semiconductor material when an external electric field is applied while the latter is based on the change of the refractive index observed for some crystals under an external electric field. A change in refractive index itself does not allow intensity modulation of a CW laser beam. However, using an interferometric structure, such as the Mach-Zehnder structure, or using resonant structures, like micro-rings, one can convert the induced phase modulation into the desired intensity modulation. Compared with direct modulation of laser, external laser sources provide an opportunity to use highly efficient external sources.

In integrated photonic-electronic systems, both direct and external modulation are viable solutions and are actively pursued by academic and industrial entities. For example, Aurrion Inc. has introduced wafer-to-wafer bonded heterogeneous lasers to implement direct modulation schemes, while Luxtera has chosen external modulation using a cheaper fabrication process. In this chapter, we are focused on external modulators and particularly resonant structures.

Figure 4.2: Conceptual block diagrams of optical transmitters. (a) Direct modulation of the laser (b) External modulation of the laser.
4.2 Mico-ring Modulators: Opportunities And Challenges

Compact, low-power, and high-speed electro-optic modulators (EOM) are one of the key components in realization of chip-to-chip optical signaling. Resonant structures such as micro-ring modulators are promising candidates due to their compact size and low power consumption. Examples of such structures (Fig. 4.3) reported in the past few years include carrier-depletion micro-ring modulators [69,70], carrier-injection micro-ring modulators [71], silicon micro-disk modulators [72], and polymer ring modulators [73].

The operation of the simplest micro-ring modulator is based on resonance wavelength shift that occurs due to carrier concentration in the guiding medium. Change in carrier concentration can be achieved in two common ways. The first is by forming a reverse biased PN junction and changing the width of its depletion region by altering the bias voltage. The second way is forming a p-i-n region that gets forward biased above threshold voltage for higher carrier concentration and biased below threshold voltage for lower carrier concentration (Fig. 4.3). With the laser wavelength chosen to be $\lambda_L$, the output optical power changes between two amplitudes, $P_0$ and $P_1$. We briefly analyze the physics behind operation of micro-ring modulators.

Figure 4.3: Index-modulation optical ring modulators and different ways of changing the index of refraction.

\[ \phi(t, \omega) = \omega \tau + \frac{\omega}{n} \int_{t-\tau}^{t} \Delta N(t') dt' \]  

(4.1)

where $\omega$ is the angular frequency of the wave, $\tau = nL/c$ is the resonance round trip, $n$ is the effective index, $L$ is the ring circumference, and $\Delta \lambda(t)$ is resonance wavelength shift due to carrier concentration changes by $\Delta N(t)$. Assuming the change in carrier concentration and refractive index
are small, we have [74]

\[ D(t) = a \exp[-i\phi(t, \omega_0)]C(t - \tau), \]  

(4.2)

where \( a \) is the attenuation coefficient. The relation between the instantaneous field amplitudes can be written as

\[ B(t) = \sigma A + i\kappa a \exp(-i\phi(t))/C(t - \tau) \]

(4.3)

\[ A + C(t) = \sigma B(t), \]

(4.4)

where \( \kappa \) and \( \sigma \) are the coupling and transmission coefficients. For a lossless coupler, we have \( \sigma^2(t) + \kappa^2(t) = 1 \). Solving the equations to get the steady-state transmission, we have [74]

\[ T_{ss} = \frac{B}{A} = \frac{\sigma - a \exp(-i\phi)}{1 - \sigma a \exp(-i\phi)} \]

(4.5)

\[ T_{ss}^2 = \frac{\sigma^2 + a^2 - 2\sigma a \cos(\phi)}{1 + a^2\sigma^2 - 2\sigma a \cos(\phi)}. \]

(4.6)

When \( \sigma = a \), the waveguide destructively interferes with the wave coupled from out of the ring to result zero transmission. This condition is called critical coupling. In order to get large a extinction ratio the micro-ring has to operate near critical coupling. As evident by equation 4.5, it is appealing to have lower loss factor, \( a \), or higher quality factor, \( (Q) \), to get higher extinction ratio.

Compact micro-ring modulators are relatively easy to implement using standard lithography techniques and they offer a low-power and high-speed solution for optical modulators. Nevertheless, there are several challenges associated with wide usage of micro-ring modulators. As explained above, it is desirable to increase the \( Q \) of a micro-ring to increase its extinction ratio for a given change in
index. This will relax the voltage swing requirement in carrier-depletion micro-rings and enhance the energy efficiency of the micro-ring. As the $Q$ of the ring increases, the photon life time increases, leading to a longer time required for energy inside the ring to decay and build up during transitions of “zero” and “one”. This results in a low-pass optical response and therefore there is a trade-off between bandwidth and quality-factor. In this chapter, we propose an entirely new structure, called the differential ring modulator, to overcome this trade-off.

Because of this trade-off, conventional ring structures are designed to have lower $Q$. In carrier-depletion mode micro-rings lower $Q$ results in a relatively high voltage swing required for an acceptable extinction ratio. This voltage is well above current CMOS-compatible voltage headrooms, which is around 1V. One way to avoid this is by using carrier-injection structure which features higher extinction ratio and its voltage requirement is CMOS-compatible. Unfortunately, carrier-injection structures have a limited bandwidth due to their slow carrier dynamics. In this chapter we explore a driver circuit that can extend this intrinsic bandwidth by a novel pre-emphasis technique.

Due to the high thermo-optic coefficient of silicon, the refractive index of the ring can change with fluctuations of ambient temperature. As a result, micro-ring modulators are highly sensitive to operational temperature. The resonance wavelength, $\lambda_0$ may change with temperature and needs calibration. Besides, temperature variations due to self-heating and off-chip thermal noise has to be actively cancelled. We will look at this issue more carefully in the following section and will propose a solution.

Figure 4.5: Diagram of micro-ring modulator with electric field in various locations.
4.3 Monolithic Silicon-photonic PTAT temperature sensor for micro-ring resonator thermal stabilization

As the resonance wavelength of micro-ring modulators is susceptible to temperature fluctuations, they require thermal tuning. Sophisticated techniques and circuitry have been proposed to stabilize the temperature fluctuations in micro-ring modulators [75]-[76]. Examples of such techniques are output optical power feedback [75], bit error-rate (BER) feedback [77], feedback through scattered light [78] and balanced homodyne detection [79]. These techniques require extra optical power on the silicon-photonic chips or complex circuitry for implementation. There have been other attempts to reduce temperature susceptibility of micro-ring resonators using negative thermo-optic materials as an overlay on micro-ring [80]. These attempts require extra fabrication steps and do not fit standard CMOS process flows. Also, any process variation causes imperfect athermalization and fluctuations in resonance frequency. As an alternative athermalization method, micro-ring modulators with integrated temperature sensor and resistive heater have been proposed to enable thermal compensation [76]. However, this effort relies on a single diode, which is prone to process variation and requires careful consideration of die temperature gradient. Temperature gradient caused by the integrated heater and ambient thermal sources can lead to inaccurate and false temperature measurements. In this paper, we propose thermal tuning through a monolithic distributed Proportional To Absolute Temperature (PTAT) sensor. Linear operation of the temperature sensor with and without operational heater is demonstrated over 125°C. Using a temperature feedback loop, the micro-ring modulator is shown to operate at 20Gb/s in the presence of emulated temperature fluctuations.

4.3.1 Structure Overview

The basic principal behind the PTAT temperature sensor is that the difference between the voltage drop of two forward-biased diodes, operating at different current densities, is linearly proportional to absolute temperature. Starting from a simple diode equation, diode current ($I_D$) vs voltage ($V_D$) is approximated to be

$$V_D = \frac{n k T}{q} \ln \left( \frac{I_D}{I_S} \right), \quad (4.7)$$

where $k$ is the Boltzmann constant, $T$ is the absolute temperature, $q$ is the charge of electron, $I_S$ is the reverse bias saturation current, and $n$ is fabrication constant typically between 1 and 2. Two forward-biased diodes with different sizes will have a voltage difference of

$$V_{D21} = V_{D2} - V_{D1} = \frac{n k T}{q} \left( \ln \left( \frac{I_{D2}}{I_{S2}} \right) - \ln \left( \frac{I_{D1}}{I_{S1}} \right) \right) = \frac{n k T}{q} \ln \left( N \frac{I_{D2}}{I_{D1}} \right), \quad (4.8)$$
where \( N \) is the ratio of diodes (i.e., ratio of their reverse bias saturation currents). The slope of the PTAT sensor can be engineered using the ratio of reverse bias saturation currents and the ratio of the currents fed to each diode. In this design, the ratio of currents is 4 and \( N = 5 \) resulting in a factor of 20. For a given voltage sensitivity, it is desirable to maximize the slope of PTAT sensor to achieve higher accuracy in temperature readings. However, if the difference in currents becomes too large, the difference between voltage-drops across the parasitic resistances of diodes cause error. A low-pass filter can significantly reduce the effect of device noise on overall PTAT voltage noise. However, this will limit the bandwidth of the feedback loop. Therefore, in high-speed applications with a limited power budget, there will be a minimum current that can be applied to diodes to satisfy the noise requirements. This leads to a maximum diode ratio \( (N) \). Designing a monolithic PTAT temperature sensor for a micro-ring modulator requires careful consideration for a number of geometric trade-offs. The temperature has to be measured accurately, which requires close proximity of PTAT temperature sensor and the ring. On the other hand, it is desirable to have the integrated heater very close to the ring and utilize maximum perimeter of the ring for modulation. Another important factor is reliability of the temperature sensor in the presence of temperature gradients and the operational heater. Bearing in mind all of the above, the structure
of Fig. 4.6 is proposed. In this architecture the heater is placed under the coupler while the rest of the ring’s perimeter is used for modulation. The PTAT temperature sensor is formed by two distributed diodes around the ring. Different segments of each diode are connected by two metal layers. Cathodes of both diodes are connected to ground and the anode of each diode is connected to a current source. The distance between the PTAT sensor and the ring has to be minimized to the extent that no leakage between the P-doped region of the modulator and N-doped region of the sensor occurs. In this design this distance is chosen to be 8µm, which ensures no leakage. The PTAT temperature sensor is used in a feedback loop to stabilize temperature of the micro-ring (Fig. 4.7). The feedback works by directly measuring the temperature of the ring and applying a heater voltage proportional to temperature error. Temperature error is defined as the difference between target temperature and rings temperature. In a fully integrated system, the feedback loop can employ a programmable gain amplifier (PGA) or an ADC-based digital feedback to maintain a constant temperature. The target temperature is set by the offset introduced in PGA or ADC. The two current sources need to be temperature independent and stable. These current sources can be implemented using bandgap current sources. We have designed a separate CMOS chip that comprises modulator drivers, bandgap current sources, and a programmable gain amplifier. In this

Figure 4.7: Concept of a feedback loop to stabilize micro-ring’s temperature.
Figure 4.8: COMSOL heat transfer simulations. (a) Temperature uniformity across distributed PTAT sensor. (b) Effect of temperature gradient due to heater on PTAT sensor.

demo, GPIB programmable SourceMeters and voltage supplies are used to test performance of the silicon photonic chip. Fig. 4.8 shows COMSOL heat transfer simulation of the proposed structure when the heater is off and the ring is being self-heated by the absorbed light (a) and when the heater is on (b). The distributed nature of the PTAT sensor makes it minimally susceptible to errors due to temperature gradients. From the first simulation, different fragments of the PTAT sensor see the same temperature isosurface from the first simulation. Also, when the heater is on with maximum power, more than 75% of the ring has less than a 10% temperature difference with components of the PTAT sensor.

4.3.2 Supporting Chip

A prototype comprising a carrier-depletion micro-ring modulator and monolithic PTAT temperature sensor is designed in OpSIS IME-5 platform. This platform enables interconnection of distributed PTAT temperature sensor using two metal layers [81]. Fig. 4.9 shows the die micrograph of the micro-ring modulator with integrated heater and PTAT sensors. The total active area of the ring and PTAT sensor is less than 100µm×100µm. The total area of the test chip including pads and grating couplers is 700µm×200µm.

4.3.3 Measurement Results

The prototype was fully characterized via both DC and high-speed measurements. The DC characteristics of the micro-ring are shown in Fig. 4.10. The full-width half-maximum of the transmission spectra is measured to be 0.33nm, resulting in a Q of around 4700 (Fig. 4.10(a)). Tunability of the ring is measured to be 0.12nm/mW (Fig. 4.10(b)). The free spectral range (FSR) of the micro-ring
is measured to be 5nm. In order to characterize the performance of the PTAT sensor, a heater is attached to the die and the difference between diode voltages versus temperature is measured. As mentioned earlier, the diodes are designed with a ratio of $N=5$. The DC currents used for the PTAT sensor are $2.5\mu A$ and $10\mu A$. Total power dissipated in the diodes of the sensor is approximately $9\mu W$. Note that it is important to keep this power dissipation low to avoid heating up the PTATs junction. The residual contact via and wire-bond resistance is estimated to be less than $45\Omega$ for diodes. To minimize voltage drop across this resistance, i.e., error in PTAT temperature reading, diode currents are kept low. The saturation currents are measured to be $2 \times 10^{-18} A$ and $10^{-17} A$ for the smaller and larger diode, respectively. For every wavelength the corresponding notch wavelength is also measured. Fig. 4.10(c) illustrates these measurements and the linear equation fitted
Figure 4.10: (a) DC static transmission of the micro-ring. (b) Measured integrated heater tunability of the micro-ring. (c) Measured PTAT voltage versus temperature. (d) Measured micro-ring resonance wavelength versus temperature.

to predict the temperature from PTAT readings. Functionality of the micro-ring modulator is first verified without ambient thermal noise. An RF probe is used to modulate the micro-ring and optical probes are used for carrying CW beam of laser to the input grating coupler and from output grating coupler. A high-speed PRBS-15 sequence is then used with a reverse bias of -3.5V and peak-to-peak
modulation depth of 5.5V at 10Gb/s and 20Gb/s (Fig. 4.11). The mico-ring achieves up to 20Gb/s of data rate with an extinction ratio of 4dB. Fig. 4.12 shows the measurement setup for high-speed measurements of the modulator. A peltier thermoelectric heater/cooler is used to emulate temperature fluctuations of the ring. The peltier heater/cooler provides a maximum temperature difference of 47°C from a maximum current of 5A. The peltier cooler’s current is modulated with a 0.5Hz square wave such that the temperature of the ring changes by 3.2°C every second. The feedback loop of Fig. 4.7 is formed using two GPIB programmable SourceMeters connected to the PTAT sensor and a programmable power supply connected to the integrated heater. The heater, which is a P-doped area under the coupling region of the micro-ring, has a resistance of 2kΩ. The SourceMeters provide constant currents of 2.5µA and 10µA for the PTAT sensor and are used to read the voltage difference between the two diodes. This voltage is then compared with a preprogrammed target voltage and the difference is multiplied by -10000 and applied to the integrated heater. The pre-programed PTAT target voltage in this case was 140.3mV corresponding to 29°C. Equation (3) summarizes the feedback loop’s operation. Without external temperature perturbations, the feedback sets the heater voltage such that the micro-ring’s temperature matches the target temperature of 29°C or PTAT voltage of 140.3mV. The heater voltage associated with this setting was measured to be 2.8V. Figs. 4.13(a) and 4.10(b) shows the peltier heater/cooler current, the PTAT voltage readings and the voltage of the integrated heater produced by the feedback loop. The heater voltage, which changes from 0.74V to 2.48V, corresponds to a 2.8mW change in the heater’s power. According to Figs 4.10(b) and 4.10(d), this corresponds to 3.2°C change in temperature of micro-ring. Note that due to the presence of PCB, the heat slowly diffuses from peltier heater/cooler to the silicon photonic die. This is the primary limit in demonstrating higher bandwidth for the feedback loop, as the PTAT sensor itself can track faster temperature fluctuations. Fig. 4.13(c) shows the output optical eye diagram in the presence of emulated ambient temperature noise without temperature stabilization feedback loop. Fig. 4.13(d) shows the output optical eye diagram when the feedback loop is turned on. In this experiment, the feedback loop was controlled by a GPIB programmable SourceMeters and power supplies connected to a computer; however, a low-power ADC-based feedback or programmable gain amplifier can be employed to control the heater voltage in the feedback loop. The power consumption of this temperature stabilization method, as other methods that use a heater, is dominated by the heater power. In this experiment the average heater power consumption was 1.3mW. But the heater power consumption depends on the target temperature, since it is possible to use the same heater to overcome process variation. The feedback loop can be used to set the temperature by ensuring PTAT voltage matches a pre-programed value. A one-time calibration can be used to set this value to adjust the resonance wavelength to wavelength of interest. In order to make sure we can set the resonance wavelength to any wavelength, the FSR of 5nm has to be covered. In this case 41mW is required to cover the entire FSR. This requirement can be significantly reduced by
selective removal of the SOI substrate to increase thermal impedance of the device [82].

![Figure 4.12: High-speed measurement setup with induced thermal fluctuations and temperature stabilization feedback loop.](image)

![Figure 4.13: (a) Peltier heater/cooler supplied current over time. (b) Closed loop integrated heater voltage and PTAT voltage. (c) Output optical eye diagram without thermal tuning feedback. (d) Output eye diagram with thermal tuning feedback.](image)
4.4 Differential Optical Ring Modulator: Breaking the Bandwidth/Quality-factor Trade-off

Improving bandwidth-density product in a fully integrated silicon photonic systems necessitates a corresponding enhancement in modulator performance. Ring resonator modulators are promising candidates to realize compact, high-speed, and low-power silicon photonic transceivers [83]. Intensity modulation is commonly achieved by index modulation or coupling modulation. It is desirable to have high-\( Q \) rings as, for a given extinction ratio, higher \( Q \) results in better energy efficiency. However, there is a trade-off between the \( Q \) of the ring resonator modulator and its optical bandwidth. As previously shown [74], the time domain dynamic transmission of the ring, \( T(t) \), can be written as

\[
T(t) = \sigma(t) + \frac{\kappa(t)}{\kappa(t - \tau)} a(t) \exp[-i\phi(t)] \times [\sigma(t - \tau)T(t - \tau)]
\]  

(4.9)

where \( \sigma \) and \( \kappa \) are transmission and coupling coefficients, \( a \) is the attenuation, \( \phi \) is the phase shift inside the ring, and \( \tau \) is the resonator round trip. Fig. 4.14 shows numerical solution of equation

Figure 4.14: Index-modulated micro-ring’s (a) static transmission, (b) optical frequency response, (c) simulated \( Q \) versus -3dB bandwidth.

Figure 4.15: (a) Index-modulated ring. (b) Coupling-modulated ring. (c) Proposed differential ring modulator.

4.9 using an iterative approach. The \( Q \)-bandwidth trade-off in an index-modulated ring is shown in Figs. 4.14 (b) and (c), which results in the low-pass response for the index-modulated ring of
Fig. 4.15 (a). Conversely, a high-pass response can be obtained using the coupling modulated ring of Fig. 4.15 (b), with a sufficiently fast variable coupler [84]. In this case, a long sequence of 1’s causes energy droop in the ring and signal degradation (Fig. 4.15 (b)). We propose a differential ring modulator that overcomes the $Q$-bandwidth trade-off in ring modulators (Fig. 4.15 (c)). This structure does not exhibit droop in the energy stored in the ring.

$$T(t) = \sigma(t) + \frac{\kappa(t)}{\kappa(t - \tau)} a(t) \exp[-i\phi(t)] \times [\sigma(t - \tau)T(t - \tau)],$$  \hspace{1cm} (4.10)

where $\sigma$ and $\kappa$ are transmission and coupling coefficients, $a$ is the attenuation, $\phi$ is the phase shift inside the ring, and $\tau$ is the resonator round trip.

4.4.1 DRM Structure Overview

The block diagram of the proposed structure (shown in Fig. 4.16) consists of two variable couplers, each of which consists of two differential phase shifters and two 3dB couplers. A Y-junction with a controllable thermal phase shifter is used to split the input beam into two beams with the same phase (A1 and A2). The variable couplers operate out of phase and when coupling of one increases the other decreases. The proposed DRM achieves considerably lower $V_\pi$ compared to a regular MZI.

Figure 4.16: Block diagram of the differential ring modulator.
modulator. Considering the electric field at various locations in the DRM, it can be shown that at
resonance the static transmission is

\[ |B_1/A_1|^2 = (a - |\cos \frac{\Delta \phi}{2}|)^2/(1 - a|\cos \frac{\Delta \phi}{2}|)^2 = (a - |\cos \frac{V\phi}{2V\pi}|)^2/(1 - a|\cos \frac{V\phi}{2V\pi}|)^2, \]  

(4.11)

where \(a\) is the loss factor and \(V_\pi\) is the voltage required to achieve a differential phase shift of \(\pi\) in
phase shifters. Critical coupling (where maximum extinction ratio is achieved) happens when

\[ V = V_{\pi,DRM} = \frac{2V_\pi}{\pi} \times \cos^{-1}(a). \]  

(4.12)

Therefore, for a \(Q\) of 32,000, \(V_{\pi, DRM}\) is 8 times smaller than \(V_\pi\). The DRM structure maintains
the energy stored in the ring constant in all conditions. This can be demonstrated by calculating
the amplitude of \(C_1\) and \(C_2\) (shown in Fig. 4.16) when data switches from 1 to 0 here:

\[ |C_1|^2 = |C_2|^2 = \sin \frac{\Delta \phi}{2}(1 + a^2/4 \cos^2 \frac{\Delta \phi}{2})/(1 - a^2/4 \cos^2 \frac{\Delta \phi}{2}). \]  

(4.13)

Intuitively, by modulating the couplers differentially, the overall coupling to the ring remains con-
stant. Thus, the variation of energy stored in the ring is minimized.

Figure 4.17: Die micrograph of the fabricated prototype. (a) Optical input grating coupler. (b) Y-
junction. (c) Heater for phase shift controller. (d) Photodiode connected to one output for testing
purposes. (e) Optical output grating coupler.
4.4.2 Supporting Chip

A prototype chip has been fabricated in OpSIS IME platform [85] and occupies less than 0.35\(mm^2\) (Fig. 4.17). Grating couplers are used for optical input and output. A heater is placed in the center of the ring for uniform distribution of temperature. A second heater is placed at one of the Y-junctions branches to calibrate phase mismatch between the two inputs of the ring. The second output of the ring is connected to an on-chip photodiode for testing.

![Diagram of measurement setup](image)

Figure 4.18: Measurement setup.

4.4.3 Measurement Results

The chip was wire-bonded to a custom designed PCB that carries high-speed and DC signals (4.18). A tunable laser source followed by an EDFA was used as the input and the output was monitored by optical sampling scope. The high-speed differential data signals were driven by a PRBS-31 sequence using a pattern generator. The voltage swing for each single-ended signal was 1.75Vp-p. Fig. 4.19 shows measured eye diagrams at the output for 5Gb/s and 10Gb/s data streams. The extinction ratio of the output optical data is measured to be 6.2dB. Some of the noise seen at the output is associated with the EDFA noise and limited sensitivity of the optical sampling scope. Fig. 4.20 (a) shows measured static transmission of the ring near one operational wavelength bias points. From this measurement, the \(Q\) of the ring is derived to be roughly 32,000. The tunability of the ring was measured by varying the input voltage of the heater, and was measured to be 12.3pm/mW.
Figure 4.19: Output optical eye diagram of the differential ring modulator operating (a) At 5Gb/s (b) At 10Gb/s.

Figure 4.20: (a) Measured steady-state transmission of the micro-ring near a resonance wavelength (b) Tunability of the micro-ring using a heater.
4.5 Carrier-injection Micro-ring Modulator Transmitter with Switched-capacitor Pre-emphasis

Recent advances in Silicon photonics have demonstrated its potential to enable next-generation low-cost high-speed optical interconnects. Electro-optic modulators that are CMOS-compatible, compact, and low power are essential elements in realization of chip-to-chip optical signaling. Carrier injection micro-ring modulators are one of the promising candidates [86], [87]. Compared with the carrier depletion micro-ring modulator, they can operate with higher extinction ratio and with CMOS-compatible drive voltages. However, the speed of carrier-injection rings is limited to slow carrier dynamics and necessitates pre-emphasis to compensate for their nonlinear transient behavior. Both types of micro-ring modulators are also susceptible to temperature variations and need wavelength stabilization loops. Here we present a hybrid-integrated CMOS-SiPh transmitter that tackles these challenges. A low-power switched-capacitor-based (SC) pre-emphasis technique that effectively compensates for the modulator bandwidth limitation is proposed. Using the proposed monolithic PTAT temperature sensor, a feed-forward bias-based wavelength stabilization technique is also presented. Carrier-injection micro-ring modulators are inherently slow and limited by recombination lifetime of carriers in the intrinsic region of the p-i-n junction (Fig. 4.21). In a forward-biased p-i-n junction, the free carrier density change is achieved through minority carrier injection. The slow carrier generation/recombination processes for the forward-biased p-i-n junction usually limits the device speed to MHz range. The rise-time of the carrier dynamics is limited by the time needed to fill the intrinsic region with enough charge, and the fall-time is limited by the extraction time of

Figure 4.21: Carrier-injection micro-ring modulator structure.
the carriers. Nonlinear pre-emphasis has proven to be an effective way of reducing carrier dynamic rise-time and fall-time [86], [87]. The optical rise-time is determined by the time needed to inject sufficient charge so that the junction reaches steady state saturation limit. Therefore, by increasing the pre-emphasis voltage, rise-time is shortened and a higher data-rate is achievable.

![Figure 4.22: Top-level block diagram of the transmitter and the proposed switched-capacitor-based pre-emphasis technique.](image)

### 4.5.1 Transmitter Architecture Overview

Prior pre-emphasis techniques relied on stacked output drivers that are highly power-inefficient and have a maximum pre-emphasis voltage drive of $2 \times \text{VDDL}$, where VDDL is the thin-oxide transistors' voltage. In this work, we use a low-power SC-based pre-emphasis technique that can boost the output voltage to $4 \times \text{VDDL}$. A top-level block diagram of the transmitter with proposed SC-based pre-emphasis technique is shown in Fig. 4.22. The driver consists of three main elements: a conventional voltage driver, and two pre-emphasis blocks for rising and falling data edges. There are two voltage
levels required for operation of this scheme, VDDL=1V, set by the standard thin-oxide transistors voltage and VDDH=2V. The conventional voltage driver provides a steady state voltage to keep the junction in forward bias when needed. The two pre-emphasis blocks work by first accumulating charge on $C_1$ and $C_2$ up to VDDH. Subsequently, these capacitors are switched so that for the rising-edge pre-emphasis the output is at $2 \times VDDH$ and for falling-edge pre-emphasis the output is at $-VDDH$. The charge on these capacitors is used as pre-emphasis to inject and extract charge from the intrinsic region of the junction. The monolithic PTAT sensor that was introduced in section 4.3 is used in a feed-forward bias-tuning wavelength stabilization scheme.

Figure 4.23: Schematic circuit details of the proposed micro-ring modulator driver with switched-capacitor pre-emphasis.
4.5.2 Supporting Circuits

Fig. 4.23 shows the schematic circuit details of the SC-based pre-emphasis technique. A 2V pulsed-cascode stage, similar to [86], is used to charge capacitor $C_1$ and $C_2$. Tunable delays are used to alter the charge time of these capacitors, and therefore adjust the strength of the pre-emphasis. A voltage driver with digitally adjustable pull-up and pull-down strengths is incorporated to maintain the junction in the forward bias region and in the off region according to the data. Another challenge for robust operation of micro-ring modulators is their sensitivity to temperature fluctuations. In this work, we propose wavelength stabilization by direct measurement of temperature through a monolithic distributed PTAT sensor. Fig. 4.24 shows the schematic block diagram of the feed-forward bias-based wavelength stabilization technique and the SiPh micro-ring modulator with on-chip PTAT temperature sensor. The monolithic PTAT temperature sensor, used for directly measuring the temperature of the ring, is described in [15]. In [15], the PTAT temperature sensor
operation was demonstrated in a carrier-depletion MRM with heater-based wavelength stabilization and without using a CMOS chip. The PTAT sensor works by measuring the voltage difference between two diodes with different current densities. The PTAT voltage is then applied to a programmable gain amplifier (PGA) implemented in the CMOS chip. For effective operation of the feed-forward wavelength stabilization, the PGA sets the bias voltage of the micro-ring modulator according to the following condition:

$$A_{PGA} \ln \frac{I_{s1}}{I_{s2}} \frac{K_B}{q} \beta_{Bias} = \beta_{Temp},$$

(4.14)

where $\beta_{Temp}$ is the wavelength versus temperature slope, $K_B$ is the Boltzmann constant, and $\beta_{Bias}$ is the wavelength versus Bias slope. Silicon has a $\beta_{Temp}$ of 0.11nm/K and $\beta_{Bias}$ varies in a range of 0.1-0.3V/K, varying in different wafers and devices. A gain of 5-10 (depending on $\beta_{Temp}$ and variation of currents) cancels the temperature dependency of micro-ring modulators notch wavelength. For linear operation of the PTAT sensor, the current of the diodes should be kept constant at different operating temperatures. The PTAT sensor currents are provided by the CMOS chip using a current
bandgap circuit, which provides the two output currents $I_1$ and $I_2$. The electronics of the optical transmitter is fabricated in a 65nm CMOS bulk process and the silicon photonic device is fabricated in the OpSIS IME-5 process. The silicon photonic micro-ring modulator with integrated PTAT sensor is connected to the CMOS chip through wirebonds, as shown in Fig. 4.25. Total active area of the transmitter including wavelength stabilization circuitry is $0.15\text{mm}^2$. The micro-ring modulator has a radius of $20\mu\text{m}$ with total active area of $0.01\text{mm}^2$, including the integrated PTAT temperature sensor.

![Static Transmission of the micro-ring](image1)

![Optical Frequency Response](image2)

![10Gb/s output eye w/o pre-emphasis](image3) ![10Gb/s output eye w pre-emphasis](image4)

Figure 4.26: Measured characteristics of the micro-ring modulator. Measured output optical eye diagram of the optical transmitter with and without pre-emphasis.

### 4.5.3 Measurement Results

Fig. 4.26 shows optical measurement results of the transmitter. The static transmission of the micro-ring modulator shows a $Q$ of $\sim6000$ and free spectral range (FSR) of $\sim5\text{nm}$. The measured optical frequency response of the micro-ring modulator in forward-bias shows a -3dB bandwidth of about $900\text{MHz}$. When a $10\text{Gb/s}$, PRBS-7 data stream is transmitted the output optical eye is completely closed without pre-emphasis. Enabling pre-emphasis opens the eye to have 7dB extinction
ratio. The transmitter consumes 3.42mW, resulting in per-bit energy of 342fJ/b. This is the lowest reported energy/bit for optical transmitters. Measurements verify that currents provided by the bandgap current source vary less than 5% in a range of 25-150°C. Note that process and voltage variation can be compensated by adjusting gain of the programmable gain amplifier (PGA). Fig. 4.27 shows measured operation of the feed-forward bias-based wavelength stabilization technique. First, the temperature dependency of the micro-ring modulators resonance wavelength is measured to be about 0.11nm/K. The linear operation of the PTAT sensor is independently verified from 25°C to 150°C. Next, the optimal PGA gain is found to be 8.2 to make the notch wavelength temperature-independent. A 500MHz PRBS-7 data stream is transmitted in presence of emulated ambient temperature noise with and without wavelength stabilization. The output optical eye diagram of the transmitter is virtually closed without wavelength stabilization and opens to have more than 7dB extinction ratio with wavelength stabilization. The maximum tuning power is 290µW for a resonance wavelength range of 0.4nm. In order to cover the complete FSR, this technique can be used as a fine-tuning in combination with heater-based thermal control as coarse-tuning [88].

Figure 4.27: Characteristics of the PTAT sensor and temperature-dependency of the micro-ring modulator's resonance wavelength. Output optical eye diagram of the micro-ring modulator with and without wavelength tuning stabilization in presence of emulated ambient temperature noise.
optical transmitter achieves energy efficiency of 342fJ/bit at 10Gb/s. The feed-forward bias-based wavelength stabilization circuit consumes 0.29mW and provides a low-cost energy-efficient technique to overcome temperature sensitivity of MRMs. Table 4.1 summarizes the system performance.

Table 4.1: Optical Transmitter Performance Summary

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ring Radius [µm]</td>
<td>20</td>
</tr>
<tr>
<td>Ring Mode</td>
<td>Carrier Injection</td>
</tr>
<tr>
<td>Ring Quality Factor</td>
<td>6000</td>
</tr>
<tr>
<td>Technology</td>
<td>65nm</td>
</tr>
<tr>
<td>TX Optical Data Rate [Gb/s]</td>
<td>10</td>
</tr>
<tr>
<td>Extinction Ratio</td>
<td>7dB</td>
</tr>
<tr>
<td>Tuning Method</td>
<td>Voltage Bias</td>
</tr>
<tr>
<td>Wavelength Tuning Range [nm]</td>
<td>0.4nm</td>
</tr>
<tr>
<td>Tuning Power Consumption [mW]</td>
<td>0.29mW</td>
</tr>
<tr>
<td>Active Area [mm²]</td>
<td>0.15mm²</td>
</tr>
<tr>
<td>TX Power Consumption [mW]</td>
<td>3.42mW</td>
</tr>
<tr>
<td>TX Energy/bit [fJ/bit]</td>
<td>342fJ/b</td>
</tr>
</tbody>
</table>
4.6 Summary

In this chapter, we presented contributions in optical transmitters based on micro-ring modulators. First, we presented a compact micro-ring modulator with integrated heater and monolithic PTAT temperature sensor. The distributed design of the PTAT sensor makes it minimally susceptible to temperature gradients. The total active area of the micro-ring including PTAT sensor (not including pads) is 100m×100m. Linearity of the PTAT sensor is shown over 125°C. A temperature stabilization loop is demonstrated to compensate for ambient thermal noise. The closed-loop system operates at data-rates up to 20Gb/s in the presence of temperature fluctuations.

Second, a differential ring modulator structure is presented that breaks the optical bandwidth/quality factor trade-off known to limit the speed of high-Q ring modulators. This structure maintains the total energy stored in the ring constant, unlike coupling modulation schemes, and hence the ring does not suffer from power droop when long sequences of 1’s or 0’s are transmitted. A prototype has been fabricated and 10Gb/s operation of the ring is demonstrated.

Finally, we presented a CMOS-SiPh optical transmitter based on carrier-injection ring modulators. It features a novel low-power switched-capacitor-based pre-emphasis that effectively compensates for the modulator bandwidth limitation. A feed-forward bias-based wavelength stabilization technique via the monolithic PTAT sensor is also presented. The optical transmitter achieves energy efficiency of 342fJ/bit at 10Gb/s and the wavelength stabilization circuit consumes 0.29mW.
Chapter 5

Clock Generation

5.1 Overview of Clock Generation for Optical Links

Frequency synthesizers have versatile application in on-chip clock generation. Frequency synthesizers in the form of frequency multipliers play a key role in design of high-speed electrical and optical links, shown in Fig. 5.1. As the aggregate bandwidth requirement for chip-to-chip interconnects grows, their respective frequency of operation increases [89]. Additionally, high-speed links timing noise heavily depends on jitter performance of clock multiplier, which takes a low-jitter low-frequency reference and generates the high frequency clock for the transceivers. This necessitates design of high-frequency clock generators that are inherently low-jitter and resilient to noise induced by outside sources such as supply noise or digital switching noise.
5.2 Prior Art In Low-jitter Clock Generation

Figure 5.2: Example of prior art. (a) Conventional multiplying PLL (MPLL). (b) Edge-combining DLL-based frequency synthesizer. (c) MDLL. (d) Injection-locked-MPLL (IL-MPLL).

Conventional multiplying phase-locked loop (MPLL) architecture, shown in Fig. 5.2(a), has been the dominant architecture in this field for many years [90]- [91]. PLLs are appealing for use as frequency synthesizers due to their low complexity and because their architecture can support a programmable multiplication rate. More recently, delay-locked loop (DLL) based frequency synthesizers have been under exploration [92]- [93]. A conventional MPLL exhibits higher jitter than a multiplying DLL (MDLL) with the same building blocks and noise environment due to higher jitter accumulation [94]. In the MPLL of Fig. 5.2(a), a step variation in the phase of the ring oscillator gets integrated over many cycles until the loop filter can respond while it only causes a constant offset in a DLL. This results in a considerably larger peak phase error for a conventional PLL compared to DLL. Reducing jitter accumulation, which has motivated this work, is particularly important in systems with large digital switching noise where a clean reference is available. Prior to presenting the proposed architecture, we briefly review other clock multiplying architectures and particularly
those that suppress jitter accumulation. Fig. 5.2(b) shows a possible architecture for a DLL based frequency synthesizer [95]. In this scheme, an edge-combining logic processes equal phases of the reference clock to produce higher frequency clock. As the reference clock is used to generate the output clock, with every new clock edge jitter from previous cycles is removed from output. Therefore, this architecture limits jitter accumulation and does not rely on high loop bandwidth for jitter correction. On the other hand, this architecture has two primary drawbacks. Delay element non-idealities such as mismatch cause duty cycle distortion and a fixed-pattern jitter. Also, in this architecture it is difficult to achieve programmability for the required multiplication ratio. Fig. 5.2(c) shows another possible scheme for clock generation called multiplying DLL (MDLL) [92]. MDLL works by replacing the rising edge of the reference clock with the output clock in every reference cycle. With a clean reference, each rising edge of the reference clock zeros the phase error of the output. Therefore, jitter accumulation is limited to one reference clock period. Since a single delay element is used to generate the edges of output clock there is no fixed pattern jitter due to delay element mismatch. Besides, the select logic can be digitally controlled to program the multiplication rate. The logic block that generates a select pulse, which opens an aperture for reference injection, limits jitter performance and the maximum output frequency of MDLL. Recent works show how an aperture position tuning technique can enhance jitter performance of MDLLs and significantly reduce reference spurs [96,97]. Also, there are other building blocks in MDLLs as well as MPDLLs, such as phase frequency detector (PFD) and charge pump (CP), that impact their overall performance. Extensive work has gone through reducing power consumption of these building blocks, enhancing their speed, and eliminating their sensitivity to device mismatch [90,98]- [100]. Note that while suppressing jitter accumulation is important in frequency synthesizers, it is not necessarily the dominant source of jitter at the output of a frequency synthesizer. In case the reference clock itself has significant jitter, unlike MDLL, an MPDLL can potentially filters out the reference clock jitter. More specifically, a PLL with a sufficiently narrow low-pass filter can filter high frequency components of the input noise. This is not always desirable as larger filter bandwidth results in better VCO noise filtering, faster settling time, and smaller filter area. Other PLL-based frequency synthesizers such as sub sampling PLLs [101,102] and PLLs with dynamic phase error compensation [103] have been proposed to increase the loop bandwidth without compromising system performance. Furthermore, there are other metrics that can be the bottleneck of frequency synthesizers performance depending on the application. For example, jitter peaking, which is the amplification of jitter transfer function over a certain frequency band, presents itself in conventional second-order PLLs with a closed loop zero. Jitter peaking in clock distribution networks with several cascaded PLLs or DLLs can cause significant performance degradation. Therefore, several techniques have been proposed to minimize or completely eliminate jitter peaking by removing the closed-loop zero [104]- [105]. Note that jitter peaking is characterized by transfer function of reference to output while jitter accumulation can
be described as how the output clock responds to noise on the control line of the VCO. [106] shows how certain types of MDLLs (e.g., Fig. 5.2(c)) suffer from jitter peaking while they exhibit jitter accumulation only during one reference period. Another class of frequency synthesizer architectures is those exploiting injection locking. Injection locked PLLs (IL-PLL) (shown in Fig. 5.2(d)), multiplyi

5.3 First-order Clock Multiplier

Figure 5.3: Proposed first-order frequency synthesizer.

An alternative approach for reduction of jitter accumulation is presented in [16] (Fig. 5.3). This architecture directly injects the rising edge of the reference clock to the output clock, resetting jitter accumulation similar to an MDLL. Also, first order frequency detection and frequency correction is used, providing unconditional stability. The first-order frequency synthesizer architecture, shown with more details in Fig. 5.4, provides programmable clock multiplication with reduced jitter accumulation compared with conventional PLLs and without challenges in producing a select signal for open aperture in a multiplying DLL [92, 110]. This architecture utilizes a phase-interpolator (PI) based reference injection proposed in [111] for a burst-mode CDR application. PI-based reference injection provides high-frequency feed-forward reference injection without using a MUX as in MDLLs [110]. In contrast to ILO-based reference injection, in the PI-based reference injection jitter accumulation suppression does not degrade as reference deviates from VCO’s natural frequency (the output rising edge is effectively replaced with the reference rising edge). Also, unlike MDLLs, the proposed architecture can utilize both LC oscillator and inverter-based ring oscillator, depending on application and jitter requirements. In the prototype presented in this paper an LC VCO has been used for low phase-noise demonstration. Frequency acquisition is achieved by a low-power implementation of rotational frequency detector [112], in conjunction with a digital coarse-tuning. Reference of the proposed architecture must be kept very clean since any jitter on this signal will pass directly to the output. Since the performance of this frequency synthesizer is highly dependent on the quality of the reference clock, in addition to electrical reference clock, as an extra feature, the prototype chip is capable of receiving a low jitter optical reference clock generated by a high-repetition-rate mode-locked laser [113]. Also, any non-linearity in the design of reference injection causes a systematic jitter at the output clock. Therefore, careful attention to the design of the sample-and-holds and phase-interpolator for reference injection is required. The proposed frequency synthesizer is implemented in a 65nm CMOS technology. It operates with a reference clock in the
range of 400MHz to 1GHz to generate an output signal in the range of 8-9.5GHz. The frequency synthesizer can be programmed to multiply the frequency by any value between 8-24, as long as the reference clock frequency and output clock frequency fall within ranges mentioned above. The prototype is fully tested with an electrical reference clock and its performance is measured. Subsection 5.3.1 explains the overall system architecture and principles of operation. Subsection 5.3.2 explains the circuit implementation details. Subsection 5.3.1 covers experimental results. Subsection 5.3.1 covers experimental results. Finally, section 5.3.4 provides an analysis for the architecture and the experimental results.

![Top-level architecture of the first-order frequency synthesizer](image)

**Figure 5.4**: Top-level architecture of the first-order frequency synthesizer.

### 5.3.1 Clock Multiplier Architecture

The basic operation of this system can be broken into three main elements. The reference clock injection restarts the phase at every rising edge of the reference clock, and will make the output clock phase independent of the VCO phase (the effect of circuit non-idealities will be discussed later in section 5.3.4). The coarse-tuning forces the output clock to have exactly $M$ rising edges in one reference clock period, and the fine-tuning, which has first order dynamics, tunes the output frequency to be exactly $M$ times the input frequency. First we explain the operation of PI-based reference injection. Following, the algorithm for digital coarse tuning is discussed. Finally, details of fine-tuning rotational frequency detection are explained. The PI-based reference injection technique is shown in Fig. 5.5. The quadrature clocks ($CK_I$ and $CK_Q$) with arbitrary phase of $\phi_0$ are sampled
at the rising edges of the reference clock (b and a). These samples are then used to interpolate between $CK_I$ and $CK_Q$ using the phase interpolator. As shown in Equation 5.3, a reference clock rising edge at $t_0$ results in an output with zero crossing at $t_0$, regardless of the absolute phase of $CK_I$ and $CK_Q$ (Fig. 5.5).

![Figure 5.5: Principle of phase-interpolation based reference injection.](image)

$$a = CK_Q(t_0) = -\cos(2\pi ft_0 + \phi_0) \quad (5.1)$$

$$b = CK_I(t_0) = \sin(2\pi ft_0 + \phi_0) \quad (5.2)$$

$$PI_{out} = b \times CK_Q - a \times CK_I = \sin(2\pi f(t - t_0)) \quad (5.3)$$

This technique has been previously used in implementation of a burst-mode CDR [111]. Reference injection resets any unwanted noise event on VCO node on the next reference cycle. Furthermore, it clears low-frequency components of the VCO phase noise by injecting the reference to the output. The flow chart in Fig. 5.6 demonstrates the operation of the coarse-tuning block. The primary function of the coarse-tuning block is to make the frequency synthesizer programmable and increase its locking range. The coarse-tuning block is a low-power digital circuit, comprising two counters and two comparators. It counts the number of rising edges of output clock in each reference clock period, and compares it to the target multiplication factor. The output of the comparator changes a 4-bit word that controls part of the capacitance of the quadrature LC VCO, adjusting its frequency of operation. Since the phase interpolator injects the rising edge of the reference, VCO’s output can have the same number of rising edges for a range of input frequencies. The coarse-tuning procedure works throughout acquisition, and as soon as the target multiplication rate is achieved, a coarse-tuning binary code is fixed. A separate block fine-tunes VCO’s frequency with an additional varactor. While the fine-tuning is active, the number of rising edges within one reference period remains constant. Therefore, coarse tuning makes no corrections after the binary code is fixed. During the few cycles that the coarse-tuning sets the binary code, the fine-tuning corrections are negligible (due to the small loop gain of fine-tuning). The basic idea of a fine-tuning frequency
detector, known as a rotational frequency detector [112], is shown in Fig. 5.7. Following is a brief explanation of the concept. The input of the fine-tuning block is the same samples taken for interpolation (a and b). Its output is UP and DN bang-bang pulses for a charge-pump that corrects control voltage of the quadrature LC VCO. The following equations describe the nth sample taken from $CK_I$ and $CK_Q$ as follows:

\[ CK_{I,S} = \cos(2\pi f_{VCO} t_n + \phi_0) \]  
\[ CK_{Q,S} = \sin(2\pi f_{VCO} t_n + \phi_0) \]  

where $t_n$ is when the nth sample is taken, $\phi_0$ is the arbitrary phase of the VCO output, and $f_{VCO}$ is the VCO oscillation frequency. $t_n$ is defined as

\[ t_n = \frac{n}{f_{ref}} \]  
\[ f_{VCO} = Mf_{ref} \pm \delta f, \]
where $\Delta f$ is the frequency error of the LC VCO. Plugging (5.6) and (5.7) in (5.4) and (5.5), we obtain

$$CK_{I,S} = \cos(2\pi Mn \pm \frac{2\pi n\delta f}{f_{ref}} + \phi_0) = \cos(\frac{2\pi n\delta f}{f_{ref}} + \phi_0)$$  \hspace{1cm} (5.8)

$$CK_{Q,S} = \sin(2\pi Mn \pm \frac{2\pi n\delta f}{f_{ref}} + \phi_0) = \sin(\frac{2\pi n\delta f}{f_{ref}} + \phi_0) = \cos(\frac{2\pi n\delta f}{f_{ref}} + \phi_0 \pm \frac{\pi}{2})$$ \hspace{1cm} (5.9)

which describe the frequency and phase relationship of the samples. As can be seen in Fig. 5.7 the frequency of $CK_{I,S}$ and $CK_{Q,S}$ is analogous to the beat frequency of two interfering waveforms with slightly different frequencies. This frequency is proportional to frequency error of the VCO:

$$f_{\text{beat}} = f_{CK_{I,S}} = f_{CK_{Q,S}} = |f_{VCO} - Mf_{ref}|.$$  \hspace{1cm} (5.10)

The absolute phase difference between $CK_{I,S}$ and $CK_{Q,S}$ waveforms is always $\pi/2$. The sign of frequency error determines which one of $CK_{I,S}$ and $CK_{Q,S}$ leads the other one. As will be discussed in the subsequent sections, this property and equation (5.10) will be used to implement rotational
5.3.2 Circuit Implementation

Fig. 5.8 shows the transistor-level implementation of the main building blocks. Fig. 5.8(a) shows the quadrature oscillator used in the system with four coarse-tuning bits and fine-tuning control voltage. The quadrature LC VCO is implemented by two matched LC VCO’s that are coupled in a quadrature VCO (QVCO) configuration. Antiphase coupling is achieved using pMOS differential pairs. LC tank’s natural frequency is varied by four bits of digital coarse tuning connected to four varactors and one analog voltage connected to a fifth varactor. Varactors are made of a pair of nMOS transistors in accumulation mode. Fig. 5.8(b) shows the master-slave S/H implementation. The master-slave S/H comprises two pass transistors, two transmission gates, and a differential buffer with source degeneration between them. The pass transistors and transmission gates act as track-and-hold elements that are triggered on opposite edges of the reference clock. While the pass transistors hold the voltage levels in parasitic capacitors of the next stage at rising edge, the transmission gates keep the voltage as the reference clock voltage falls. The buffer with source degeneration is used to minimize kickback [114] and charge-sharing from the output transmission gates to the input pass gates [111]. Non-linearity of S/H induces jitter at the output of the phase-interpolator. Source degeneration in the buffer ensures that this systematic jitter is minimized [115]. The pass-transistor is sized to minimize charge injection to lower the reference sours in the output clock. The S/H is optimized to have maximum bandwidth and linearity. Fig. 5.8(c) shows the differential phase interpolator architecture. The interpolation coefficients \((a)\) and \((b)\) are converted from voltage to current using a differential trans-conductance stage. The operation of interpolation is performed in current-mode and is converted to voltage at output nodes with resistors. Clock samples as well as quadrature waveforms are differential to minimize supply noise and phase mismatches.

The circuit implementation of the bang-bang rotational frequency detector is shown in Fig. 5.9. The fine-tuning circuit uses the beat frequency to set the correction rate proportional to the frequency error and phase relationship between the signals. The principle of fine-tuning is based on the beat frequency and phase difference of samples taken from \(CK_I\) and \(CK_Q\). These samples have the beat frequency described in equation (10). In this design D flip-flops are used to determine the phase relationship between \(CK_{I,S}\) and \(CK_{Q,S}\). By extracting the phase information of \(CK_{I,S}\) and \(CK_{Q,S}\), bang-bang UP/DN pulses associated with each case are generated to increase and decrease the fine-tuning control voltage of LC quadrature VCO.

The frequency acquisition behaves as a first order system since corrections are proportional to deviation of output frequency from \(M\) times frequency of the reference clock. Schmitt triggers are used to prevent metastability and unwanted pulses when frequency of oscillator is close to \(M\) times frequency of the reference clock. Note that, unlike conventional PFDs, where UP/DN signals turn
on at the same time, in this frequency detector UP/DN signals are independent.

Fig. 5.10 shows an electrical/optical reference generator. A MUX chooses between the two signals. The electrical reference generator, which is used during the test, is simply a CML/CMOS circuit retrieving the signal from an off-chip source. As an additional feature, the chip can also use optical reference. The optical reference generates sharp rising edges from high-repetition rate mode-locked laser. A tunable feedback circuit resets the voltage before the next pulse arrives. An on-chip tunable RC delay controls the duty-cycle of the reference waveform. Since only the rising edge of the reference is injected to the output clock the duty-cycle and jitter of the falling edge do not play a role in the quality of the output clock.

5.3.3 Experimental Results

The proposed frequency synthesizer is fabricated in 65nm CMOS technology. Fig. 5.11 shows real-time acquisition of the system at 8GHz while locking to a 400MHz reference clock. As coarse-tuning
and fine-tuning loops correct the frequency of the generated clock, the operation of the reference injection can be seen at three different time steps. While the rising edge is always injected with the same relative phase, frequency of the quadrature VCO changes until lock is acquired. The frequency synthesizer has an operating range of 8-9.5GHz limited by LC VCO’s frequency range. Fig. 5.12 (a) shows reference spurs that are measured to be 64.3dB lower than the main carrier frequency. The overall jitter seen during measurement of the proposed frequency synthesizer can be separated to

Figure 5.9: Details of fine-tuning frequency-detection.

Figure 5.10: Electrical/optical reference generator.
Figure 5.11: Real-time acquisition for a reference clock of 400MHz and output clock of 8GHz.

Figure 5.12: (a) Measurement showing reference spurs at 8GHz with 400MHz reference. (b) Output clock jitter histograms. (c) Total jitter measurement. (d) Phase noise measurement.

Jitter due to device noise and jitter due to non-linearity of different components:

\[ \sigma_{\text{total}}^2 = \sigma_{\text{random}}^2 + \sigma_{\text{systematic}}^2 \]  

(5.11)

where \( \sigma_{\text{random}} \) is the variance of a normal Gaussian distribution of phase errors that is the result
of the circuit and system thermal and device noises. The $\sigma_{\text{systematic}}$ is the RMS of the systematic phase errors due to non-linearity of phase interpolator and S/H, phase mismatch in quadrature LC VCO, jitter due to loop dynamics, and bang-bang frequency correction error. In order to test the chip, a clean reference clock, generated by signal generator, is divided by a power splitter. One output of the power splitter is used to trigger real-time oscilloscope and the other output is used as the reference for the test chip. Fig. 5.12 (b) and (c) show absolute jitter measurements, obtained by triggering the oscilloscope with the clean reference and measuring the phase jitter of the output clock. Fig. 5.12 (b) shows the histogram and total peak-to-peak jitter of the system to be 9ps and total RMS jitter to be 680fs. A real-time scope can distinguish Gaussian random jitter from the rest of jitter components. Fig. 5.12 (c) shows the measured breakdown of the jitter performance. The output clock has a random RMS jitter of 490fs and the periodic peak-to-peak jitter is measured to be 2.06ps. Fig. 5.12 (d) shows phase-noise measurements indicating -127dBc/Hz of phase noise at 1MHz offset. For the same reference frequency of 400MHz, jitter for three other multiplication factors (21, 22, and 23) have also been measured and are shown in Fig. 5.13.

In order to characterize different components of jitter, different test-structures on the chip have been measured separately to show different jitter contributions. Fig. 5.14 (a)-(d) shows simulation and measurement results of the quadrature LC VCO test-structure in room temperature. When the temperature is increased to 125°C, the simulated resonance frequency of the LC tank changes by less than 3.5%. The coarse-tuning range and steps are shown in Fig. 5.14 (d). The fine-tuning simulation verifies a range between 420-510MHz depending on frequency of operation. The bang-bang frequency corrections have steps of 1.6MHz-2.5MHz. Phase noise of the open-loop oscillator is measured to be -91.8dBc/Hz at 1MHz offset. The maximum quadrature phase mismatch of the quadrature LC VCO is measured to be 2.7%. This quadrature phase error appears both on phase-interpolator coefficients ($a$ and $b$) and corresponding clocks. For small enough phase mismatch

![Figure 5.13: Total RMS jitter of output clock versus frequency.](image-url)
(\(\Delta t \ll \frac{1}{f_{\text{ref}}}\)), the output amplitude error remains small and the subsequent CML-to-CMOS stage removes any amplitude degradation. Also, the quadrature phase error has no effect on zero-crossing and therefore systematic jitter of the output of the phase-interpolator. These conditions are shown as follows:

\[
PI_{\text{out}}(t) = b_{cr} \times CK_Q(t) - a \times CK_{I,cr}(t) = \sin(2\pi f(t_0 - \Delta t)) \cos(2\pi ft) - \\
\cos(2\pi ft_0) \sin(2\pi f(t_\Delta t)) = \cos(2\pi f\Delta) \sin(2\pi f(t - t_0)) \approx \sin(2\pi f(t - t_0)) \quad (5.12)
\]

\[
PI_{\text{out}}(t_0) = \cos(2\pi f\Delta) \sin(2\pi f(t_0 - t_0)) = 0. \quad (5.13)
\]
A similar argument can be used to show that amplitude mismatch in the quadrature LC VCO does not affect the performance of phase interpolation either, as it appears in both the samples and the quadrature waveform as follows:

\[
PI_{\text{out}}(t) = b_{\text{er}} \times CK_Q(t) - a \times CK_{I,\text{err}}(t) = A_{I,\text{err}} \sin(2\pi ft_0) A_Q \cos(2\pi ft) - A_Q \cos(2\pi ft_0) A_{I,\text{err}} \sin(2\pi ft) = A_{I,\text{err}} A_Q \sin(2\pi f(t - t_0)).
\]  
(5.14)

\[
PI_{\text{out}}(t_0) = A_{I,\text{err}} A_Q \sin(2\pi f(t_0 - t_0)) = 0.
\]  
(5.15)

The quadrature phase mismatch \(\phi_{\text{er}}\) of the LC VCO output translates to quadrature phase mismatch of the samples taken from \(CK_I\) and \(CK_Q\):

\[
CK_{I,S} = \cos(2\pi f_{VCO} t_n + \phi_0) = \cos(\frac{2\pi n \Delta f}{f_{\text{ref}}} + \phi_0)
\]  
(5.16)

\[
CK_{Q,S,\text{err}} = \sin(2\pi f_{VCO} t_n + \phi_0 + \phi_{\text{er}}) = \cos(\frac{2\pi n \Delta f}{f_{\text{ref}}} + \phi_0 + \phi_{\text{er}} \pm \frac{\pi}{2}).
\]  
(5.17)

Also, any mismatch in the delay of \(I_{S,\text{dig}}\) and \(Q_{S,\text{dig}}\) and duty-cycle distortion (DCD) in their waveforms further degrades their quadrature characteristic required for operation of the fine-tuning. These non-idealities can be summed up in \(T_{\text{er}}\), defined in

\[
|T_{\text{er}}| = \frac{\phi_{\text{er}}}{2\pi \times \Delta f} + \frac{1}{4} \frac{D_I}{2\pi \times \Delta f} + \frac{1}{2} \frac{D_Q}{2\pi \times \Delta f} + |T_{\Delta}|.
\]  
(5.18)

\(D_I\) and \(D_Q\) are DCD of \(I_{S,\text{dig}}\) and \(Q_{S,\text{dig}}\), respectively, and \(T_{\Delta}\) is their delay mismatch. Fine-tuning can only operate as long as rising edges of \(I_{S,\text{dig}}\) and \(Q_{S,\text{dig}}\) (shown in Fig. 5.7) are distinguishable by flip-flops used in fine tuning Fig. 5.9. Distinguishability by flip-flops is determined by the following condition:

\[
\frac{1}{4 \times |\Delta f|} - T_{\text{er}} > \max(t_{su}, t_h),
\]  
(5.19)

where \(t_{su}\) and \(t_h\) are setup time and hold time of the flip-flops. The maximum pull-in range is limited by the non-idealities of circuit implementation, as described in equation 5.19. As the system gets close to lock and \(\Delta f\) gets smaller, the time-domain difference between rising edges of \(I_{S,\text{dig}}\) and \(Q_{S,\text{dig}}\) get easier to detect. The delay mismatch in UP/DN bang-bang pulses has no effect on the loop operation near lock as correction pulses are independent and do not rely on UP/DN overlap. Bang-bang UP/DN corrections of the fine-tuning block cause a small error in VCOs frequency. Note that as the rising edge of the reference is injected to the output, the output clock will have zero average frequency error. However, as the injection occurs at a fraction of the VCO’s frequency, error in VCO’s frequency causes a periodic shift in VCOs’ phase compared with output phase, which leads to deterministic jitter and reference spurs. When the system is at lock, the difference between
UP/DN pulses causes a frequency error $\Delta f_{err}$. The periodic phase shift in VCO’s phase compared with the output phase can be written as

$$\Delta T_{err} = \frac{1}{f_{ref}} - \frac{N}{N f_{ref} - \Delta f_{err}} \approx \frac{\Delta f_{err}}{N f_{ref}^2}.$$

(5.20)

Monte Carlo simulations show worst-case $\Delta f_{err}$ to be around 0.5MHz. From 5.20, for a multiplication factor of 20 and reference clock frequency of 400MHz, $\Delta T_{err}$ is 0.16ps. This value of deterministic jitter gets larger with larger $N$ or smaller reference frequency.

Fig. 5.15 shows measured non-linear characteristics of the S/H and phase interpolator. An off-chip reference clock and quadrature waveforms have been applied to a replica test structure that only contains phase interpolator and S/H. Shifting the relative phase of the reference and quadrature signals, the change in output phase has been measured. Maximum deviation from the ideal line is measured to be 4.7. This deviation means the output clock phase will not be completely independent of the VCO phase and the reference injection induces a systematic jitter. For an output clock of 8GHz the peak-to-peak systematic jitter associated with this non-linearity is 1.63ps. As can be seen, the measured 2.06ps periodic peak-to-peak jitter is largely associated with this non-linear effect.

Fig. 5.16 (a) shows the measured power consumption and energy efficiency of the frequency synthesizer while operating at different frequencies. Energy efficiency of the frequency synthesizer is 0.312mW/GHz at 8GHz, which drops to 0.305mW/GHz at 9.5GHz. As the frequency of operation increases, power consumption increases linearly primarily due to the power consumption of digital elements. The power consumption measurement includes LC quadrature oscillator, sample-and-holds, phase interpolator, coarse-tuning, fine-tuning, and output CML-to-CMOS converter. Fig. 5.16 (b) shows the power breakdown of the frequency synthesizer at 8GHz. More than half of the
power is consumed by quadrature LC VCO.

Fig. 5.17 shows the chip micrograph. The frequency synthesizer occupies an active area of 0.044mm$^2$, including the inductors. The octagonal structure of the inductor design, the digital coarse tuning block layout, and placement of the rest of the circuits between the two inductors can also be seen in subfigures A, B, and C of Fig. 5.17, respectively. Table 5.1 summarizes performance of the frequency synthesizer and compares it to the state-of-the-art. While this table provides a glimpse at how different design topologies perform in terms of jitter, power, and reference spurs, the noise environment and reference quality play an important role in final output jitter. The proposed architecture is a low-power alternative for on-chip frequency synthesizers limiting jitter accumulation to one reference period.

Table 5.1: Frequency Synthesizer Performance Summary

<table>
<thead>
<tr>
<th>Technology</th>
<th>65nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency [GHz]</td>
<td>8-9.5</td>
</tr>
<tr>
<td>Reference [MHz]</td>
<td>400-1000</td>
</tr>
<tr>
<td>Reference Spurs [dBc]</td>
<td>-64.3</td>
</tr>
<tr>
<td>Integrated Jitter/RMS [ps]</td>
<td>0.49</td>
</tr>
<tr>
<td>Total Jitter RMS/PP [ps]</td>
<td>0.68/9.11</td>
</tr>
<tr>
<td>Power [mW/GHz]</td>
<td>0.312</td>
</tr>
<tr>
<td>Area [mm$^2$]</td>
<td>0.044</td>
</tr>
</tbody>
</table>
Figure 5.17: Die micrograph and layout details of the implemented prototype. (a) Design of inductor. (b) Digital coarse-tuning block. (c) Placement of the rest of the circuitry between the inductors.

5.3.4 Analysis

Fig. 5.18 shows the block diagram of the system, including frequency acquisition, reference injection and the effect of reference injection on phase noise in the simplified case of noise-less reference. The system is modeled when the coarse-tuning code word is settled and the system is within fine-tuning acquisition range. First, we will investigate the effect of reference injection on filtering the phase noise of the frequency-locked loop (FLL). Reference injection is modeled as shifting the phase of the FLL output by the difference between periodic samples of the FLL output phase and reference clock phase. In frequency domain, this translates to convolution of an impulse train with the phase noise spectrum of the FLL output followed by a sinc function, shown as

\[
\phi_{\text{out}}(j\omega) = \phi_{\text{FLL}}(j\omega) + \phi_{\text{err}}(j\omega) = \phi_{\text{FLL}}(j\omega) - \left\{ \sum_{k=\infty}^{k=\infty} \phi_{\text{FLL}}(j\omega - \frac{2\pi k}{T_{\text{ref}}}) \right\} \times T_{\text{ref}} \text{sinc}(T_{\text{ref}}(j\omega)).
\]  

(5.21)
Due to the shape of the sinc function, reference injection filters out most of the low-frequency components of the noise (Fig. 5.18). Another interesting case is when the reference is not clean. In that case there is a direct path for reference noise to appear at the output. The impulse train, by which the FLL output noise is convolved, will have the corresponding jitter and the sinc function will also be distorted depending on the shape of the reference clock’s phase noise. Nonetheless, the low-frequency phase noise of the FLL output will still get filtered (now less effectively) as long as the reference jitter is small. In the presence of a noisy reference, the FLL output will carry low-frequency components of the reference phase noise (high-frequency components get filtered out in the loop). The following equation shows the relation between reference clock and VCO frequencies:

\[
\frac{f_{VCO}}{f_{ref}} = \frac{K_{CP} K_{VCO}}{s + \frac{K_{CP} K_{VCO}}{M}} = \frac{K_{CP} K_{VCO}}{s + \frac{K_{CP} K_{VCO}}{M}} = \frac{\omega_1}{s + \frac{\omega_1}{M}}.
\]

(5.22)

where \(M\) is the multiplication factor and \(\omega_1/M\) represents the loop bandwidth. In order to verify the loop dynamics of the system, a step frequency is applied to the reference, changing the reference from 400MHz to 403.5MHz. As the loop corrects the output frequency, the output clock is measured in real time and its frequency is measured over 300ns. Fig. 5.19 shows the measured loop response, the simulated loop response, and the curve fitted with the measurement results. The fitted first order step response time constant is 28ns (showing 5.7MHz loop bandwidth).
Figure 5.19: Loop response measured by applying a step frequency to the reference clock.
5.4 Summary

In this chapter we reviewed clock generation and different topologies previously used for on-chip clock generation. We presented a novel first-order frequency synthesizer. In this architecture, injection of the rising edge of the clock limits jitter accumulation to one reference cycle, and first-order dynamic of the system ensures acquisition without stability concerns. Reference injection is implemented via phase-interpolation. The frequency acquisition consists of digital coarse-tuning and rotational frequency detection for fine-tuning. A prototype has been implemented in a 65nm CMOS process, and has been fully verified via measurements and simulations. Experimental results validated functionality of the system with an electrical input clock. The total active area including inductors is 0.044mm². The test chip operates in the range of 8-9.5GHz. At 8GHz, with a multiplication factor of 20, it consumes 2.49mW, and exhibits 490fs RMS integrated jitter and 2.06ps peak-to-peak periodic jitter. At this operating frequency, the reference spurs are measured to be 64.3dB below the carrier frequency. The first order characteristics of the frequency acquisition have been examined and demonstrated via measurement. Different factors contributing to the jitter have been analyzed and supporting measurements have been presented. This architecture is well suited for dense parallel links and noisy environments where a clean reference clock is available.
Chapter 6

Conclusion

Optical interconnects are being driven by two major forces in semiconductor industry. On the computing side, the quest for extending our processing power has resulted in increasing the number of cores, which results in an ever increasing demand for higher interconnect bandwidth density. On the networking side, hundreds of thousands of servers in data centers require any-to-any server communication that entails an immense network interconnected through a vast array of CMOS switching chips. A vast majority of these CMOS switching chips are now connected through fibers using optical transceiver modules but aspects of current fiber-optic transport technologies, such as bulky form-factor and intrinsic incompatibility for integration with electronics, make them unscaleable as networking traffic grows. Co-design and co-integration of electronics and photonics provide a unique opportunity for entirely new optical interconnect architectures for computing and networking applications. In this light, as the system involves many trade-offs that are intertwined between electronics and photonics, a hollistic design approach is nessesary for optimal optical interconnect design. In this dissertation, we looked at primary building blocks of optical interconnects with an emphasis on their hollistic design.

In the first part of this dissertation we described a 3D-integrated CMOS-Silicon photonic optical receiver that employs a novel integrating low-bandwidth TIA front end. The 3D integration is based on Copper Pillar flipchip technology, allowing low parasitic capacitance and 40µm pitch for interconnection. A measured prototype achieves -15dBm of sensitivity and 170fJ/bit of energy efficiency at 25Gb/s. We study different trade-offs in designing an optical receiver and how to choose between a full-bandwidth TIA front-end and integrating architecture using a resistive front-end or a low-bandwidth TIA front-end. The design methodology is supported by measurements of two 3D-integrated prototypes based on a conventional TIA and a double-sampling integrating receiver.

In a follow-up work, a 3D-integrated parallel optical receiver is presented. The receiver’s source-synchronous clocking scheme is based on a quadrature-locked loop which generates accurate clock phases for a 4-channel optical receiver using a forwarded clock at quarter-rate. We presented an adaptive body biasing circuit to maintain the per-bit energy consumption across wide data-rates.
The prototype measurements showed the lowest reported power consumption of 4.87mW per-channel at 32Gb/s (153fJ/bit). The receiver sensitivity was measured to be -8.8dBm at 32Gb/s.

In the second part, we introduced three solutions for primary challenges associated with implementation of micro-ring modulator-based transmitters. We presented a novel ring-modulator structure, a differential ring modulator (DRM), which breaks the optical bandwidth/quality factor trade-off that limits the speed of high-Q ring modulators. The DRM maintains a constant level of energy stored in the ring and does not suffer from power droop when long sequences of identical bits are transmitted. A prototype has been fabricated and measured to operate up to 10Gb/s. The speed of this prototype is primarily limited by RC bandwidth limitation of contacts. Next, we presented a scheme for thermal stabilization of micro-ring resonator modulators that has become feasible by co-design and co-integration of electronics and photonics. The wavelength stabilization works by direct measurement of ring temperature using a PTAT temperature sensor on the silicon photonic chip. The measured temperature is used in a feedback loop to adjust the thermal tuner of the ring. In this scheme, there is no need for constantly tapping a portion of output optical power for monitoring or complex circuitry. The closed-loop feedback system is demonstrated to operate in the presence of thermal perturbations at 20Gb/s. Finally, we present a CMOS-silicon photonic optical transmitter based on carrier-injection ring modulators. Carrier-injection mode modulators required that voltage swing is CMOS compatible and lower than carrier-depletion mode modulators, but they have intrinsically lower bandwidth. This design features a new low-power switched-capacitor-based pre-emphasis that effectively compensates for the modulator bandwidth limitation. A feedforward wavelength stabilization technique via direct measurement of ring temperature using a monolithic PTAT sensor is also presented. The optical transmitter achieves energy efficiency of 342fJ/bit at 10Gb/s and the wavelength stabilization circuit consumes 0.29mW.

Finally, as the last contribution of this dissertation, we presented a first-order frequency synthesizer that is suitable for high-speed on-chip clock generation for electrical and optical transceivers. The proposed design featured an architecture combining an LC quadrature VCO, two sample-and-holds, a PI, digital coarse-tuning, and rotational frequency detection for fine-tuning. Unlike multiplying DLLs this architecture can use LC oscillators with lower phase noise and does not rely on opening an aperture for reference clock injection. This new architecture is suitable for implementation in noisy environments such as large SoCs applications, where the digital switching noise propagates through the substrate and power distribution networks. In such systems, unlike multiplying PLLs, this frequency synthesizer cleans the accumulated jitter every period of the reference clock. We tested the prototype chip using an electrical reference clock but the prototype chip was capable of receiving a low jitter optical reference clock generated by a high-repetition-rate mode-locked laser. The output clock at 8GHz has an integrated RMS jitter of 490fs, peak-to-peak periodic jitter of 2.06ps, and total RMS jitter of 680fs. The reference spurs are measured to be 64.3dB below the
carrier frequency. At 8GHz the system consumes 2.49mW from a 1V supply. When electronics and photonics are closely integrated they provide a great promise for improving interconnect performance and reduce cost. Hollistic design of co-integrated optical interconnects provides a unique opportunity to design entirely new architectures and bring the performance of current systems to unprecedented levels. This dissertation presented such breakthroughs in optical links by electronic-photonic co-design of the primary elements of an optical link: receivers, transmitters, and clocking.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWGN</td>
<td>Additive white Gaussian noise</td>
</tr>
<tr>
<td>BB</td>
<td>Body Biasing</td>
</tr>
<tr>
<td>BER</td>
<td>Bit error rate</td>
</tr>
<tr>
<td>BOX</td>
<td>Buried oxide layer</td>
</tr>
<tr>
<td>CDR</td>
<td>Clock-data recovery</td>
</tr>
<tr>
<td>CMOS</td>
<td>Complementary metal-oxide-semiconductor</td>
</tr>
<tr>
<td>DJ</td>
<td>Deterministic jitter</td>
</tr>
<tr>
<td>DFE</td>
<td>Decision-feedback equalizer</td>
</tr>
<tr>
<td>DLL</td>
<td>Delay-locked loop</td>
</tr>
<tr>
<td>DRM</td>
<td>Differential ring modulator</td>
</tr>
<tr>
<td>DOM</td>
<td>Dynamic offset modulation</td>
</tr>
<tr>
<td>EC</td>
<td>Embedded clock</td>
</tr>
<tr>
<td>ER</td>
<td>Extinction ratio</td>
</tr>
<tr>
<td>FBW</td>
<td>Full bandwidth</td>
</tr>
<tr>
<td>FC</td>
<td>Forwards clock</td>
</tr>
<tr>
<td>FDSOI</td>
<td>Fully depleted silicon on insulator</td>
</tr>
<tr>
<td>FIR</td>
<td>Finite impulse response</td>
</tr>
<tr>
<td>FLL</td>
<td>Frequency locked loop</td>
</tr>
<tr>
<td>FOM</td>
<td>Figure of merit</td>
</tr>
<tr>
<td>FPGA</td>
<td>Field-programmable gate array</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Gb/s</td>
<td>Gigabit-per-second</td>
</tr>
<tr>
<td>IC</td>
<td>Integrated circuit</td>
</tr>
<tr>
<td>IO</td>
<td>Input/output</td>
</tr>
<tr>
<td>IL</td>
<td>Injection locking</td>
</tr>
<tr>
<td>ILO</td>
<td>Injection locked loop</td>
</tr>
<tr>
<td>IL PLL</td>
<td>Injection locked phase locked loop</td>
</tr>
<tr>
<td>ISI</td>
<td>Inter symbol interference</td>
</tr>
<tr>
<td>LBW</td>
<td>Low bandwidth</td>
</tr>
<tr>
<td>LPF</td>
<td>Low-pass filter</td>
</tr>
<tr>
<td>LTI</td>
<td>Linear time invariant</td>
</tr>
<tr>
<td>MQPE</td>
<td>Mean quadrature phase error</td>
</tr>
<tr>
<td>MRM</td>
<td>Micro-ring modulator</td>
</tr>
<tr>
<td>MZI</td>
<td>Mach zehnder interferometer</td>
</tr>
<tr>
<td>MZM</td>
<td>Mach zehnder modulator</td>
</tr>
<tr>
<td>PFD</td>
<td>Phase frequency detector</td>
</tr>
<tr>
<td>PI</td>
<td>Phase interpolator</td>
</tr>
<tr>
<td>PLL</td>
<td>Phase-locked loop</td>
</tr>
<tr>
<td>PRBS</td>
<td>Pseudo-random bit sequence</td>
</tr>
<tr>
<td>PTAT</td>
<td>Proportional to absolute temperature</td>
</tr>
<tr>
<td>PVT</td>
<td>Process, voltage, temperature</td>
</tr>
<tr>
<td>QLL</td>
<td>Quadrature locked loop</td>
</tr>
<tr>
<td>RJ</td>
<td>Random jitter</td>
</tr>
<tr>
<td>RO</td>
<td>Ring oscillator</td>
</tr>
<tr>
<td>RMS</td>
<td>Root mean square</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-noise ratio</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>TIE</td>
<td>Time interval error</td>
</tr>
<tr>
<td>UI</td>
<td>Unit interval (one bit-time in a data stream)</td>
</tr>
<tr>
<td>VCDL</td>
<td>Voltage-controlled delay line</td>
</tr>
<tr>
<td>VCO</td>
<td>Voltage-controlled oscillator</td>
</tr>
<tr>
<td>VCSEL</td>
<td>Vertical-cavity surface-emitting laser</td>
</tr>
<tr>
<td>WDM</td>
<td>Wavelength division multiplexing</td>
</tr>
</tbody>
</table>
Bibliography


[4] Saman Saeedi, Sylvie Menezo, and Azita Emami. A 25gbps 3d-integrated cmos/silicon pho-
tonic optical receiver with 15dbm sensitivity and 0.17 pJ/bit energy efficiency. In Optical

Chemical Heritage Foundation, 2006.

[6] Xuezhe Zheng, Frankie Y Liu, Jon Lexau, Dinesh Patil, Guoliang Li, Ying Luo, Hiren D
Thacker, Ivan Shubin, Jin Yao, Kannan Raj, et al. Ultralow power 80 Gb/s arrayed cmos silicon
2012.


[8] Sylvie Menezo, Gabriel Pares, Stephane Bernabe, Olivier Castany, Corrado Sciancalepore,
Karim Hassan, Benjamin Blampey, Benoit Charbonnier, Julie Harduin, Sonia Messaoudene,
et al. (keynote) silicon photonics technology for optical communications with high bandwidth
density requirements (1tbit/s and 1,000 gbit/s/cm²). In Meeting Abstracts, number 29, pages

[9] Azita Emami-Neyestanak, Meisam Hoarvar Nazari, and Saman Saeedi. Double-sampling re-
ciever with dynamic offset modulation for optical and electrical signaling, May 6 2013. US


123


[125] Yu-Li Hsueh, Lan-Chou Cho, Chih-Hsien Shen, Yi-Chien Tsai, Tzu-Chan Chueh, Tao-Yao Chang, Jui-Lin Hsu, and Jing-Hong Conan Zhan. 28.2 a 0.29 mm 2 frequency synthesizer in


