

DESIGN OF CMOS RECEIVERS FOR PARALLEL
OPTICAL INTERCONNECTS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Azita Emami-Neyestanak
August 2004

© Copyright by Azita Emami-Neyestanak 2004
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Mark A. Horowitz
(Principal Adviser)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

David A. B. Miller

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Bruce A. Wooley

Approved for the University Committee on Graduate Studies.

Abstract

The growing demand for high-bandwidth communication between integrated circuit chips calls for large numbers of high-speed inputs and outputs (IOs) per chip. IO data rates have increased to the point where electrical signaling is now limited by the channel properties. In order to achieve multi-Gb/s data rates, complex designs that equalize the channel are necessary.

Using optics for chip-to-chip interconnections is promising since the optical channel dispersion and cross-talk are small. In this work we demonstrate the possibility of building small and low-power optical receivers that facilitate large numbers of IOs. A new double sampling/integrating front-end is proposed and implemented. Unlike prior designs, this receiver removes the need for a gain stage that runs at the data rate, making it suitable for low-power implementations. This front-end allows a time-division multiplexing technique to support very high data rates. The dynamic range of the integrating input node can be improved by a proposed decision-directed common-mode control loop, which reduces the dependency of the dynamic range on the power supply voltage.

The required receive clock can be generated in many ways. While the standard oversampled clock recovery is possible, it needs extra clock phases in the middle of main data samples. In order to reduce the power, a baud rate clock recovery technique is proposed and implemented as part of a transceiver array test-chip. The resulting transceiver consumes less than 150mW per channel at 5.0Gb/s in a 0.25 μ m CMOS technology. If projected to a 90nm CMOS technology, 15Gb/s data rate and 30mW power per IO are possible, which allow more than 10Tb/s chip-to-chip bandwidth, with up to one thousand IOs per chip.

Acknowledgements

During my graduate studies in Stanford University, I had the privilege of learning from the best teachers and receiving support from the most caring friends. This work would not have been possible without any of the two.

First of all, I would like to thank my Ph.D. advisor Professor Mark Horowitz. Mark's knowledge, vision and personality make him the greatest advisor and teacher one could ever wish for. Every single meeting with Mark helped me to go another step forward in my research, and motivated me to learn more. Mark's enthusiasm and care were the driving forces for creation of new ideas in this work. He also effectively helped me to improve my public speaking and writing skills. I sincerely thank him for his kindness and patience.

I gratefully acknowledge my co-advisor Professor David Miller for his invaluable advice and help. Professor Miller was the one who initiated this research and inspired this project as a collaboration between the two research groups. Working in his laboratory and with his students was a great opportunity for me to expand my knowledge to the area of photonics. I also would like to thank him for being a member of my oral defence and reading committees.

I would like to extend my gratitude to Professor Bruce Wooley for his advice, encouragements and help throughout my years at Stanford, and for serving as members of my oral defence and reading committees. I sincerely appreciate his kindness.

Faculty of Electrical Engineering department at Stanford are among the most brilliant teachers in the world and I had the opportunity of learning from many of them. Particularly, I would like to express my gratitude to Professor Fabian Pease, chair of my orals committee and Professor Tom Lee, who both helped me during my

first year at Stanford.

With no doubt, a great aspect of being one of Mark's students was to be in a friendly and highly cooperative research group. This thesis builds upon the work done by many former members of the Horowitz group. I would like to thank Samuel Palermo, Hae-Chang Lee and Elad Alon for helping in the design of my chip. I also thank Jaeha Kim and Dean Liu for their technical help and for answering my numerous questions. I am also grateful for technical discussions with Vladimir Stojanovic, Ken Mai, Ron Ho, Kun-Yung Chang, Ken Yang and Bill Ellersik. I appreciate the friendship of other former and current students in the Horowitz group, Bennett Wilburn, Michal Smulski, Evelina Yeung, Gu-Yeon Wei, Amin Firoozshahian, Francois Labonte, Alex Solomatnikov, Vicky Wong and Dinesh Patil.

Members of Professor Miller's group greatly helped me in the optical testing of my chips. I would like to thank Aparna Bhatnagar, Gordon Keeler, Noah Helman and Diwakar Agarwal for the integration of optical devices, lab set-ups and technical discussions.

This work was possible with the generous support from National Semiconductor, Vitesse Semiconductor, DARPA and MRCO IFC. I also would like to thank CIS staff members, computer administrators and Mark's administrators, Teresa Lynn, Penny Chumley and Taru Fisher for creating an amazing work environment for us. I sincerely thank CIS former students Hiran Samavati, Joel Dawson and Lalit Nathawad who took the responsibility of National tape out.

The highly academic environment at Sharif University of Technology helped me to build the required background in engineering and encouraged me to pursue my graduate studies abroad. I sincerely thank my undergraduate advisors Professor Sharif-Bakhtiyar, Professor Fotovat, and my teacher Professor Jahanbeglo.

My first exposure to science and engineering goes back to my high-school years. I was extremely fortunate to go to Farzanegan, with the most dedicated and caring teachers. I would like to thank, Ms. Poorsaeed, Mr. Niusha, Mr. Helli, Ms. Mokhtari, Ms. Rohani, Mr. Kazemi and Mrs Haerizadeh.

I would like to thank my friends Valeria Bertacco, Vace Shakoori, Fatemeh Jalayer, Yasamin Mostofi, Mahmood Reza Kasnavi, Dara Ghahremani, Farid Nemati, Ali

Hajimiri, Ramin Farjad-rad, Amy Droitcour, Ardavan Maleki, Mina Matin, Parisa Gholami and Nogol Rashidi for making Stanford a fun place to live and work.

I sincerely thank my best friend Kaveh Hosseini who helped me to go forward in every stage of my graduate studies. His encouragements and friendship brought peace and happiness to my life.

My brother Sohrab has been an amazing mentor. He always inspired and motivated me to do my best. All these years, my lovely sisters Maryam and Mitra eased the hardship of being away from home by their love, beautiful gifts and letters. My deepest love and gratitude go to these three.

I dedicate this thesis to my parents, Mr. Akbar Emami-Neyestanak and Mrs. Khatoon Hadavi-Neyestanaki for their lifelong efforts to provide the best for me. I deeply appreciate their endless love, support and sacrifices.

At the end, I cherish the memory of Masoomeh Hadavi and Zhila Asghari.

Contents

Abstract	iv
Acknowledgements	v
1 Introduction	1
1.1 Organization	3
2 Background	7
2.1 Electrical Link Basics	7
2.1.1 Channel Termination	10
2.1.2 Transmitter Design	12
2.1.3 Receiver Design	13
2.1.4 Time Division Multiplexing	14
2.2 Clock Generation and Recovery Loops	15
2.2.1 Phase-Locked Loop	16
2.2.2 Voltage and Timing Margins	18
2.2.3 Clock Recovery	19
2.3 Channel	24
2.3.1 Equalization	27
2.4 Summary	31
3 Receiver Design for Optical Interconnects	33
3.1 High-Speed Optical Interconnects Overview	34
3.1.1 Photodetectors	36

3.2	Prior Art in Design of Optical Receiver Front-Ends	38
3.2.1	Front-End Design Challenges	38
3.2.2	Transimpedance Amplifiers	40
3.2.3	TIA Design Consideration	46
3.2.4	Integrating Front-Ends	48
3.3	Double Sampling/Integrating Front-End	52
3.3.1	Receiver Design Overview	52
3.3.2	Sampling and Comparison	55
3.3.3	Filter and Current Feedback	60
3.3.4	Supporting Circuits	62
3.3.5	Performance Analysis	64
3.4	Receiver Testing	68
3.5	Results and Performance Comparison	70
3.6	Summary	73
4	Scaling and Common-Mode Control	75
4.1	Double Sampling Front-End Scaling	76
4.2	Decision Directed Current Control	79
4.3	TIA Scaling	84
4.4	Summary	86
5	Clock Generation and Timing Recovery	87
5.1	2X OverSampled Clock Recovery	88
5.2	Baud Rate Clock Recovery	89
5.2.1	Principles of Baud Rate Clock Recovery	90
5.2.2	Phase Detector Performance Analysis	92
5.3	Optical Transceiver Test-Chip	94
5.3.1	Clocking High-Level Architecture	95
5.3.2	CDR Building Blocks	97
5.4	Experimental Results	104
5.5	Design Improvements	106
5.6	Summary	110

6	Conclusions	112
A	TIA Analysis	115
	Bibliography	119

List of Tables

3.1	Chip performance summary	72
5.1	4-bit patterns with phase information for baud rate CDR	92
5.2	Chip performance summary	105

List of Figures

1.1	Parallel optical chip-to-chip interconnection over free space	3
1.2	Array of GaAs optical devices, flip-chip bonded to the silicon CMOS chip	4
2.1	Components of a basic electrical link	8
2.2	Plesiochronous serial link	9
2.3	Parallel source-synchronous link	10
2.4	Electrical signaling (a) with no termination, (b) properly terminated .	11
2.5	Data transmission schemes (a) low-impedance, voltage mode drive, (b) high-impedance, current mode drive	12
2.6	Current mode driver	13
2.7	Data resolution at the receiver using a slicer	14
2.8	Time division multiplexing and parallelism	15
2.9	Phase-locked loops (a) a VCO-based PLL, (b) a VCDL-based DLL .	17
2.10	Receiver voltage and timing margins	19
2.11	Timing recovery from the received data	20
2.12	2x over-sampled CDR with two clock phases for data and phase resolution	21
2.13	2x over-sampled phase detection, loop error signal and implementation	22
2.14	Channel pulse response and criteria for baud rate recovery	24
2.15	Signal path and channel components in a typical backplane	25
2.16	Typical frequency transfer function of a backplane channel	26
2.17	Dispersion in a band-limited channel	26
2.18	Signaling methods (a) single-ended signaling, (b) differential signaling	27
2.19	Linear equalization, flattens the frequency response	28

2.20	Linear equalization at the receiver with FIR filtering	29
2.21	Linear equalization at the transmitter with FIR filtering	29
2.22	Decision feedback equalization	30
3.1	Synchronous optical transmission over free space	36
3.2	Equivalent electrical model of a reverse-biased photodiode	37
3.3	Simplified model of a photodiode	39
3.4	Simple resistive optical front-end	40
3.5	Common-gate TIA and its noise sources	41
3.6	Regulated cascode TIA input satge	42
3.7	Shunt-shunt feedback TIA with limiting amplifiers	44
3.8	Shunt-shunt resistive-feedback TIA model	44
3.9	Common-source shunt-shunt feedback TIA designs	45
3.10	Capacitive network feedback TIA	46
3.11	Inductive peaking for limiting amplifiers	49
3.12	Receiver-less front-end with totem-pole and clamp diodes	50
3.13	Sense-amp-based front-end	51
3.14	Block diagram of double sampling/integrating front-end	53
3.15	Data resolution and input voltage waveform of the integrating front-end	54
3.16	Double sampling/integrating front-end implementation	54
3.17	De-multiplexing double sampling/integrating front-end	55
3.18	Double sampler and comparator circuits	56
3.19	Rise-time of the NMOS sampler in response to a 10mV step voltage as a function of input voltage level	57
3.20	Output voltage swing of the sampler with a 10mV step voltage at the input as a function of input voltage level	58
3.21	PMOS adjustable capacitors for offset compensation	59
3.22	Input-referred offset as function control signals and input common- mode voltage	60
3.23	Feedback loop with low-pass-filter to adjust I_{DC}	61
3.24	I_{DC} variations with the DC-balance range of input data	62

3.25	Voltage to current conversion for I_{DC} loop	63
3.26	Comparator and following stages of sense-amp and SR latch	63
3.27	Chopping the clock for 1:2 de-multiplexing	64
3.28	Multi-phase clocking, which automatically allows time for comparison	65
3.29	Sampler and half circuit of the StrongArm latch comparator	66
3.30	Required optical energy per bit versus C_s , C_i and C_p assuming $n = 5$, $R = 0.5A/W$, $SNR=36$ (BER = 10^{-10}), and $A_c=1$	68
3.31	Analog sampler for monitoring the input voltage waveform	69
3.32	Effect of sending unbalanced input data stream to the receiver, (a) balanced input data (b) unbalanced input data, $\Delta V_b = 2\Delta V_{b0}$	70
3.33	Top level block diagram of the first receiver test-chip	71
3.34	Fabricated test-chip with the flip-chip bonded devices	72
4.1	Block diagram of decision-directed common-mode control	80
4.2	Input voltage waveform with the decision-directed current control	80
4.3	Data resolution for DDCC by adding offset, four possible cases for different values of $D[n]$ and $D[n-m]$ are shown	81
4.4	Simulated loop dynamic for DDCC	82
4.5	Simulated loop dynamic for DDCC with comparator noise and offset, $\sigma_n = 1$, offset=3	83
5.1	2x-oversampled phase deetction for the integrating front-end	89
5.2	Integrating input waveform and baud rate phase detection	91
5.3	Samplers and comparators for baud rate CDR	91
5.4	Bangbang CDR loop architecture used for the performance analysis	93
5.5	Percentage of phase correction commands vs. phase misalignment for the 2x-oversampled and baud rate techniques	94
5.6	Percentage of phase correction commands after a majority filtering over every 5 bits	95
5.7	Optical transceiver chip block diagram	96
5.8	Multiphase clock generation for the transmitter and clocked integrating front-end	97

5.9	Coupled ring-oscillator [1]	99
5.10	Phase-frequency detector for the global PLL [2]	99
5.11	Charge pump and loop filter	100
5.12	Linear voltage regulator for low noise VCO voltage control	101
5.13	Phase and pattern detector for baud rate CDR	102
5.14	Receiver VCO fine control, integral and proportional gains	103
5.15	Receiver PLL charge-pump	103
5.16	Transceiver test-chip for parallel optical interconnection	105
5.17	Receiver recovered clock signal	106
5.18	Simulated jitter and measured jitter vs. input voltage	107
5.19	Jitter vs. input voltage swing with low loop gain and offset	108
5.20	Dual Loop Clock and Data Recovery Loop	109
A.1	Shunt-shunt resistive feedback TIA model	116

Chapter 1

Introduction

In most applications today, integrated circuit (IC) chips need to communicate with many other ICs or modules in the system [3] [4]. The increasing speed of on-chip data processing and computation creates a growing demand for high-bandwidth input and output (IO) on these chips [5]. The required bandwidth is achieved by both increasing the signalling rate of each IO pin and increasing the number of IO pins on the chip.

Until recently, technology scaling facilitated faster transceiver circuits and on-chip clocking, allowing IO rates to scale with the technology [6] [7]. Unfortunately, the nature of the IO design problem has changed. Today internal circuits can run at 10's of Gb/s, but the performance of the link is limited by the characteristics of the channel - the electrical path from one die to the other. In order to achieve desired data rates over existing channels, many multi-Gb/s links use complex signal processing to get around the channel limitations [8].

Instead of continuing to increase the links complexity, a different approach for scaling performance of IOs is to change the signaling method and the channel media. The electrical loss in the copper wires has long motivated the use of optical fiber communication for data transmission over long distances. Reduced number of amplifiers in the signal path, higher bandwidth, improved signal-to-noise-ratio (SNR) and effectively lower cost have made optical fiber communication the favored choice for links longer than 10m.

The possibility of using optics for interconnection at short distances recently has

been a subject of considerable research and analysis [9] [10]. By providing a high-capacity channel, optical signaling can potentially close the gap between the interconnect speed and on-chip data processing speed. This dissertation investigates the challenges of designing electronics for short-haul optical links and proposes a number of solutions to enable optical IOs. We focus on techniques to design simple, small and low power receivers suitable for dense parallel optical interconnects. In order to achieve low power and area a novel receiver front-end using a double sampling/integrating technique is proposed and the supporting circuits are presented. This design facilitates a number of interesting solutions such as parallelism and demultiplexing, an efficient baud-rate clock and data recovery (CDR) and a decision directed common-mode control technique to enhance the dynamic range of the receiver.

Although optical signaling involves electrical-to-optical (EO) conversion and vice versa (OE), its large channel bandwidth can simplify the design of the transceiver electronics. The frequency dependent loss and dispersion in optical signaling over short distances are negligible and the channel itself can support very high data rates [11]. The maximum data rate of an optical link is in fact limited by the performance of the optical devices and the speed of on-chip electronics. Although the data rate per channel might not be significantly higher than electrical signaling, the overall design might have lower power allowing more IOs built on die. In many systems, the maximum power consumption and area are among the limiting factors for the number of IOs possible on-chip. Moreover for parallel optical signaling at short distances, one can either use fiber-bundles or free space to send collimated beams in parallel from one chip to the other. In both situations the cross-talk among the beams is negligible, avoiding another problem with large numbers of electrical IOs. Note that optical signaling does not have impedance matching and pin allocation restrictions either. A three dimensional configuration for parallel optical interconnection in free space is shown in Figure 1.1. The integration of dense two dimensional (2D) arrays of optical devices with standard complementary-metal-oxide-semiconductor (CMOS) ICs has been demonstrated, and allows a huge chip-to-chip interconnection bandwidth [12] [13] [14] [15] [16]. Figure 1.2 shows an example of hybrid integration of

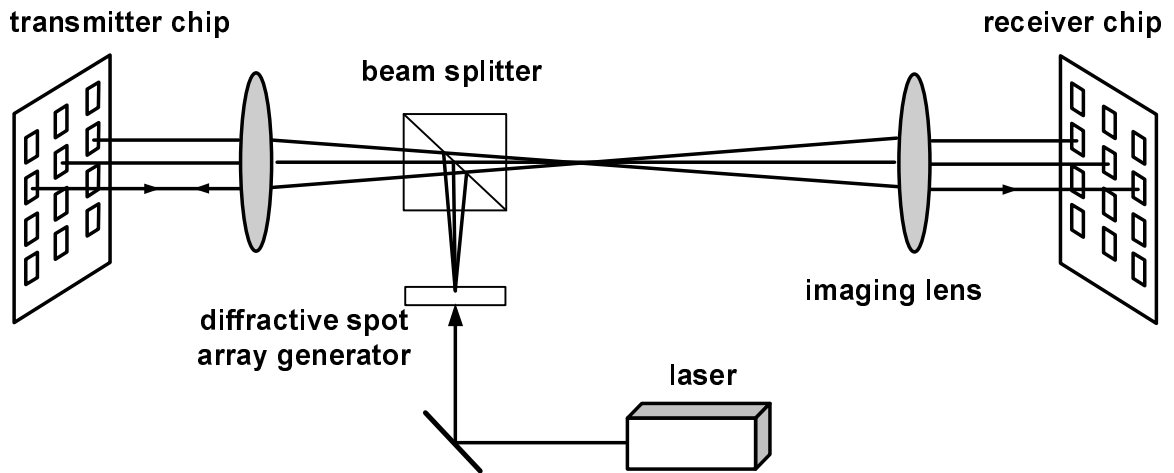


Figure 1.1: Parallel optical chip-to-chip interconnection over free space

optical devices to the surface of silicon, developed in Stanford University.

1.1 Organization

Link designers have solved many of the problems associated with high-speed electrical signaling. Understanding these issues and solutions for electrical links is critical for us for a number of reasons. First, all optical links have an electrical link embedded in them. They surround that electrical link with an electrical-to-optical (EO) and optical-to-electrical (OE) conversion. Second, since optical links are proposed as an alternative to the electrical links, recognizing the limitations and performance of electrical links is essential to evaluate the utility of optical links. Chapter 2 of this thesis is dedicated to reviewing electrical links. That chapter provides a background in high speed data transmission systems and motivates the application of optical signaling.

Optical links can have significant advantages over electrical links only if large number of parallel optical beams can interface with each IC. Scaling of parallel optical interconnects to hundreds and thousands of links on a single chip requires receiver and transmitter circuitry that are very small and have very low power consumption at

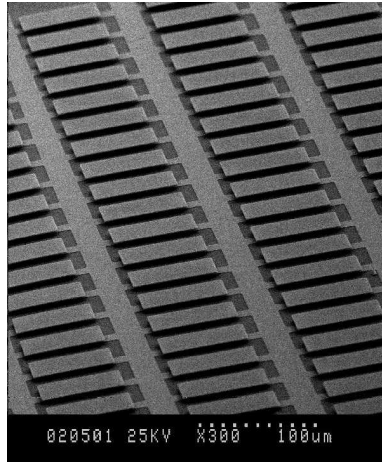


Figure 1.2: Array of GaAs optical devices, flip-chip bonded to the silicon CMOS chip

high data-rates. This is a different requirement than that for long-haul communication links, where sensitivity and bandwidth are the most critical issues and complex designs are allowed. For parallel optical interconnects, the design of a low power receiver front-end is particularly challenging. A photodetector, usually a reverse-biased photodiode, converts the optical power to a small proportional electrical current. The front-end receiver tasks mainly include converting this current to voltage, amplification and data detection. The receiver must add minimum noise and distortion to the signal at high data-rates and operate over a wide range of input optical power.

Chapter 3 of this thesis focuses on the receiver design for short-haul optical interconnects. It starts by reviewing optical to electrical conversion and optical devices used for this purpose. Understanding the properties of these devices is essential in designing an optimized and high performance front-end. Chapter 3 continues by exploring the prior art in optical receiver design. Optical signaling has been used for many years for long-haul communication and considerable effort has been dedicated to the design of the front-end for those applications. Investigation of the existing solutions can help us to understand the challenges of optical front-end design. Finally Chapter 3 presents a new low-power optical receiver front-end that uses a double sampling/integrating technique for data resolution. Unlike most prior designs, this receiver avoids having any linear/analog gain in the active-path and does not rely on

a high gain-bandwidth product voltage amplifier.

The eventual goal in optical interconnect design is to have thousands of transceivers in a single chip. The continuing scaling of feature sizes in the CMOS technology, allows smaller and faster circuitries. Chapter 4 investigates how the performance of the proposed front-end scales with the advanced technologies. In particular we look into what factors will eventually limit performance scaling. While the data-rate, power consumption and area of electronic circuits improve with the scaling, the reduced power supply voltage and increased leakage current can introduce new problems to any design. The integrating nature of our receiver brings concern regarding the dynamic range of input optical power, as well as the acceptable data formats for the correct operation of the front-end. These concerns are even more serious with the scaled power supplies of the advanced CMOS technologies. In this chapter we propose a decision directed control scheme that significantly reduces these effects.

In any synchronous data transmission scheme, clocking and synchronization are among the most challenging problems. The proposed double sampling/integrating receiver is a clocked front-end and needs a synchronous clock signal to perform the sampling and comparison. Multi-phase clock generation and synchronization issues are discussed in Chapter 5. The timing of the bits is precisely controlled by a phase-locked loop at the transmitter. The optimal sampling time at the receiver is maintained by clock generation and timing recovery circuits. The clocking circuits, both at the transmitter side and receiver side generate multi-phase clocks for the de-multiplexing. The integrating front-end allows an efficient baud rate clock recovery technique that reduces the power consumption and complexity compared to the standard clock recovery techniques.

The techniques proposed in this thesis were implemented in two test-chips fabricated in $0.25\mu\text{m}$ CMOS. The first chip demonstrated the possibility of very low-power receiver design that achieves a high bandwidth and sensitivity [17]. The performance of this chip is discussed in Chapter 3. The low-power consumption of the front-end makes it an excellent candidate for building dense receiver arrays on-chip. We implemented a two dimensional array of optical transmitters and receivers in the second

test-chip. The receiver design was improved to achieve higher data-rates and implemented the clock recovery techniques described in Chapter 5 [18].

Finally, in Chapter 6 we summarize the conclusions of this work.

Chapter 2

Background

Point-to-point parallel electrical links have been widely used in short-distance applications such as multiprocessor interconnections [19] [3], networking and communication switches [4] [20], and consumer products [21]. As we mentioned in Chapter 1, in this thesis we explore a number of techniques that facilitate using optical signaling as an alternative in applications that demand high bandwidths. While the traveling signal in optical links are light beams, they still interconnect electrical IC chips. The complete link always starts with an electrical signal at the transmitter and results in an electrical signal at the receiver side. Therefore many internal blocks and principles of electrical links and optical links are identical. Before focusing on optical link design as a replacement for electrical signaling, it is essential to understand the latter. This chapter provides a background in electrical link design. We investigate the basic structure of electrical links, transmitter and receiver design, as well as clock recovery techniques. Since channel capacity and noise are the major limitations for increasing the data rate in most systems, we continue this chapter with reviewing the channel properties and techniques developed to solve the problems associated with them.

2.1 Electrical Link Basics

Electrical links can provide high communication bandwidths between chips, and consist of three major components as shown in Figure 2.1. The transmitter converts the

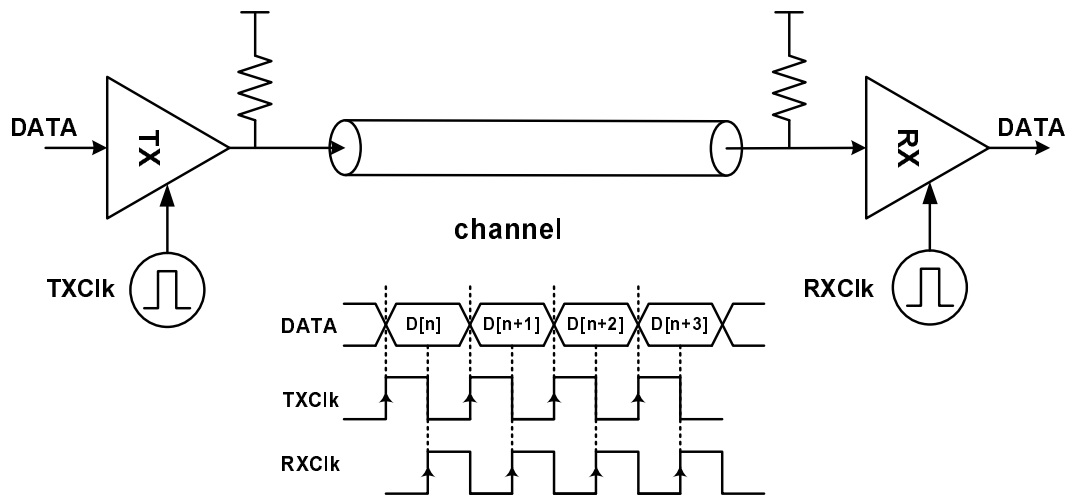


Figure 2.1: Components of a basic electrical link

digital data into an electrical signal that travels through the channel. The electrical channel is the complete electrical path from one die to the other. This channel can consist of traces on a printed circuit board (PCB), coaxial cables, shielded or unshielded twisted pairs of wires, traces within chip packages, and the connectors that join these various parts together. A receiver then converts the incoming electrical signal back into digital data.

The conversion of a discrete-time digital signal into a continuous-time analog signal is called modulation. Here we limit ourselves to the simple non-return-to-zero (NRZ) modulation format, where the data is sent directly on the channel, and the signal levels are represented by different electrical voltages. This modulation technique is called pulse amplitude modulation (PAM). 2-PAM is the simplest signaling scheme where the transmit symbol is a binary signal. The electrical signaling is often low swing to reduce the power consumption.

In most electrical links a synchronous or plesiochronous transmission scheme is adopted, where a clock signal at the transmitter is used to define uniform time periods for sending the data signals successively one after another. At the receiver side a similar clock signal, synchronized with the data is used to sample the incoming signal. While the synchronous data transmission scheme can be high-bandwidth, it requires

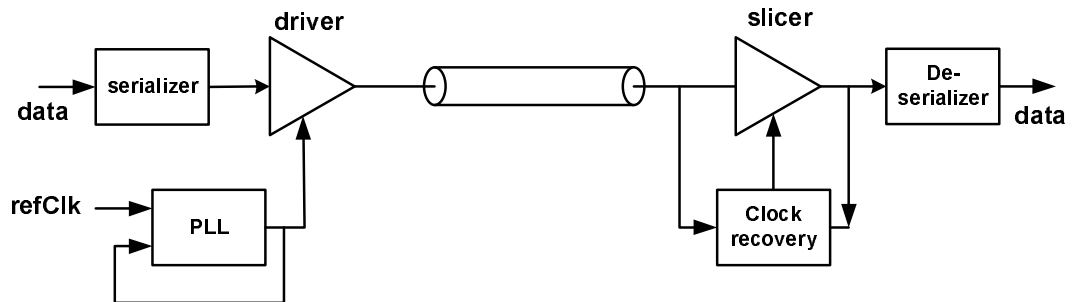


Figure 2.2: Plesiochronous serial link

very precise control on the timing of the signals at the transmitter and receiver. In most systems, phase-locked loops (PLLs) or delay-locked loops (DLLs) are employed to achieve a high level of timing accuracy [22] [23] [24]. The receiver sampling clock should be optimally positioned to minimize the bit error rate (BER) and achieve very high data rate. The optimum sampling point is usually close to the middle of the bit-period. The adjustment of the receiver clock frequency and phase is called clock and timing recovery.

Point-to-point data transmission techniques historically have been divided in two groups: serial and parallel links. Figure 2.2 illustrates a typical serial link, where the transmit data is serialized and sent through a single, high-data-rate link to the receiver, where the data is de-serialized to parallel slower sets of bits. In this scheme the receiver needs to recover the clock and timing from the transitions embedded in the incoming data stream [25] [26]. This clocking scheme is called plesiochronous. Many serial links are designed to support very high data-rates over relatively long distances [20] [27] [28]. In order to meet these objectives and perform the accurate clock recovery, serial links may need relatively complex designs [8] [29].

Figure 2.3 on the other hand illustrates a typical parallel link. This is a common architecture to enable high bandwidth communication between two chips that integrates several parallel sets of data links whose delays through the channels match [30]. A separate reference clock, synchronized with the data, is then sent from the transmitter to the receiver. This clocking scheme is called source-synchronous. The transmission of reference clock (RefClk) simplifies the design of the receiver clock recovery.

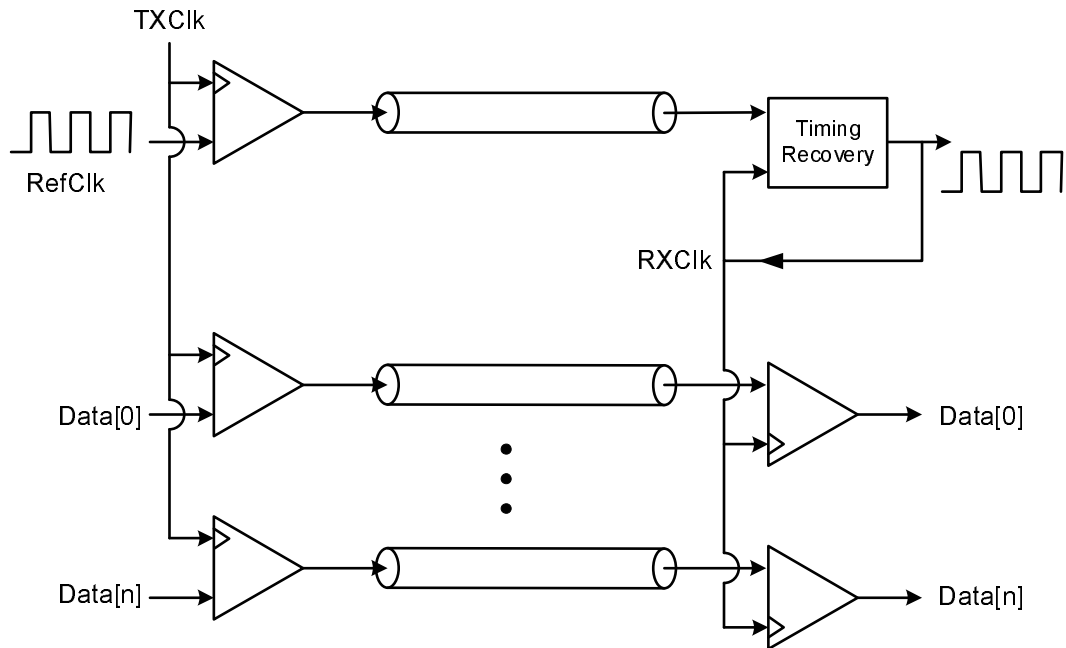


Figure 2.3: Parallel source-synchronous link

A simple delay-locked loop can be used to align the edge of the RXClk to the midpoint of the RefClk. Simple IO design is crucial when many numbers of IOs per chip is required.

As system requirements change over time, the design goals, features, and applications of modern serial links and parallel links are converging. In parallel links, due to the process mismatches a precise delay matching between different pins is not practical and can limit the data rate. Many parallel links now employ traditional serial link techniques with per pin clock and data recovery (CDR) [31] [32], in the quest for higher bandwidth. On the other hand reducing power consumption and design complexity are now among the goals of serial link designers.

2.1.1 Channel Termination

The electrical channel used for high speed signaling is normally impedance-controlled and is modeled as a transmission line with the intrinsic impedance Z_0 . Termination

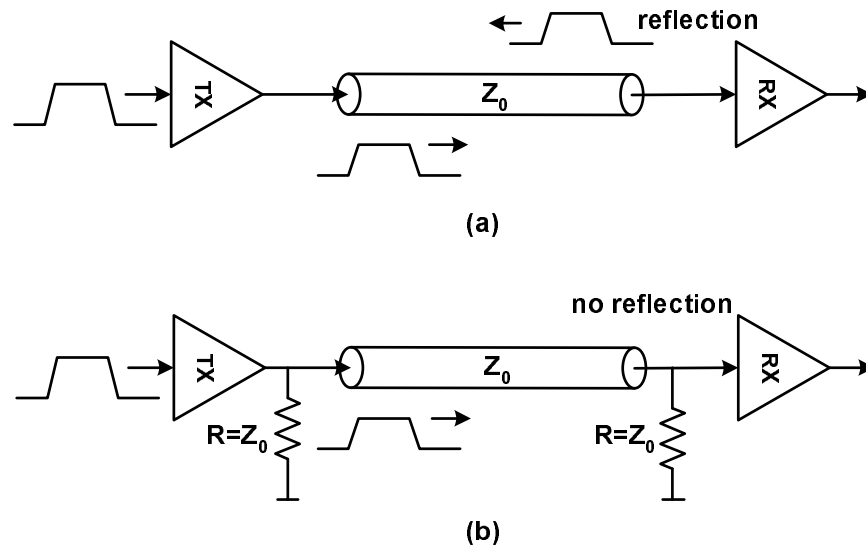


Figure 2.4: Electrical signaling (a) with no termination, (b) properly terminated

of the channel to an impedance matched with the channel is critical in design of high speed links. Termination effectively suppresses reflection, which can cause interference and limit the signalling rate. Figure 2.4 illustrates the cases when the channel is properly terminated and when it is not. When the channel is not terminated, the signal that arrives at the receiving end can bounce back to the opposite direction. With no termination at the transmitter side the signal keeps bouncing back and forth until it dissipates in the channel. In this situation, for the correct data decision, the transmitter may wait until the reflections are very small to send the next bit. However, this will reduce the data rate significantly.

With proper channel termination at both transmitter side and receiver side, the signal is fully absorbed and therefore consecutive bits can be sent through the channel with no delay imposed by the reflections. In fact the next bit can be sent even before the previous bit reaches the receiver. The termination at the transmitter is necessary since deviations from ideal matching at the receiver can cause some reflections that should be absorbed at the transmitter side.

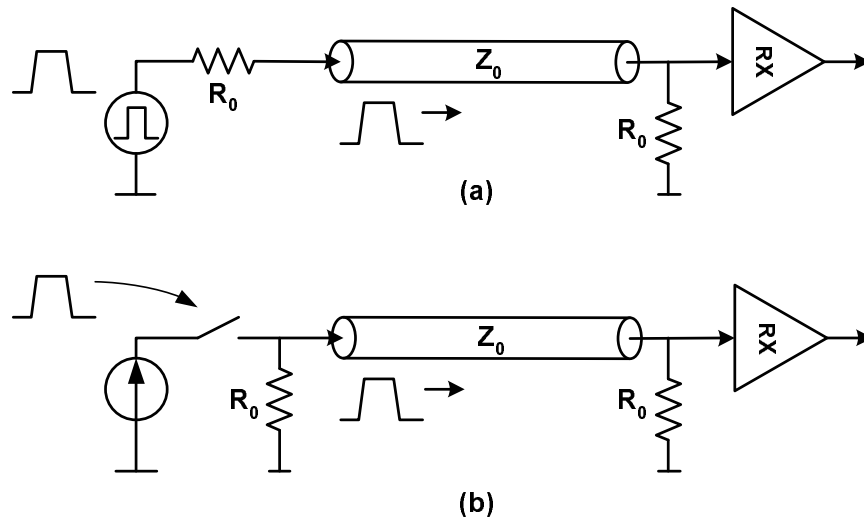


Figure 2.5: Data transmission schemes (a) low-impedance, voltage mode drive, (b) high-impedance, current mode drive

2.1.2 Transmitter Design

A signal transmitter converts digital data into electrical signals that propagate through the impedance-controlled channel to a receiver at the opposite end. For high-speed data communication links, this must be done with accurate signal levels and timing. The receiver chooses a threshold voltage to resolve the data sent from the transmitter. A common voltage reference is then required to correctly resolve the data value. Usually ground is chosen to be the common voltage reference for communication.

The signal transmission over the line can be done with either a low-impedance driver or a high-impedance driver, shown in Figure 2.5. High-impedance signaling is the most common type of signaling in today high-speed links [29].

In high-impedance signaling, the output signals are generated via a current source that turns on and off depending on the polarity of the transmitted data. The voltage swing at the output depends on the termination and the size of the current source. In order to control the voltage swing and proper termination - to avoid reflections - the current source should be kept in the saturation region and the termination resistor should be carefully designed and controlled [33] [34] [35]. Figure 2.6 shows a possible

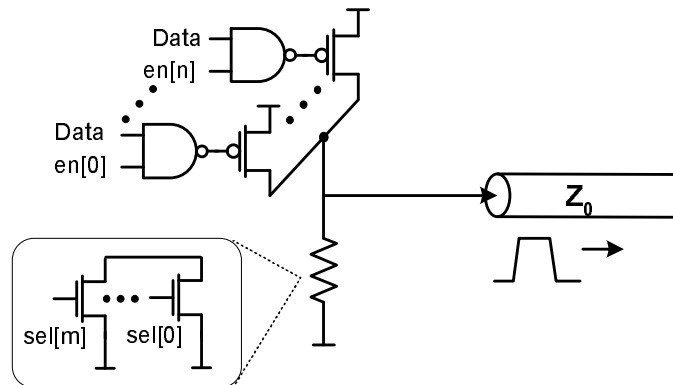


Figure 2.6: Current mode driver

implementation of the transmitter with PMOS transistor current driver and NMOS transistors in linear region as termination resistor.

For high-speed applications, it is crucial to maintain a robust voltage swing and slew rate at the output of transmitter. The properties of transistor and passive elements that set these values are strongly dependent on temperature and process. Therefore, in many designs the switching current source is determined via a feedback loop to monitor and control the output voltage swing.

2.1.3 Receiver Design

A conventional receiver design is a slicer that samples the incoming data in the middle of the bit-period and compares that sample with a threshold voltage chosen to be between the signal levels of the adjacent symbols; see Figure 2.7. For high-speed applications, designers try to avoid linear amplification of the input signal - which requires a high-bandwidth amplifier - by launching enough signal power into the channel. Therefore the slicer samples the raw input signal with no pre-amplification.

The general requirements of a receiver in high-speed applications are high bandwidth, high gain, low noise and low offset. The noise and offset of the comparator plus the coupled noise can increase the bit error rate significantly.

Many receiver designs have been implemented and published in the literature [36] [37] [38]. A different approach that offers good noise filtering is the integrating receiver

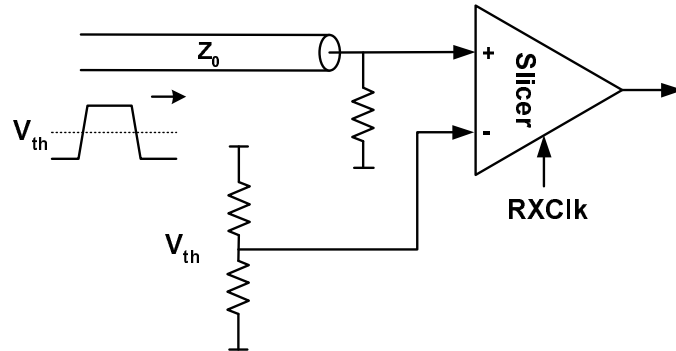


Figure 2.7: Data resolution at the receiver using a slicer

proposed by Sidiropoulos et al. and described in [39]. The integration of the received signal over the bit-period rejects high frequency noise. However, it requires accurate phase alignment between the clock and data.

2.1.4 Time Division Multiplexing

In all transmitter and receiver designs mentioned in the previous sections, the data rate is dictated by the maximum on-chip clock frequency. The clock is used to generate the bit stream at the transmitter and to sample the input signal at the receiver side. In any CMOS technology the maximum clock frequency that can be buffered and transferred robustly across the chip is limited by the switching speed of transistors. Technology speed is sometimes evaluated by the delay of an inverter buffer with a fanout-of-four (FO4). With no loss in amplitude, a chain of CMOS inverters can propagate pulses that are as short as 3-4 FO4 inverter delays. Therefore the minimum clock period is limited to 6-8 FO4 inverter delays. The FO4 inverter delay in picoseconds will decrease as the device feature size shrinks.

The data transmission rate can significantly increase if we use parallelism by effectively implementing a time-division multiplexing. This is possible by using multiple branches of transmitters and receivers as well as multiple phases of the clock, which are equally shifted over time [25] [26] [20]; see Figure 2.8. Assuming that the number of branches is M , in this scheme each branch is driven by a clock, which is M times

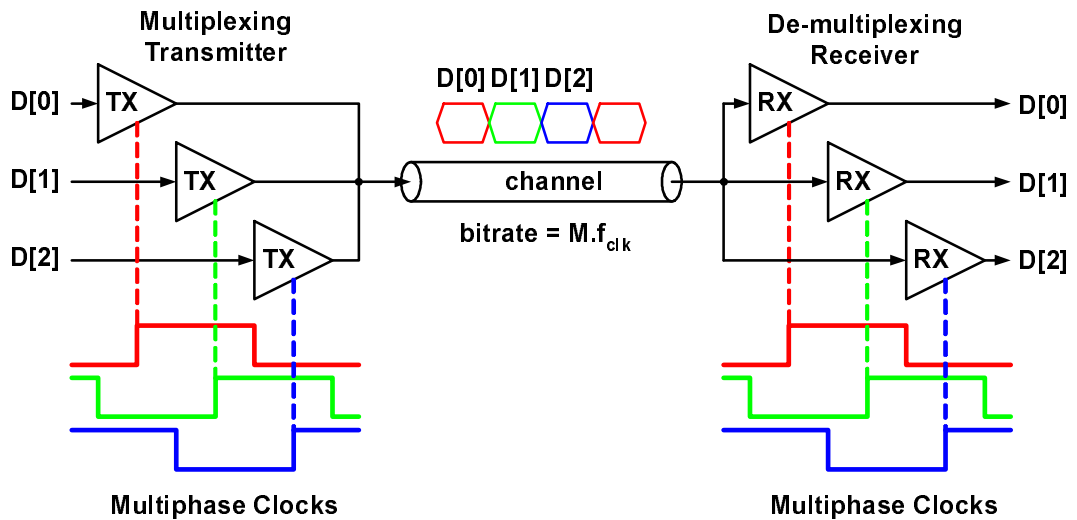


Figure 2.8: Time division multiplexing and parallelism

slower than the data rate. Multiple transmitters connected in parallel convert low-frequency parallel data streams into a single high-frequency stream on the channel. Multiple phases of the lower-frequency clock control the on- and off-timing of each transmitter. Similarly, the parallel receivers convert the high-frequency data stream back to the low-frequency parallel data streams. Note that multiphase clocks evenly divide a clock period and set the receiving window of each receiver. The bandwidth requirement of the receiver front-end is relieved by the de-multiplexing scheme since there is no amplifier running at the bit rate. Instead each slicer operates at the clock frequency. With this parallelism approach, bit-periods as low as 1 FO4 inverter delays are possible [25].

2.2 Clock Generation and Recovery Loops

Most high speed links use clock signals to generate a defined bit-time at the transmitter and correctly recover the data at the receiver side. The timing uncertainties of clocks or data signals are referred to as jitter, the AC variation of the period of the waveform over time, and skew, the DC component of the timing misalignment

between the waveform and a reference clock. Phase-locked loops are commonly used at the transmitter side to generate and synthesize a very low-jitter clock from a periodic reference. PLLs can be used for frequency multiplication when the on-chip clock frequency needs to be higher than the reference clock coming from the PCB boards. PLLs are also used at the receiver side to generate a synchronous clock for data resolution. This section reviews the basics of PLL and clocking and timing recovery in high-speed links.

2.2.1 Phase-Locked Loop

A PLL controls the phase of an output clock so its phase is aligned to an incoming signal. In other words, the output clock tracks the phase variations of the input signal. The phase tracking behavior in a PLL is implemented through a negative feedback loop. A PLL can be built around a voltage-controlled oscillator (VCO) or a voltage-controlled delay line (VCDL). Phase-locked loops with VCOs are simply called PLLs while the ones with VCDL are commonly referred to as delay-locked loops (DLLs). Figure 2.9 shows typical block diagrams of a PLL and a DLL.

A VCO-based PLL consists of a phase detector (PD), a loop filter, a VCO and possibly a frequency divider. The phase difference between the output clock and the input signal generates a proportional error signal at the output of the phase detector. The loop filter, which is usually a single-pole low-pass filter (LPF), smooths out this signal and generates a control voltage, called V_{ctrl} in Figure 2.9.

Since in a PLL we measure the phase but control the frequency, a VCO acts as a phase integrator. The resulting PLL transfer function has at least two poles and needs a stabilizing zero in the LPF. The positioning of the zero is critical and careful design is required to ensure the stability when process variations are considered. DLLs on the other hand use a VCDL, which has a linear phase transfer function, therefore DLLs are inherently first order loops and stability concerns are relaxed in this case. However, a PLL can filter out the jitter of the incoming signal and generate a clean clock from a noisy input, while any jitter in the input of a DLL will directly appear at the output.

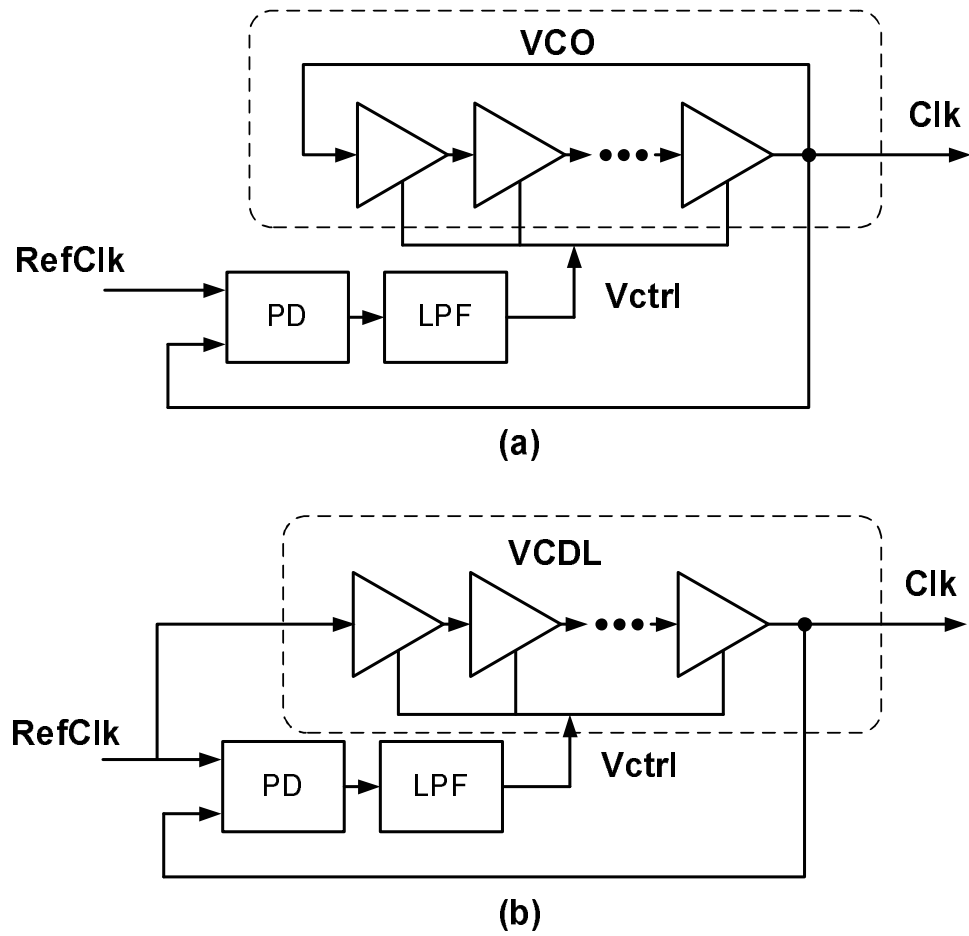


Figure 2.9: Phase-locked loops (a) a VCO-based PLL, (b) a VCDL-based DLL

The general purpose of a PLL is to generate very low-jitter clock signals that track the phase of the input signal. The jitter characteristics depend on the noise of the internal blocks such as delay elements and oscillators, as well as loop bandwidth and phase margin. The most common approaches to design VCOs are ring oscillators and LC oscillators. The noise performance of these VCOs are investigated by Hajimiri: [40]. Although noise performance of the LC oscillator is generally superior to the ring oscillator, rings have been used widely in high speed IOs due to their simplicity, smaller area and easily accessed multi-phase outputs. A differential delay element proposed by Maneatis in [41] uses a replica biasing technique to ensure linearity and high power supply noise rejection. Simple inverters with regulated power supply are also used as adjustable delay elements for building DLLs and PLLs [42]. A regulator essentially isolates the VCO from the noisy power supply in order to reduce the VCO clock jitter.

2.2.2 Voltage and Timing Margins

In most high-speed signaling systems the received signal contains a significant amount of amplitude and timing noise. The noise is added both at the transmitter and during the travel time along the channel. For the receiver that samples, both timing and amplitude uncertainty in the input signal translate into a voltage noise in the sampled value. The quality of the received signal is important since the receiver itself is not ideal. An eye diagram is usually used to examine the quality of the received signal, see Figure 2.10. The jitter and phase offset in the receiver clock as well as the comparator noise and offset calls for a wide eye-opening in the received signal. Voltage and timing margins at the receiver are defined in Figure 2.10.

With conventional electrical signaling, the voltage and timing margins are decreasing rapidly as the data rate increases. For such systems generation of very accurate synchronous timing at the receiver is crucial to sample the signal where noise is minimum.

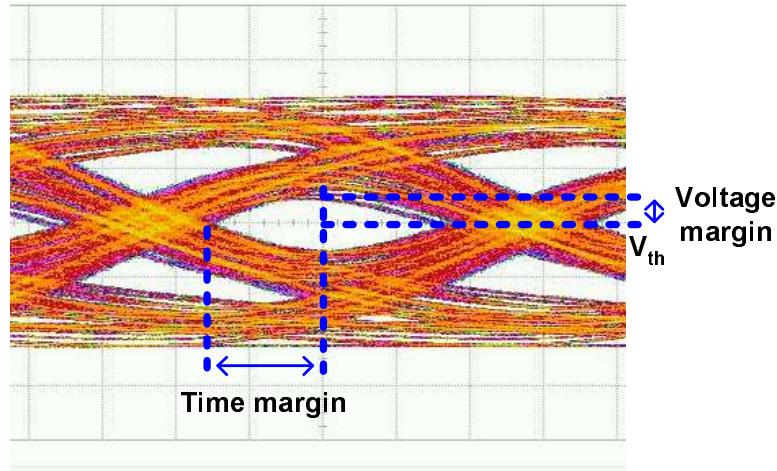


Figure 2.10: Receiver voltage and timing margins

2.2.3 Clock Recovery

In typical high-speed links, due to the process mismatches, time variations and undefined delays in the signal path, the received data can have an undefined phase and frequency. As we mentioned before, either a separate clock signal is sent along the data signal for timing information or the clock should be recovered from the incoming data signal.

The transmitter's limited bandwidth, frequency dependent losses in the channel and reflections all create inter-symbol interference (ISI). This ISI combines with the transmitter clock jitter and cross-talk to add amplitude noise and timing jitter to the received waveform. A clock recovery block should extract the clock component of the incoming signal and filter out the timing jitter. Finding the best sampling time with low variance and phase shift is particularly critical for bandlimited channels to minimize the BER. In most systems direct estimation of a sampling time that minimizes the probability of error in the data resolution is not a practical task. Instead sub-optimal practical solutions are adopted to define the best sampling point. The most common approach is to assume that the best sampling time is where the overall ISI is minimum. At this point the vertical eye-opening in the eye-diagram is maximum as is shown in Figure 2.10.

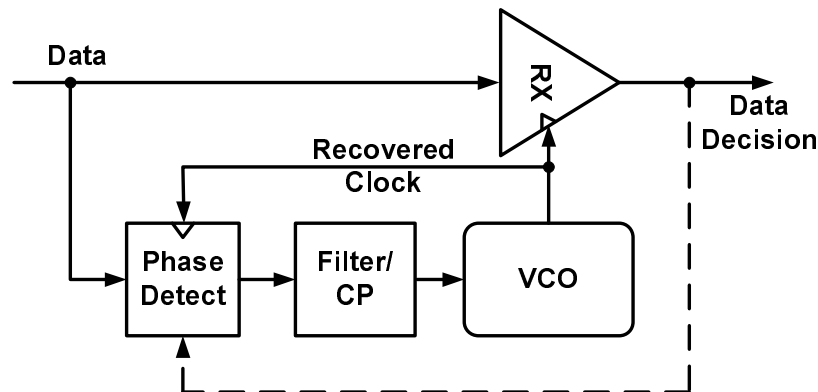


Figure 2.11: Timing recovery from the received data

Figure 2.11 illustrates a typical timing recovery loop for a serial link. In case of phase mismatch between the clock and received data, the phase detector generates an error signal to adjust the VCO's phase and frequency. The output of the phase detector can be an analog signal for an analog loop or a digital correction commands for a bangbang-controlled loop. In a bangbang-controlled PLL, the phase and frequency of the VCO are corrected by constant steps in two different directions depending on the decisions of the phase detector. The correction commands are called "up" and "dn" commands in this thesis. As we will see in the following sections, a decision-based phase detector is used in many CDR techniques. In these CDRs, the data decisions are usually used to determine the type of transition that occurred and then use that information to find the correction needed.

A number of different techniques for phase detection and definition of best sampling time have been proposed [25] [26] [43]. The over-sampling technique described in [25] takes 3 or more samples per each bit and performs the data resolution for all of them. Then, by looking at the sequence of resolved values, it decides which sample is the most reliable one, which is simply the one farthest from the transitions. This technique is very robust but requires a considerable hardware, power and area overhead for the oversampling. It also needs phase spacings that are at least three times shorter than a bit-period. For systems that run at data rates close to the technology limitations, such sampling schemes are not practical.

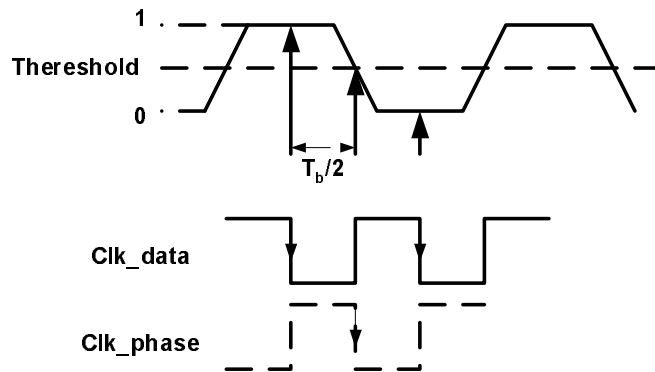


Figure 2.12: 2x over-sampled CDR with two clock phases for data and phase resolution

A common CDR technique in electrical signaling is the 2x-oversampling scheme [26]. This technique is based on the assumption that if the best sampling time with minimum ISI is at time τ , at any "one-zero" (10) or "zero-one" (01) transition, an edge sample at time $\tau - T_b/2$ is expected to be equal to the threshold value. T_b here is the bit-period shown in Figure 2.12. Any phase error will cause deviation of edge-sample from the threshold value. The difference can be used as an error signal in a phase correction feedback loop. This error signal can indicate both the magnitude and direction of phase error. The effect of an early clock is shown in Figure 2.13. Depending on the direction of transition, 10 or 01, the error signal can be positive or negative for an early clock. A simple block diagram of a 2x-oversampling bangbang-controlled CDR is shown in Figure 2.13. The $P[n]$ signals are generated by comparing the edge sample with the threshold value. $P[n]$ and the resolved data values before and after the edge sample, $D[n-1]$ and $D[n]$ provide complete information for phase detection. As shown in this figure, a replica of front-end slicer, clocked with an extra clock phase that is shifted by half a bit-period, can generate the phase information, P signals.

The 2x-oversampling clock recovery technique is used widely in high-speed PAM links. This technique is called zero-forced-detection (ZFD) in some literature. Clearly for binary 2-PAM the phase information is valid at all transitions, which occur 50% of times. For higher levels of PAM, specific types of transitions that ideally cross

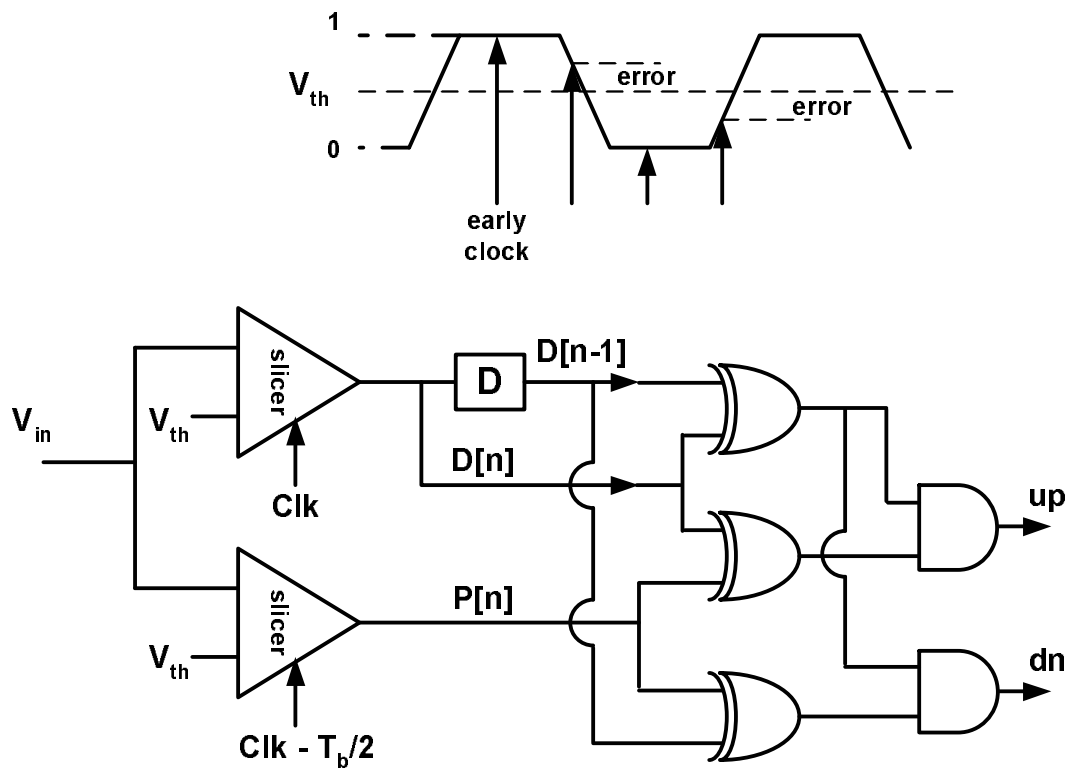


Figure 2.13: 2x over-sampled phase detection, loop error signal and implementation

the threshold at time $\tau - T_b/2$ are the ones that generate accurate phase information [26] [44]. One of the concerns with this clock recovery technique is the uncertainties of edge samples in the presence of ISI, as well as the inaccuracy of comparison with V_{th} . In other words the error rate in P signals could be much higher than the error rate in resolved data D , which means heavy filtering of phase information is needed for good performance. Moreover, 2x-oversampled CDR needs an extra clock phase that is accurately positioned in the middle of main phases for data samples. Generation and distribution of extra phases are particularly challenging and power hungry when a time division multiplexing technique is used.

Baud rate clock recovery refers to techniques that use only main data samples for both data recovery and adjusting the clock frequency and phase. Thus the baud rate CDR does not need the extra clock phases and edge samples. This implies that this technique potentially reduces the complexity and power consumption compared to the over-sampled techniques. There are a number of different techniques proposed for baud rate clock recovery in digital data communication systems. The technique proposed by Mueller and Muller in [43] provides phase information by identifying certain properties of the pulse response of the ISI channel. In this technique we can find a relationship between the baud rate samples of the channel pulse response that is true only if the timing of the samples are aligned with the pulse. If this relationship is defined as $f(\tau)=0$, then τ is the best sampling time and a phase misalignment generates a positive or negative $f(\tau)$. Figure 2.14 shows a typical channel pulse response, $h(\tau)$ and two different ways of identifying the best sampling time. In the first technique point A at time $\tau = \tau_A$ is defined as a solution to $f(\tau) = h(\tau + T_b) - h(\tau - T_b) = 0$, where T_b is the bit-period. Point B on the other hand is the solution to $f(\tau) = h(\tau + T_b) = 0$.

During the data transmission the baud rate samples of the incoming waveform are not equal to the samples of this pulse response. However, each baud rate sample of the incoming signal is equal to the sum of the pulse response sample at time τ and other samples at $\tau + kT_b$, $k=1,2, \dots$ which act as ISI. All these samples are modulated by their corresponding bit values. Therefore, theoretically, if we know the bit values, it is possible to reconstruct a similar relationship between the baud rate samples of the

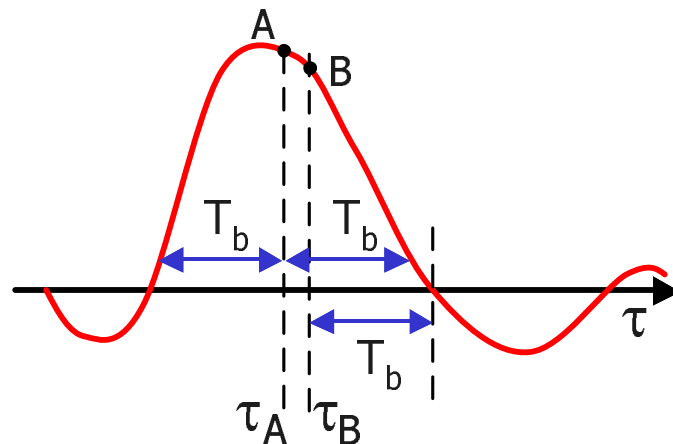


Figure 2.14: Channel pulse response and criteria for baud rate recovery

incoming signal that has timing information for clock recovery. The expected value of this relationship can be used as an error signal in a CDR loop, and it is chosen to minimize the variance. While baud rate CDR is a very attractive solution, there are a number of challenges associated with this approach. First, the defined relationship between the samples and resolved data values requires linear and accurate arithmetic processing of the analog samples. This processing is usually hard and needs to be fast. Moreover in order to reduce the phase variance of the phase correction loop, it might be necessary to choose only certain data patterns or for different patterns calculate different functions. This process adds to the complexity of the baud rate CDR techniques for regular electrical signaling.

2.3 Channel

At the data rates required in many systems today, the filtering imposed by the electrical channel is among the most challenging problems. The performance of the channel strongly depends on the application. As an example a typical backplane link and its components are shown in Figure 2.15 [45]. Loss per unit length of PCB-traces increases with the frequency due to the dielectric loss and skin effect. Different trace

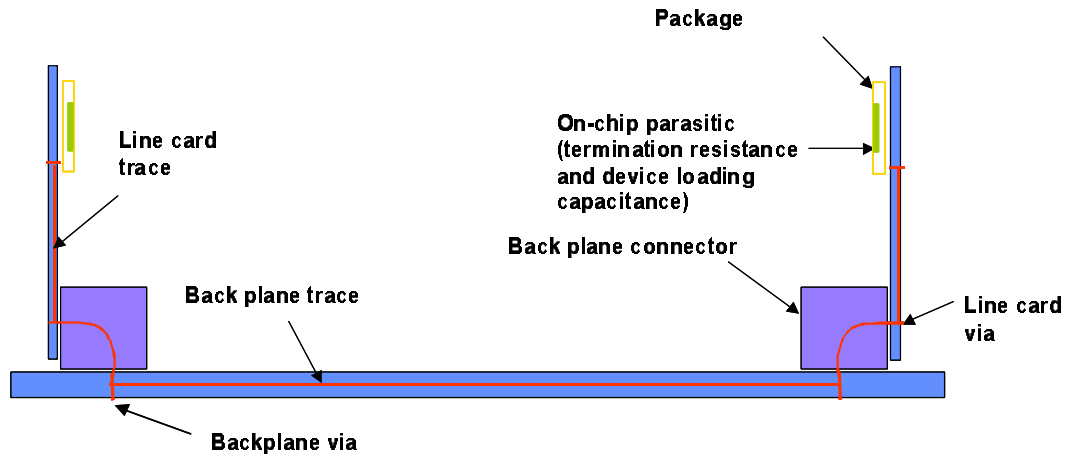


Figure 2.15: Signal path and channel components in a typical backplane

lengths and backplane material properties, as well as types of connectors, vias and routing layers, cause significant variation in channel transfer function both among different boards and among channels in the same backplane.

The typical transfer functions for channels within a single backplane are illustrated in Figure 2.16 [8]. Since the loss in the channel increases with the frequency, the channel acts as a low-pass filter. The filtering effects leads to spread of narrow pulses originally confined to a bit-period as shown in Figure 2.17. This effect is called dispersion. The tail of the pulse acts as an additive noise for the next bits and is referred to as inter-symbol interference or ISI. Dispersion is enhanced by the filters formed by unintended transmission line impedance discontinuities caused by via stubs and connections. In the time domain these discontinuities cause reflections, which also lead to ISI. Crosstalk or co-channel interference is the other problem that occurs in dense interconnects. Both far and near end cross-talk (FEXT and NEXT), are important in such systems.

One way to reduce the cross-talk problems is to use differential signaling instead of single-ended. A differential signalling scheme transmits both the signal and its complement on a differential channel as it is shown in Figure 2.18. A drawback of differential signaling is the need for an additional pin per IO and potentially increased power consumption. However, reduction in the noise possible through differential

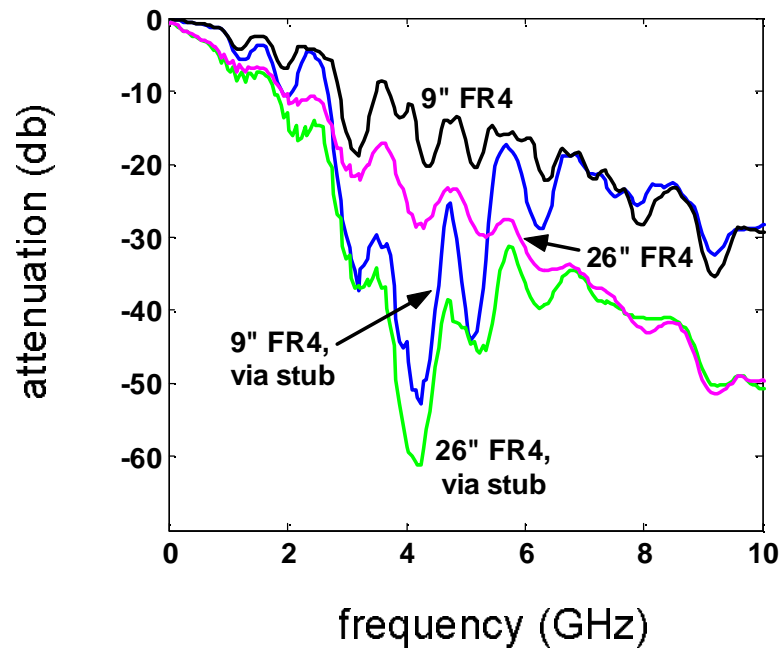


Figure 2.16: Typical frequency transfer function of a backplane channel

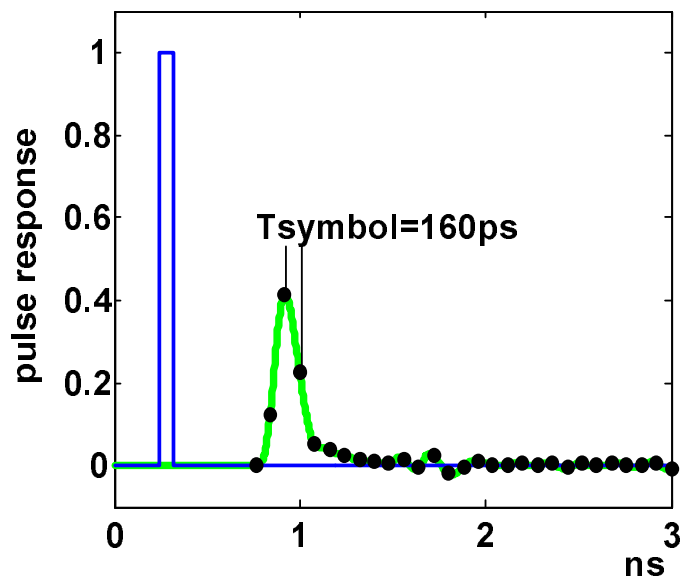


Figure 2.17: Dispersion in a band-limited channel

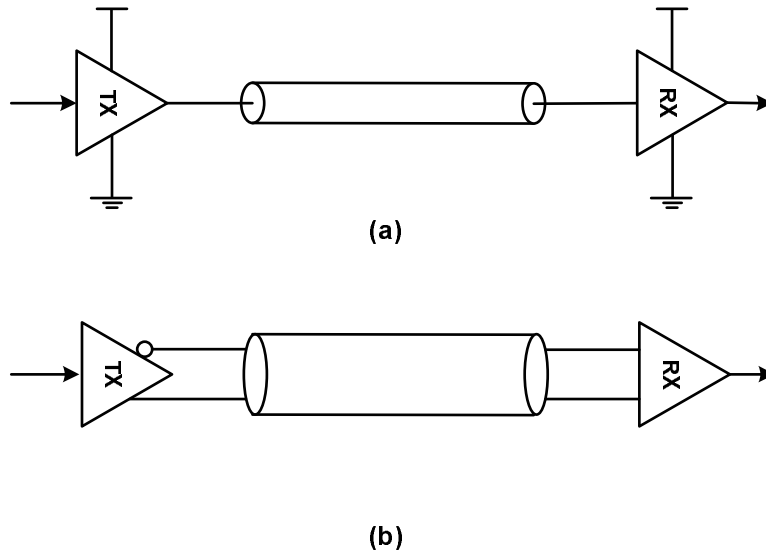


Figure 2.18: Signaling methods (a) single-ended signaling, (b) differential signaling

signalling may enable lower signaling levels to reduce the overall power consumed, relative to single-ended signalling.

Even with differential signaling the dispersion problems will not go away. Also by simply increasing the signal intensity the SNR does not improve significantly since the ISIs depend on the signal itself. In fact these residual errors are not random but deterministic in nature. This means that if the channel pulse response and input sequences are known, it is possible to at least partly correct the channel dispersion at the receiver and transmitter. These techniques are referred to as equalization, and explored in the next section.

2.3.1 Equalization

Equalization techniques have been used increasingly in high speed links in recent years [46] [47] [38] [26]. An equalizer subtracts the ISI in the time domain or equivalently flattens the frequency response of the channel. In order to flatten the frequency response we can boost the high frequencies relative to the low frequencies, or attenuate the low frequencies; see Figure 2.19. These techniques are known as linear

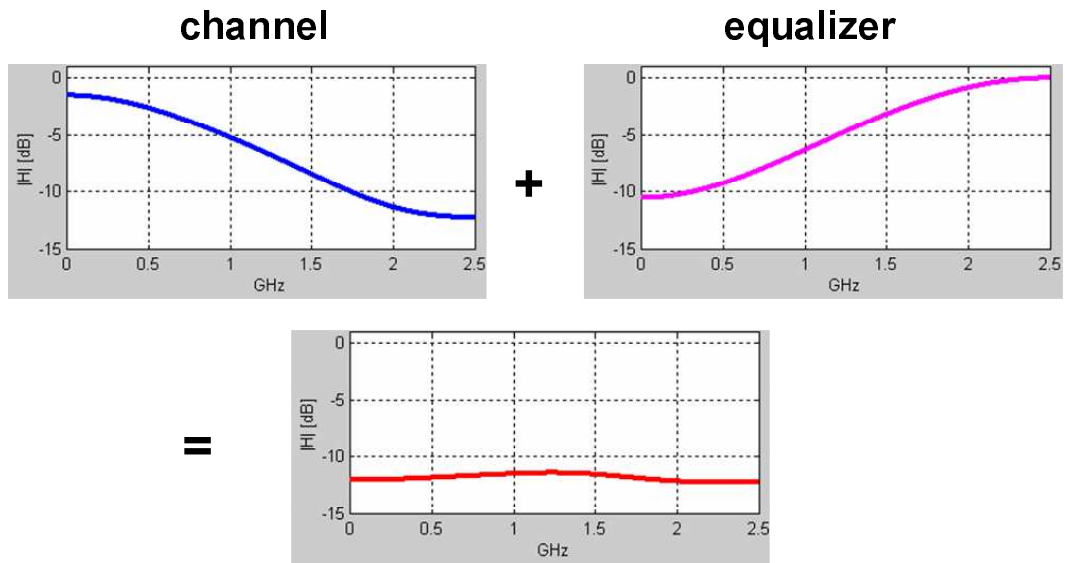


Figure 2.19: Linear equalization, flattens the frequency response

equalization. A linear equalizer at the receiver side can be built with an FIR filter shown in Figure 2.20. The FIR filter could be either digital or analog. This equalizer effectively amplifies the high frequencies attenuated by the channel. The main issue with such a solution is that the noise at high frequencies is also amplified. The coefficients W_i , can be set by using adaptive algorithms such as least-mean-square (LMS). In an analog FIR filter, the precision of the arithmetic operations as well as building the analog delay lines are among the challenges of building the FIR filter.

Alternatively the equalizer can be built at the transmitter side. The transmitter equalizer shown in Figure 2.21, attenuates the low frequency components of the incoming signal. In this design, due to the limited output power, the amplitude of the output signal should be adjusted accordingly. The FIR filter here is in fact a digital to analog converter (D/A) while the FIR at the receiver is a analog to digital converter (A/D). In general the design of an D/A is simpler and one can achieve better precision in design of transmitter equalizer compared to the receiver equalizer. Setting the equalization weight at the transmitter side is possible but is not as straightforward [48].

In all techniques discussed so far, a linear equalizer tries to invert the channel

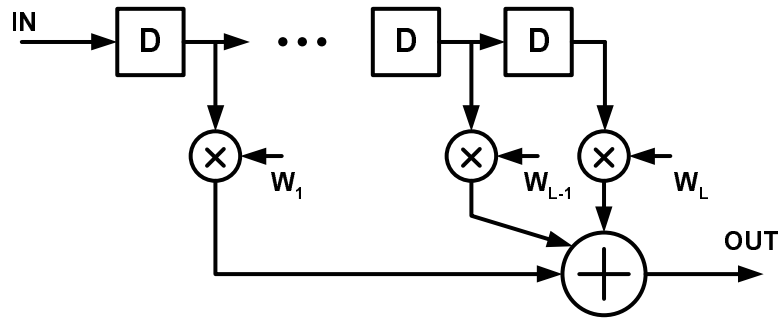


Figure 2.20: Linear equalization at the receiver with FIR filtering

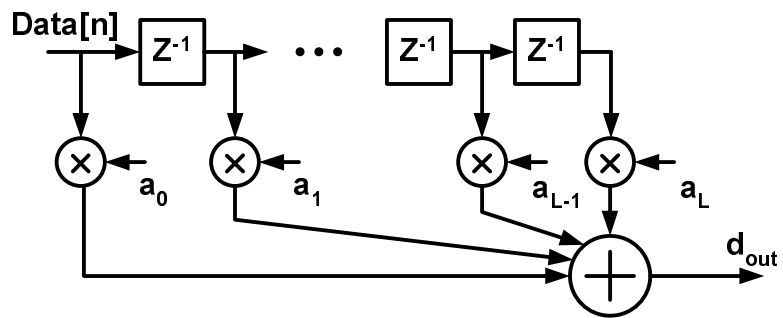


Figure 2.21: Linear equalization at the transmitter with FIR filtering

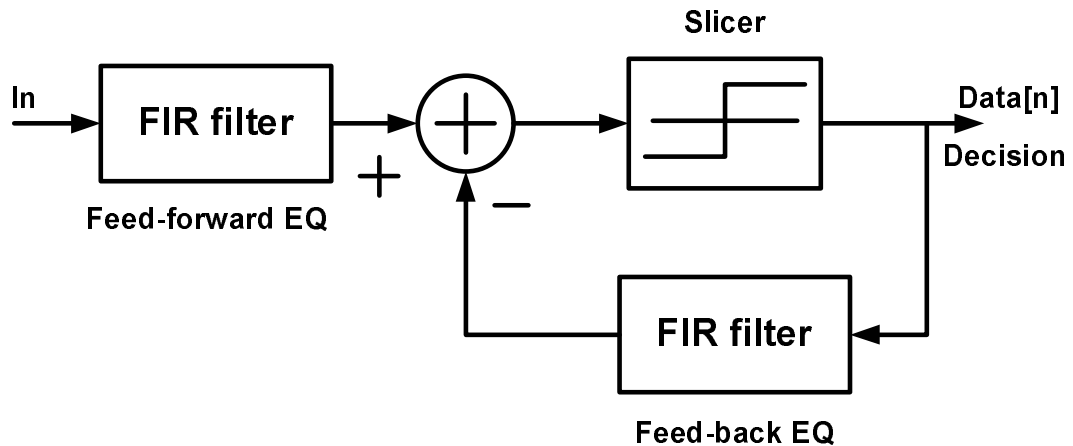


Figure 2.22: Decision feedback equalization

and compensate for attenuations at high frequencies. A different class of equalizers called decision-feedback-equalizers or DFE, focuses on directly removing the ISI from the input signal. Removing the ISI is possible if the characteristics of the channel as well as the data decisions are available. In communications, DFE has been used heavily instead of linear filtering, to circumvent the problem of noise amplification [49]. Recently similar techniques has been applied to serial links [50] [44].

A possible DFE implementation is shown in Figure 2.22. Here the slicer resolves the data value and the decision, and older bits are fed to an FIR filter that then drives a DAC whose output is subtracted from the input signal [51]. The feedback filter of the DFE only removes the ISI caused by previous bits, while a feedforward filter is necessary to remove the pre-cursor ISI. The feedforward FIR is normally implemented at the transmitter side, where we have access to the bits not sent to the channel yet.

Unfortunately the above approach suffers from latency problems at the receiver side. For a 10Gb/s binary link, we have only 100ps to resolve the input, drive the DAC, and have the DAC outputs settle to the required precision. The latency problem has motivated an interesting approach called loop unrolling. While the decision is made for the last bit, the feedback is computed two times in parallel, assuming the incoming data is either "1" or "0". As soon as the data is resolved, one of the outputs is chosen by using a fast multiplexer. This technique was proposed by Parhi

and Kasturia [52] [53]. This approach doubles the number of comparators in a 2-PAM scheme if it is applied to one tap of the equalizer. This method can be applied to two or more taps of feedback, though the number of decision blocks increases exponentially.

DFE is capable of canceling the ISI caused by reflections from the discontinuities in the signal path. This type of ISI is usually created by the bits that are far enough from the current bit, and latency in DFE is not a problem anymore. However a very high number of taps adds to the complexity and parasitics at the input node, as well as to power consumption and area. In most systems only a limited number of taps are available.

2.4 Summary

Electrical signaling is the most natural approach for interconnecting electronic chips since no conversion of energy is required and achieving high levels of integration is possible. As a result of extensive design efforts as well as access to faster transistors, electrical signaling has experienced a significant performance improvement in the last two decades. In many electrical signaling applications, most of channel components, for instance the PCB boards, are kept unchanged to avoid extra cost. In these systems considerable effort is dedicated to the design of transceivers capable of high-speed signaling through the same channel to meet the requirements of today systems. However, the high-speed signaling is now running into the limitation of channel bandwidth and discontinuities in the signal path. Moreover the cross-talk between adjacent signals puts a hard limit on the density of high speed IOs in the system. As we discussed in this chapter, advanced timing recovery techniques as well as equalization techniques have allowed the maximum data rates to continue to scale. However these techniques heavily add to the complexity, power consumption and area of the transceivers.

The complexity of transceiver design can be reduced if the channel characteristics are improved. A fundamentally different approach is to use a carrier at much higher frequencies than the signal bandwidth. If the channel can support that carrier, there

is a high chance to have a flat frequency response over the signal bandwidth, which avoids dispersion. An excellent example of such carrier is a coherent light traveling in a transparent medium, which has led to fiber optics communication systems. However, for chip-to-chip interconnections, switching to optics, which involves energy conversions, is beneficial only if one can build simpler transceivers that can result in higher number of IOs per chip. In the rest of this thesis we explore different techniques that can facilitate this goal.

Chapter 3

Receiver Design for Optical Interconnects

The power consumption and area of transmitter and receiver electronics limit the number of IOs possible on-chip. As we discussed in Chapter 2, the electrical channel limitations require complex transceiver designs for supporting very high data rates. Although optical interconnections between electronic chips require the overhead of OE conversion, the high quality of optical channels prompts us to investigate the possibility of building small and simple electronics that allow large numbers of optical IOs per chip. Using optics is particularly promising because of the emergence of techniques for building dense 2D arrays of optical devices hybrid-integrated to the silicon chips [12] [13] [14] [54] [55] [56] [57] [58] [59].

This chapter is focused on the receiver design for optical interconnects. The overall performance of high speed optical links depends on the design of the receiver front-end. Receiver circuitry plus the OE conversion devices directly affect the maximum data rate, required optical power, electrical power consumption and area of the link. Receiver design objectives are strongly dependent on the application and overall system configuration. Optics has been used widely for long-haul high-speed data communication. Optical fibers confine and guide the beam from the transmitter side to the receiver. Since in optical signaling the maximum modulation frequency (data rate) is much smaller than the carrier frequency (frequency of the light), the

frequency dependent power loss in the fiber is very small. A flat response over the frequencies of interest corresponds to very small dispersion in the channel. In long-haul optical fiber communication, other types of dispersions such as modal dispersion can cause ISI [60]. Moreover, although the fiber loss is much smaller than the loss in copper wires, an optical beam that travels over many kilometers, experiences a significant power loss. Thus the main requirements for receivers in long-haul communications are very high sensitivity and bandwidth. Like in electrical serial links, long-haul optical links can afford high levels of design complexity to achieve good performance [60] [27] [28].

For short-haul parallel interconnections, the channel distortion and loss is negligible [10] [11]. However, limitations on power consumption, area, and noise generation become important constraints in the design of large numbers of high-speed IOs on a single chip. Integration of optical devices as well as cross-talk between different signals can add to the difficulties of having thousands of beams connecting different chips.

We start this chapter by briefly reviewing the basics of optical links and optical devices used at the receiver side. Since the receiver circuitry directly interfaces with the photodetectors, understanding the operation and characteristics of these devices is essential for an optimum design. In the second part of this chapter, we focus on the design of the receiver electronics. First we examine the prior art in front-end design for optical communication. Investigating the existing designs provides a motivation for the third part of this chapter, which describes the proposed double sampling/integrating front-end.

3.1 High-Speed Optical Interconnects Overview

With the recent advances in the optical technology, optical devices that can handle 10's of Gb/s data rates are available. These high performance optical devices facilitate very high data rates in optical interconnects, if high bandwidth transceiver circuitries, optimized for the characteristics of optical devices are designed.

In most optical data transmission systems, a coherent light from a laser is amplitude modulated with the transmit data. The amplitude modulation of the light can be done either by directly changing the current of the laser or by using an external optical modulator. A suitable laser choice for the direct modulation in a 2D array system is a vertical cavity surface emitting laser, VCSEL. While using direct modulation avoids the need for an external light source, VCSELs have a number of problems that are subjects of ongoing research. In order to avoid the turn-on delay of the laser, a bias current above the threshold is required. The power consumption and heat generation associated with biasing and modulation of VCSEL can change the properties of the lasers in large arrays. Thus external modulators are preferred in some designs. A promising modulator device is a multiple quantum-well (MQW) $p-i-n$ diode [61]. The absorption properties of the quantum well structure in the i region are dominated by quantum-confined Stark effect (QCSE) [62]. With constant reverse voltage bias across the diode, the light absorption in the i region has a relatively narrow peak at a certain wavelength λ_1 . Because of the QCS effect, changing the reverse voltage across the diode causes the absorption peak to shift to a different wavelength. This means that for a certain wavelength, changing the voltage across the diode causes significant change in the light absorption, which results in a modulation effect. These modulators can achieve data rates higher than 40Gb/s.

In most optical systems a 2-PAM modulation scheme is adopted. The modulated light is then sent to the receiver through an optical channel. For short-haul chip-chip interconnection, where a line of sight (LOS) is possible, the optical beam can be sent directly via the free space to the receiver, which was shown in Figure 1.1. For longer distances or places where a LOS is not available, optical fibers and waveguides are employed.

The data transmission scheme in the optical links is very similar to the electrical links, see Figure 3.1. A clock signal at the transmitter is used to define uniform time periods for sending the data signals successively one after another. At the receiver side a photodiode converts the optical signal to an electrical signal proportional to its input optical power. The receiver circuitry following the photodetector is responsible for resolving the data from the incoming electrical signal. Similar to electrical receivers,

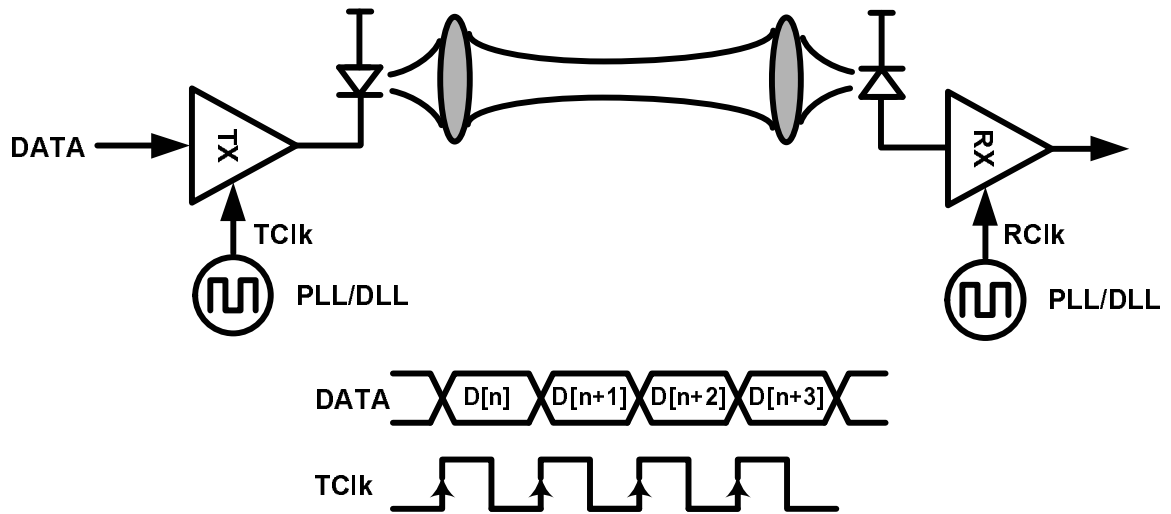


Figure 3.1: Synchronous optical transmission over free space

a clock signal, synchronized with the data is used for sampling and data decision.

In the next section we investigate some of the properties of photodetectors used in high-speed optical interconnects. We are particularly interested in the surface normal devices that can be used for building dense 2D arrays, integrated or hybrid integrated with CMOS chips.

3.1.1 Photodetectors

The two commonly used devices for OE conversion are $p-i-n$ diodes and metal-semiconductor-metal MSM diodes. In both types of diodes, an electrical field in a semiconductor material, drives the electrons and holes generated by the incident photons to the terminals. The result is a current proportional to the number of photons absorbed per second. In a $p-i-n$ diode, a reverse-bias across the diode ensures a strong field in the i region and a very small current in absence of light.

Figure 3.2 shows a simple electrical model for a photodetector. The optically generated current I_{op} is proportional to the input optical power P_{op} . The diode responsivity R is defined as $R = I_{op}/P_{op}(\text{A/W})$. If the wavelength of the light is

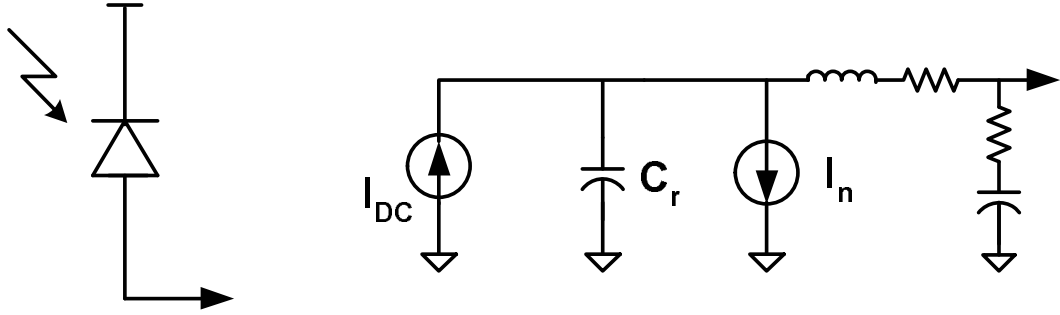


Figure 3.2: Equivalent electrical model of a reverse-biased photodiode

λ , the responsivity of a detector can be expressed as $R = q\eta\lambda/hc$ ¹ where η is the diode quantum efficiency. An optimally designed $p-i-n$ diode, has $\eta \simeq 1$, which at 850nm implies $R = 0.68\text{A/W}$. Stand-alone GaAs and silicon $p-i-n$ detectors have responsivities close to this value. The noise current I_n in Figure 3.2 is mainly due to the diode shot noise. Other sources of noise are the thermal noise of the series resistance and the background illuminations. In most designs the diode noise is much smaller than the noise of the receiver front-end circuitry.

Detector capacitance is usually the dominant input load for a receiver, impacting the sensitivity, electrical power consumption, and for some designs, the bandwidth of the front-end. Additionally, the detector integration scheme can affect the overall footprint and density of the front-end, which is particularly important for parallel interconnects. Photodetectors can be hybrid-integrated to CMOS chips after chip manufacture via a number of techniques, which are reviewed by Krishnamoorthy et al in [15] and Miller et al. in [58]. The advantage of hybrid integration is that the material and design of the photodetector is independent of the transistor technology. The most mature hybridization techniques are wire bonding and flip-chip bonding. Flip-chip bonding is a manufactureable technology enabling the integration of large device arrays. The performance tradeoffs between these two techniques are evaluated in [15]. Wire bonding has greater parasitic inductance and capacitance, reducing performance at high bit-rates as compared to flip-chip bonding. Hence flip-chip bonding

¹ hc/λ is the photon energy, where h is the Plank constant and c is the speed of light.

is more suitable for high-performance applications requiring minimum front-end capacitance. The capacitance of detectors themselves reduces as the width of the high field depletion region is increased. For $p-i-n$ detectors this width is the width of the intrinsic region and for MSMs, it is the finger spacing. Increasing this critical dimension however, lowers the field and can make the detector slower, unless greater bias voltages are available. Capacitance can also be reduced by removal of excess substrate or the use of insulating substrates. For a given design, capacitance scales with the detector size, which is limited by the practical ability to focus light to spots smaller than 5 to 10 microns diameter. Wire-bonded off-the-shelf $p-i-n$ detectors present a front-end capacitance in the range of 200fF or higher depending on the detector size. In [63] the measured capacitance of $15\mu\text{m}\times 15\mu\text{m}$ flip-chip bonded GaAs $p-i-n$ detectors is reported to be 52fF.

A precise photo-detector model is essential for optimum receiver circuit design. However, if a flip-chip bonding technique is used the inductor and resistors in the model can be neglected and the parasitic elements can be reduced to a single capacitor (the diode capacitance plus the parasitic capacitance of the connecting pads and bonds). Therefore for initial design and analysis, we can simplify our model to a current source $I_{op} + I_n$ and a parasitic capacitor C_p .

3.2 Prior Art in Design of Optical Receiver Front-Ends

A number of different designs have been used for optical receivers on chip. These designs include transimpedance amplifier (TIA) [64] [65] [66], diode-clamped or receiver-less front-end [67] [68] and clocked sense-amplifiers [69] [70]. In this section, we examine these different approaches to the design of receiver front-ends.

3.2.1 Front-End Design Challenges

As we mentioned earlier in this chapter, if a flip-chip bonding technique is used, the photodiode model can be simplified to an optically generated current source, a noise

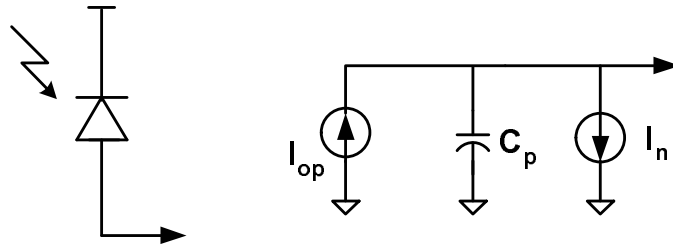


Figure 3.3: Simplified model of a photodiode

current source and a parasitic capacitance; see Figure 3.3.

The task of the receiver front-end is to resolve the value of the incoming signal by sensing the changes in the magnitude of photodiode current. In most applications, with the limited transmit optical power and/or loss in the path, the optically generated current is very small, in the order of $5\text{-}50\mu\text{A}$. In order to achieve a robust data resolution with low BER, the total input-referred noise current from the circuitry and the diode itself should be well below the optically generated current. In general, design of a low-noise front-end with a very high bandwidth is difficult and requires high electrical power consumption or advanced fabrication processes or both.

In all the receivers discussed in this section, the optically generated current is eventually converted to a voltage swing. The simplest solution for this conversion is to add a resistance to the input node to convert the current to a voltage signal. A voltage amplifier then amplifies the voltage swing for the following data resolution slicer block shown in Figure 3.4. Assuming that the voltage amplifier has a high bandwidth, the bit rate of such a front-end is limited by the $1/(R \times C_p)$ where C_p is the diode capacitance. The RC_p time constant of the input node sets a maximum limit on the resistor R . On the other hand the maximum possible voltage swing at this node is equal to $\Delta V = R \times I_{op}$ where I_{op} is the input photocurrent. It is clear that lower R values degrade the signal-to-noise-ratio (SNR) at the input, since the voltage noise variance is constant² and equal to kT/C_p . Hence, there is a strong trade-off between the sensitivity and the bandwidth, both depending on R . A possible solution

²The resistor thermal voltage noise power density is $4kTR$, where k is the Boltzmann constant and T is the temperature in Kelvin. If we integrate it over the RC_p bandwidth, results in a noise power with variance kT/C_p .

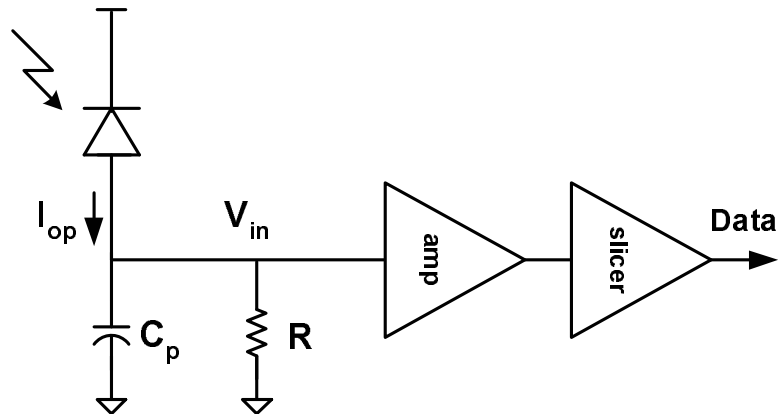


Figure 3.4: Simple resistive optical front-end

to increase the bandwidth is to add an equalizer following the voltage amplifier to compensate for the low-pass filter effect of the R and C_p . However, the design of the equalizer itself at high data rates is very challenging. As we will see in the next few sections, the current to voltage conversion and data resolution can be done in more efficient ways, either by using TIAs or integrating front-ends.

3.2.2 Transimpedance Amplifiers

The strong trade-off between the bandwidth and SNR 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 component to the design, resulting a transimpedance architecture. A transimpedance amplifier (TIA) is an analog front-end with reduced input impedance and a relatively high current-to-voltage gain. The addition of active components, like transistors, will add to the noise. However, with a careful design, very high SNRs are possible at the output of an optimized transimpedance amplifier. Detailed analysis of TIAs are covered in numerous publications [71] [72] [73]. In this section, we briefly discuss the performance and trade-offs of a number of different TIAs. For TIAs, like any other receiver, the most important specs are bandwidth, sensitivity, power consumption, area and dynamic range.

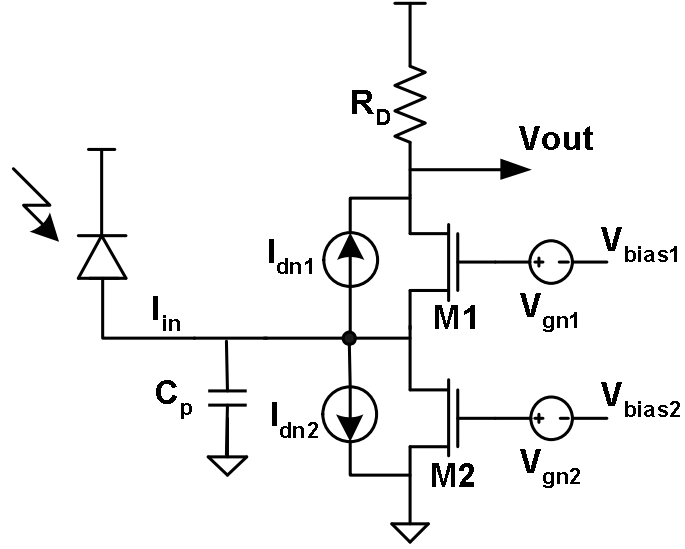


Figure 3.5: Common-gate TIA and its noise sources

Common-Gate TIA

In order to achieve a low input impedance and at the same time a high gain, one can use a common-gate (CG) topology. The common-gate TIA, shown in Figure 3.5 isolates the diode capacitance C_p from the gain resistor R_D and therefore has a wide bandwidth. The effective input impedance is $1/g_m$ of the input transistor³ M_1 , and the transimpedance is R_D . The noise sources for this topology is shown in Figure 3.5. The sensitivity of this front-end is less than optimum due to the direct additive noise of the resistor R_D , and the bias transistor M_2 . The equivalent input-referred current noise power spectral density is given in Eq(3.1). γ is the excess noise coefficient of transistor M_2 and for $0.25\mu\text{m}$ technology is about 2.5. The noise contributions from M_1 and R_D rises as frequency increases and the diode capacitance shunts the input node [74].

$$\bar{i}_n^2 = \bar{i}_{nd2}^2 + \bar{i}_{nr}^2 = 4kT\left(\frac{1}{R_D} + \gamma g_m\right) \quad (3.1)$$

³ g_m is the transistor transconductance.

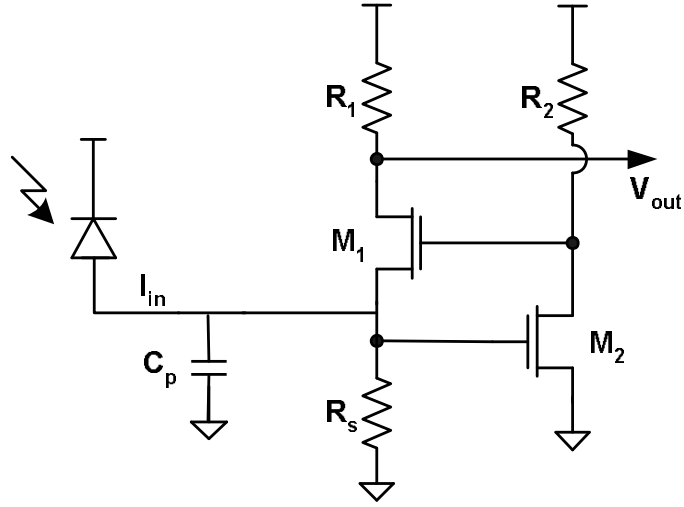


Figure 3.6: Regulated cascode TIA input satge

Regulated Cascode TIA

The CG input configuration relaxes the effect of large input parasitic capacitance on the bandwidth of the front-end. However, the poor device characteristic of MOS transistors such as small g_m , cannot totally isolate the parasitic capacitance. Also, this small g_m deteriorates the noise and stability performance of the amplifier. A regulated cascode (RGC) configuration addresses these issues and is used in some TIA designs [75] [76]. The RGC input mechanism enhances the effective transconductance significantly. As a result, the input node of the amplifier can sit at virtual ground and higher bandwidths are feasible.

Figure 3.6 shows the schematic diagram of the RGC circuit. The optically generated current is amplified to be a voltage at the drain of M_1 . The M_2/R_2 stage operates as a local feedback and thus reduces the input impedance by the amount of its own voltage gain. With a simple small-signal analysis, the input resistance of the RGC circuit can be approximated by Eq(3.2).

$$R_{in} \simeq \frac{1}{g_{m1}(1 + g_{m2}R_2)} \quad (3.2)$$

Clearly the input impedance is $(1 + g_{m2}R_2)$ smaller than the CG configuration.

However, the local feedback stage inherently produces a second pole, causing a peaking in the frequency response. In order to avoid this peaking and stability concerns, either the resistance R_2 or the gate width of M_1 should be reduced. Reducing R_2 decreases the input transconductance G_m almost linearly. 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. Reducing the width of transistor M_1 decreases G_m more slowly. However, it may lead to the increase of the channel thermal noise contribution from M_1 due to smaller g_m . Note that in RGC TIAs a second stage voltage amplifier usually follows this first stage RGC.

The noise performance of RGC stage is analyzed by Park in [77]. Similar to CG topology the noise of bias resistor and gain resistor directly contribute to 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 [75].

Common Source and Shunt-Shunt Feedback TIA

An alternative design with more relaxed noise-headroom trade offs is a shunt-shunt feedback TIA [64] [65] [71]. A TIA with resistive feedback, followed by a chain of post amplifiers is the most common type of receiver design and is shown in Figure 3.7. Some detailed analysis of this TIA is given in the Appendix A, using a simple model shown in Figure 3.8. While the transimpedance gain is R_f , the input impedance at DC is reduced to about R_f/A , where A is the gain of the voltage amplifier. In most designs the overall TIA bandwidth, BW , is limited by the pole at the input node as expressed by Eq(3.3) and not by the voltage amplifier. C_{in} here is the input capacitance of the amplifier.

$$BW \simeq \frac{A}{R_f(C_{in} + C_p)} \quad (3.3)$$

The sensitivity of the TIA depends on the total input-referred noise of the TIA and the noise of the diode itself. The TIA noise is mostly due to the thermal noise of the feedback resistor R_f and the input-referred noise of the amplifier. Appendix

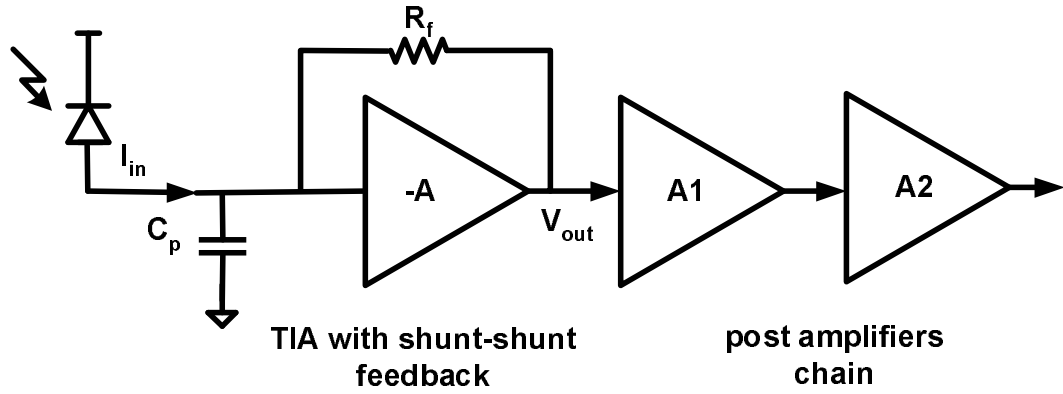


Figure 3.7: Shunt-shunt feedback TIA with limiting amplifiers

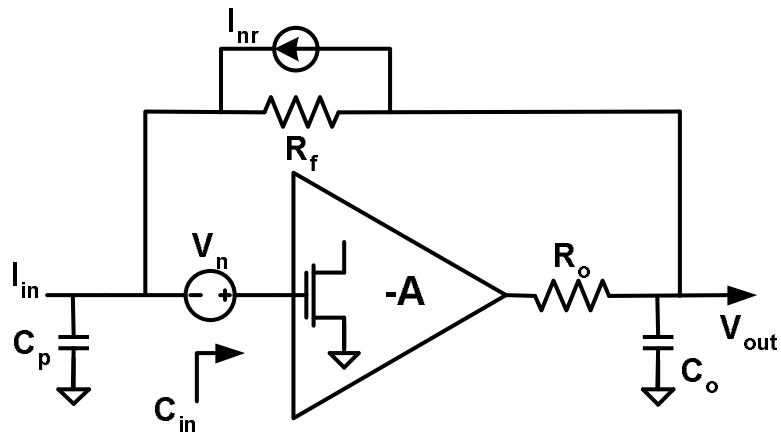


Figure 3.8: Shunt-shunt resistive-feedback TIA model

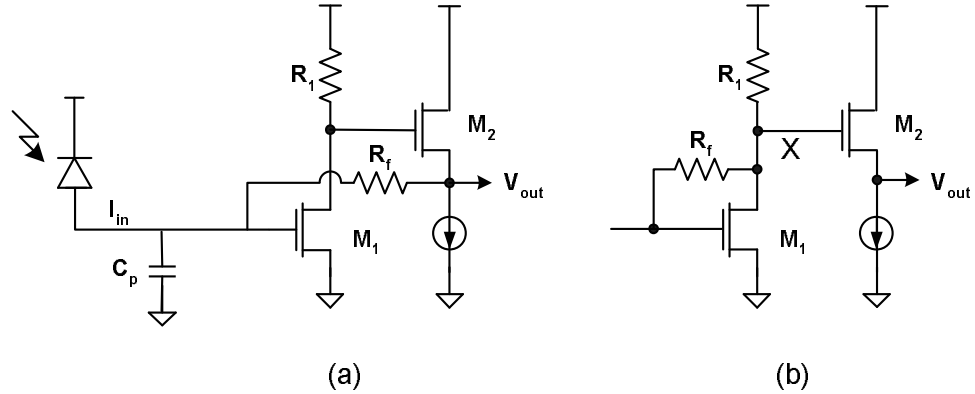


Figure 3.9: Common-source shunt-shunt feedback TIA designs

A provides a detailed discussion on TIA sensitivity. The sizing of the input stage of TIA can be optimized for the lowest input-referred noise [72]. In the optimized design the C_{gs} of the input common-source transistor turns out to be 0.5-1 times the diode capacitor C_p .

Two simple TIA designs with resistive feedback are shown in Figure 3.9. In the first design (a), 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 directly translates to small open-loop gain of the amplifier, higher noise and lower bandwidth. Moreover in this design if the output node is strongly capacitive, the introduction of another pole can cause stability and overshoot problems. The load capacitance can be isolated by adding an extra source follower in parallel to the existing one, which drives the output capacitor. The design shown in Figure 3.9 (b) allows a greater voltage drop across R_1 . In this second order system the bandwidth is maximized when the system is slightly under-damped. Therefore the pole at node X can be chosen to increase the overall bandwidth by up to 40%. Dynamic tuning of this capacitance might be necessary due to process variations.

For shunt-shunt TIAs, it is possible to use capacitors instead of resistors to implement the feedback [78]. Figure 3.10 shows a TIA design using capacitive network

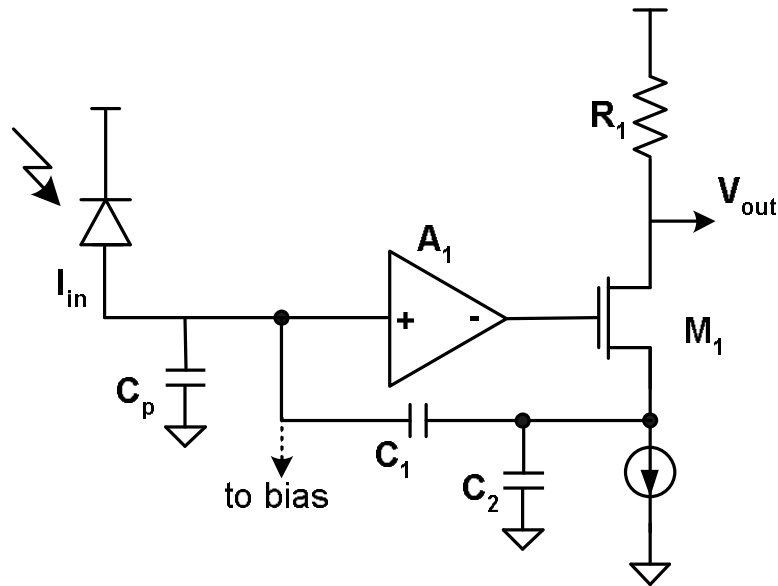


Figure 3.10: Capacitive network feedback TIA

feedback. In this topology, C_1 senses the voltage across C_2 and returns a proportional current to the input node. If $A_1 \gg 1$ then this design provides a transimpedance equal to $(1 + C_2/C_1) \times R_1$. The claimed advantages of using this topology are, no noise contribution from the gain defining elements (C_1, C_2), the capacitance seen at the input node only reduces the DC loop gain and does not degrade the stability of the TIA, and finally by correctly choosing the capacitance values, one can achieve lower amplifier noise contributions [78]. However, this design needs special biasing as shown in Figure 3.10.

3.2.3 TIA Design Consideration

All the TIAs discussed in the previous sections rely on a voltage amplifier that has high gain and high bandwidth. The high gain is required to achieve a low input impedance and low noise allowing higher values of R_f . As we mentioned before, the TIA bandwidth is usually set by the dominant pole at the input node. The amplifier creates the second pole, which should be far enough away to ensure stability and avoid gain roll off.

A popular figure of merit for high frequency performance of transistors is ω_T , a frequency at which the current gain of a common-source amplifier falls to unity. The most common expression assumes that the drain node is shorted and results $\omega_T \simeq g_m/C_{gs}$. On the other hand the gain-bandwidth product of a common-source amplifier with a load capacitance of C_L is equal to g_m/C_L . Since C_L is related to the sizing of common-source transistor (C_{gs}), we can claim that the gain-bandwidth product is approximately proportional to ω_T . Since the amplifier bandwidth is set by second pole it should be at least 3-4 times higher than the TIA bandwidth and one can simply say $gm/C_{gs} \simeq \alpha \times BW \times A$. α here is a constant, at least 3-4 and BW is the overall TIA bandwidth. BW should be at least 0.5 times the data rate to avoid dispersion at the output. Therefore high data rates require very high g_m and thus DC current in CMOS technology. The TIA current is roughly estimated in the Appendix A by Eq(A.12).

It is important to mention that so far we limited our discussion to single ended TIAs. However single-ended TIAs are very sensitive to supply noise and in most high-speed designs, a differential topology with high common-mode rejection is preferred [79] [79]. A differential scheme further increases the total power consumption of the TIA by almost a factor of two.

There are a number of different techniques proposed to increase the bandwidth of the TIA without sacrificing the sensitivity or increasing the current of the voltage amplifier. The second pole in TIAs is usually due to the output node of the voltage amplifier. Inductive peaking can be used to compensate the output node capacitance. This technique effectively increases the gain-bandwidth product of the amplifier. A drawback of this technique, particularly for the array application, is the significant area increase for the receiver as well as isolation issues. In [79] by using a common-gate topology the dominant pole is moved to the output node, thus the bandwidth enhancement achieved by inductive peaking is even more effective. As it was mentioned before, the common-gate topology degrades the sensitivity of the wide-band receiver by adding the channel noise of the bias and input transistor. Adjustment of the second pole (capacitive peaking) can also help to maximize the bandwidth [80]. In this technique, the second pole in the transfer function can be controlled by adding

an extra capacitance to the output of the first stage to achieve a damping factor (ζ) of 0.707.

Limiting Amplifiers

For the minimum input optical power, the voltage swing at the output of the TIA is usually very small. The voltage swing of a typical TIA is roughly calculated in Appendix A and can be less than 20mV. This voltage needs to be amplified for the following clock and data recovery (CDR) circuitry. In most designs, several stages of gain called limiting amplifiers follow the first stage TIA. Limiting amplifiers are usually a cascade of differential pairs with enough bandwidth and linear phase response. Limiting amplifier nonlinearity can introduce unwanted ISI and jitter to the output signal that goes to the CDR stage. The design of broadband limiting amplifiers with low supply voltage and limited headroom is also a challenging task. Inductive peaking shown in Figure 3.11, is used in many designs to meet the specs. Limiting amplifiers require careful design regarding stability and power supply noise. They also add to the overall power consumption of the front-end. The power consumption in the limiters is usually as high as the TIA itself [72].

3.2.4 Integrating Front-Ends

Integrating front-ends have been used to reduce the power consumption and area of the front-end by eliminating the need for TIAs with analog current/voltage amplification. In this type of front-end, the input impedance of the receiver is designed to be purely capacitive within the frequency range of the input data. The optically generated current from the photo-detector is then integrated onto the capacitor seen 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 will 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 period. The voltage swings should be enough for the receiver to correctly resolve the data. The sensitivity directly depends on the size of the capacitor C_{tot} . If a single diode is simply connected

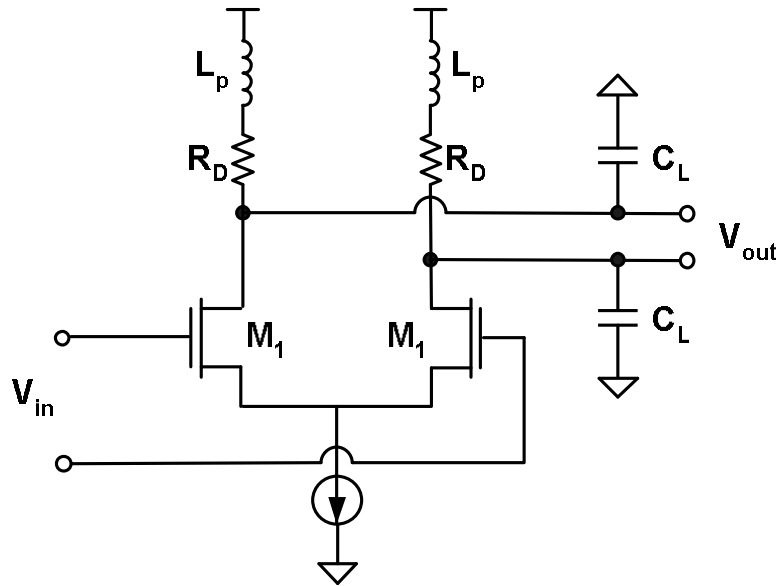


Figure 3.11: Inductive peaking for limiting amplifiers

to a capacitive node, I_0 and I_1 will be both positive or both negative. I_0 is usually small and kept close to zero. It is clear that upon 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" shown in Figure 3.12. Upon shining the light beams to each diode, one charges the input node, and the other one discharges this nodes. 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" or "recless" [68], is very simple and consumes very low electrical power. The required input modulation optical power for a full voltage swing is $P_{op} = C_{tot}V_{dd}/RT_b$ where R is the diode responsivity and T_b is the bit-time. The minimum optical power is proportional to C_{tot} , calling for very small photodiode capacitances as well as a small voltage buffer that follows it.

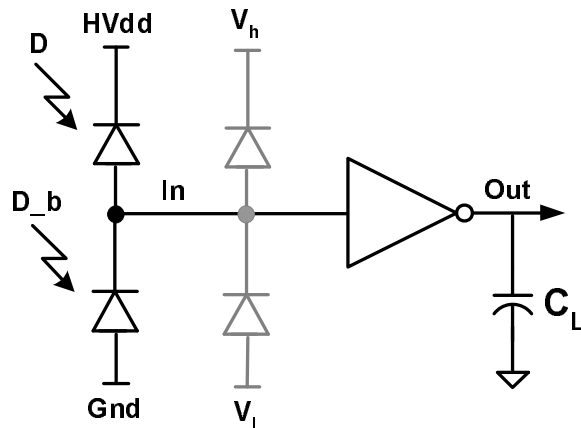


Figure 3.12: Receiver-less front-end with totem-pole and clamp diodes

If input optical signals are not short pulses and optical energy is uniform over the bit-period, for a fast rise and fall time at the input node, the optical power needs to be higher than the minimum value for full swing. In that case and upon receiving constant ones or zeros, the voltage of the input node may saturate to $V_{in} = V_{dd} + V_{int}$ or $V_{in} = -V_{int}$. V_{int} is the diode intrinsic or built-in voltage. For GaAs, this voltage is higher than 1V and may cause stress for the input transistor of the buffer stage, particularly at deep submicron technologies. In order to limit the voltage swing at the input node, one can add clamping diodes shown in Figure 3.12, [67]. The problem with diode-clamped design is the adjustment of the voltages of the protecting diodes to right values, as well as addition of extra capacitances to the input node.

The receiver-less approach has advantages, particularly for low noise timing and clock generation. A sharp rise-time can be created by using short-pulse lasers to charge and discharge the input node instantly. A short-pulse laser beam can have sub-picosecond width in the time domain. By minimizing the electrical circuitry at the front-end, the optical-to-electrical latency is reduced to the detector response time [70]. Depending on the design of the detector, the rise times can be less than 10 picoseconds, faster than the electrical FO4 delay of today's technologies. Since sharp rise times and large signal swings reduce supply noise sensitivity, this approach is suitable for generating low jitter multi-GHz clocks with precise skew control [81].

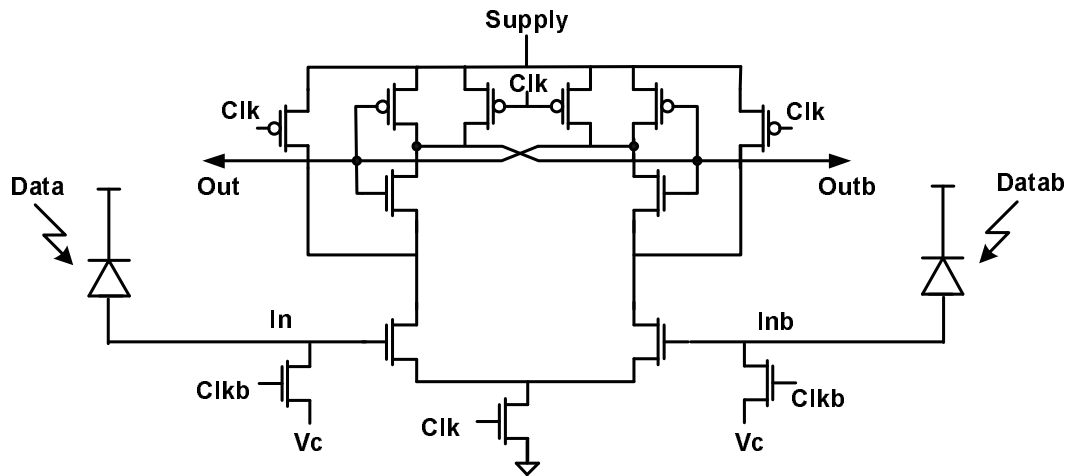


Figure 3.13: Sense-amp-based front-end

With the existing photodiodes, the receiver-less front-end needs a relatively high input optical power to generate a full swing voltage. The required voltage swing at the input node can be reduced by replacing the inverter stage with a sense-amplifier [69] [70] shown in Figure 3.13. The sense-amplifier-based front-ends either need two photodetectors and a differential pair of optical beams or a precise reference current, in order to resolve the data. For each bit-time, the receiver has two phases, integration phase and reset phase. The input optical power is used only during the integration phase, the data is evaluated and then both integrating nodes are reset to initial voltage. This calls for operation in a return-to-zero (RZ) data transmission scheme. A synchronous internal clock is used to define different phases, to integrate in half a bit-period and evaluate during the next half. The data rate is then limited to the on-chip clock frequency.

The sense-amplifier based front-end improves the sensitivity of the front-end compared to the receiver-less topology, and has lower power consumption compared to the TIAs. However, it requires differential beams as well as a reset phase (RZ data stream), which reduces the effective data rate of the system. In the next section we propose an integrating front-end to solve some of these problems.

3.3 Double Sampling/Integrating Front-End

As we mentioned in the previous section, integrating front-ends eliminate the need for a voltage amplifier that has a large gain-bandwidth product. Therefore, they potentially have lower power consumption and area. In this section we will show that by applying a different sampling scheme and adding simple circuits to the front-end, we can achieve high sensitivity and bandwidth. This receiver removes the need for an extra optical beam and resetting the input node. It also allows the use of time-division de-multiplexing with a single photodiode to achieve data rates much higher than the on-chip clock frequency.

3.3.1 Receiver Design Overview

As mentioned before, the photodiode capacitor C_p can integrate the optically generated current of the detector over time. If the input of the front-end receiver is also capacitive with a capacitance C_{in} , the voltage of the input node at the end of each bit-time, V_n , is always a sum of the incoming signal and the voltage of the input node just before that bit $V_n = V_{n-1} + I_{in}T_b/(C_p + C_{in})$, where T_b is the bit time and I_{in} is the effective input current to the receiver. I_{in} depends on the value of the input data and as we will explain can be set to be positive when the bit value is "one" and to be negative when the bit value is "zero". Therefore, if we have the voltage samples at the end of each bit-period, V_n and V_{n-1} , we have enough information about the input signal at time t_n to determine whether it was a one or a zero.

Figure 3.14 illustrates the top level block diagram of this receiver. The input signal from the photo detector is single-ended, with a positive current. The injected charge is higher if the bit value is "1" but it is not necessarily zero when the bit value is "0". Therefore, in order to have a bipolar voltage change at the input of the receiver we need to subtract a constant charge for every bit from the input capacitor. This is done by subtracting an adjustable current from the input. The DC current I_{DC} is adjusted by a feedback loop looking at the DC value of the voltage of the input node. The feedback loop not only adjusts the DC current but also sets the average voltage of the input node. A bipolar voltage change at the input allows us to decide

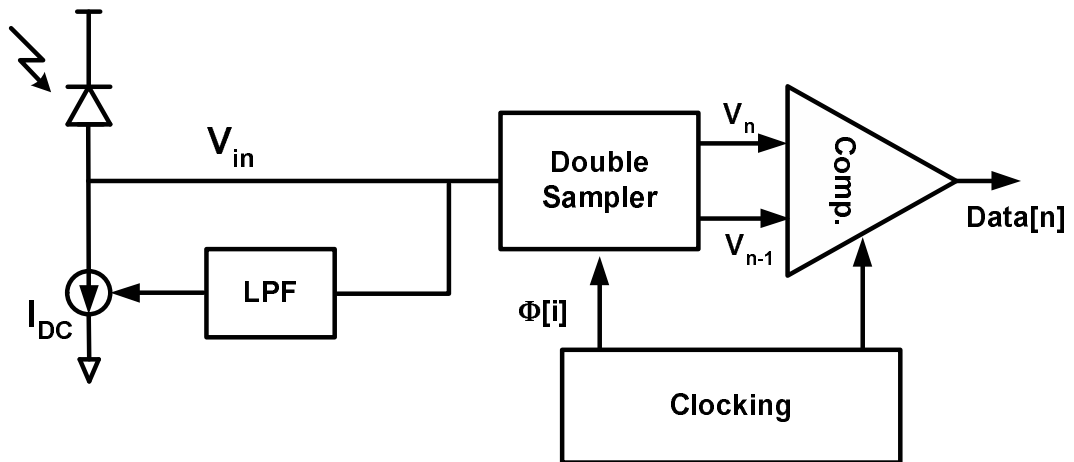


Figure 3.14: Block diagram of double sampling/integrating front-end

the input value by comparing two adjacent samples of the input voltage. If the new sample is higher, the input signal is "1", otherwise, it is "0". Figure 3.15 illustrates how V_{in} varies with time when I_{DC} is set correctly, assuming constant currents during each bit period.

Figure 3.16 illustrates a possible implementation of this front-end. Two non-overlapping clock phases, $\phi[0]$ and $\phi[1]$ perform the double sampling, which results in a bit rate twice the on-chip clock frequency. This design has only two samplers and each sampled value is used twice by the two comparators. The two interleaved comparators are triggered once every clock cycle, after every other sampling period. As we will show in the following sections this scheme works because the kick-back from the comparators cancel each other. It is also possible to use four different samplers and use each sample only once. The two-sampler scheme is preferred to minimize the power consumption and area of the receiver.

One can extend this double sampling front-end to use n clock phases to provide a de-multiplexing receiver with a bit rate n times higher than the clock frequency. A multi-phase oscillator, like a ring oscillator, can provide the required clock phases. The de-multiplexing front-end is shown in Figure 3.17. In this design, $n = 5$ different phases are used to increase the data rate to five times more than the on-chip clock frequency.

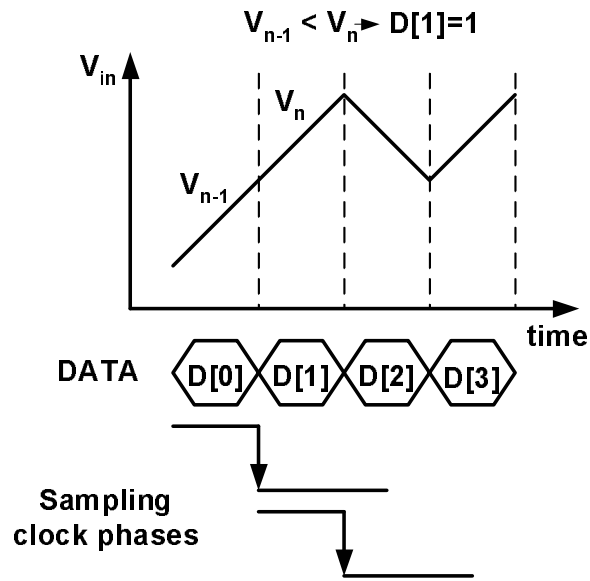


Figure 3.15: Data resolution and input voltage waveform of the integrating front-end

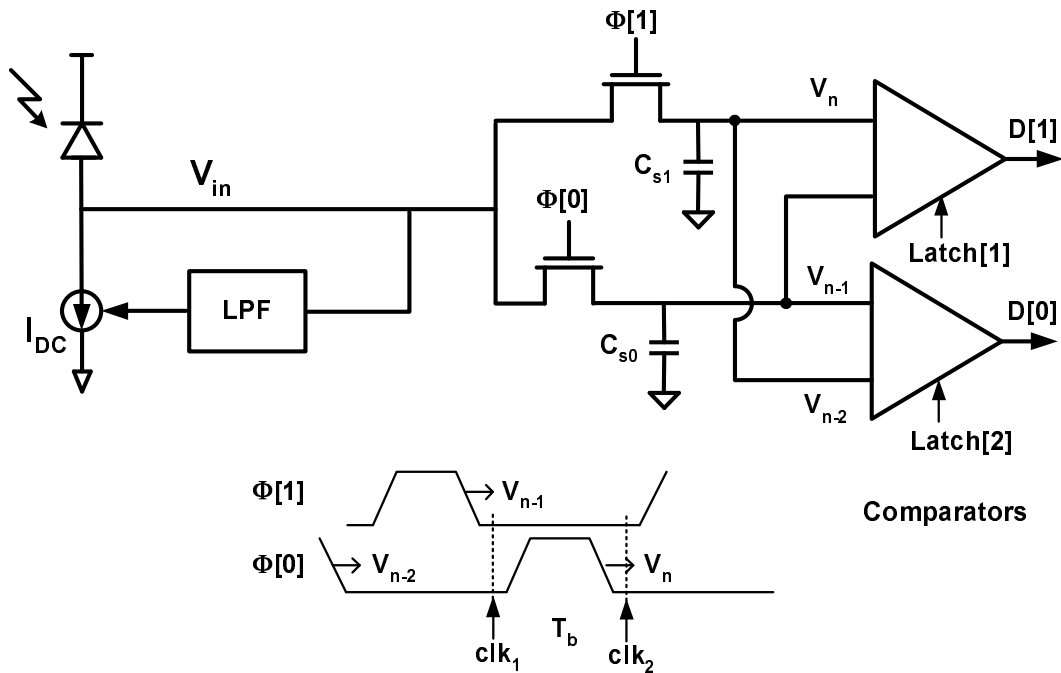


Figure 3.16: Double sampling/integrating front-end implementation

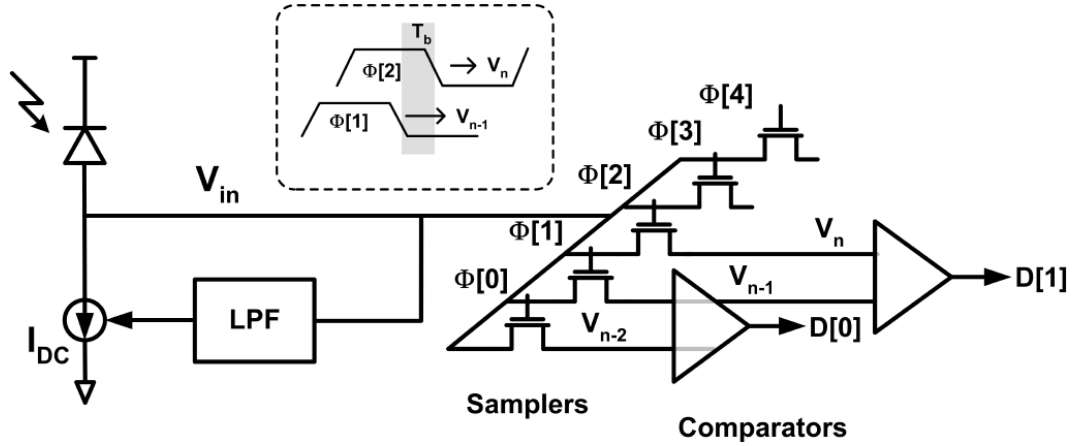


Figure 3.17: De-multiplexing double sampling/integrating front-end

With this brief introduction to the principles of the double sampling/integrating front-end, the next few sections are dedicated to details of the receiver implementation, its performance analysis and design optimization.

3.3.2 Sampling and Comparison

The double sampling is done with a pair of samplers driven by two different clock phases. The samplers are implemented by NMOS pass-transistor switches, which sample the voltage onto the sampling capacitor C_s and then hold it after the end of the bit-period. In this implementation the sample/hold (SH) capacitor is in fact the input capacitance of the next stage comparator as well as the parasitic capacitor of the NMOS switch. After holding the two samples, at the end of the second bit-period a clocked sense-amplifier is triggered to compare the two samples and resolve the bit value. The sense-amplifier in this design is a StrongArm regenerative latch [82], shown in Figure 3.18. The two sampling phases $\phi[1]$ and $\phi[2]$ are separated by one bit-period. Since the sense-amplifier is triggered after the falling edge of both $\phi[1]$ and $\phi[2]$, all the charge injections from the clocks to the sampling nodes are equal and act as common-mode signals.

The data rate in this design is limited by the speed of the samplers with time

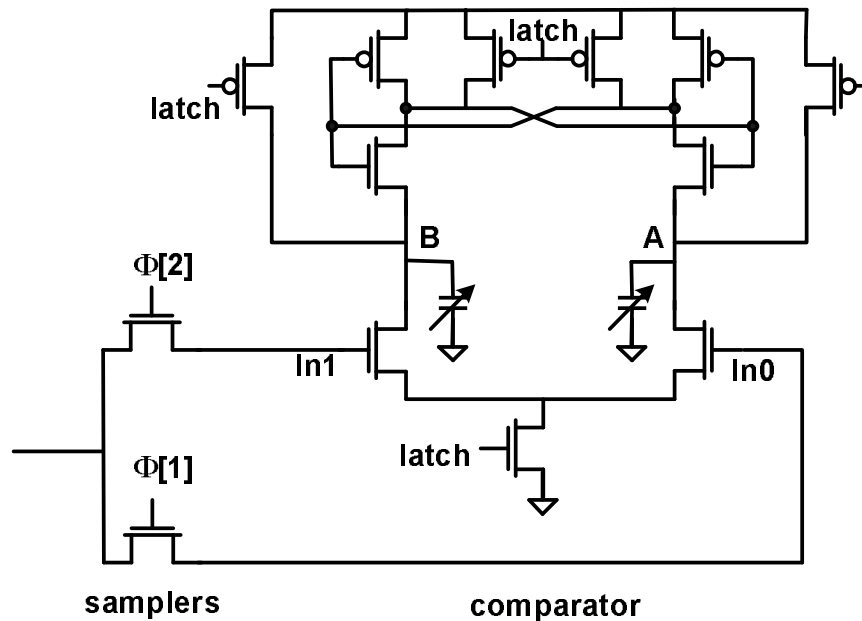


Figure 3.18: Double sampler and comparator circuits

constant of $\tau^{-1} = 1/R_s C_s$. The sampling capacitance C_s is dominated by the input capacitance of comparators connected to the sampling nodes, and the sampling resistance R_s is the ON resistance of NMOS transistor. Figure 3.19 shows the sampler rise time in response to a 10mV step voltage at its input as a function of initial input voltage level, with a minimum sized NMOS pass-transistor and $C_s = 15\text{fF}$ in a $0.25\mu\text{m}$ CMOS technology. It is clear from this figure that bit-times smaller than 200psec, less than 2FO4 inverter delay in this technology, are possible. This bit-time is much smaller than the minimum clock period that can be reliably generated and distributed on-chip. However a de-multiplexing technique, using multiple clock phases, allows such data rates. While smaller C_s can increase the speed of the sampler even more, as we will see it degrades the sensitivity of the receiver due to the sampler's thermal noise with voltage variance of $\sigma_{ns}^2 = kT/C_s$. Moreover, the input capacitance of the comparator has a lower limit, set by the speed and offset requirements of the comparator. The other limitation on the maximum value of n , the de-multiplexing factor, is the precision of on-chip multiple clock phases.

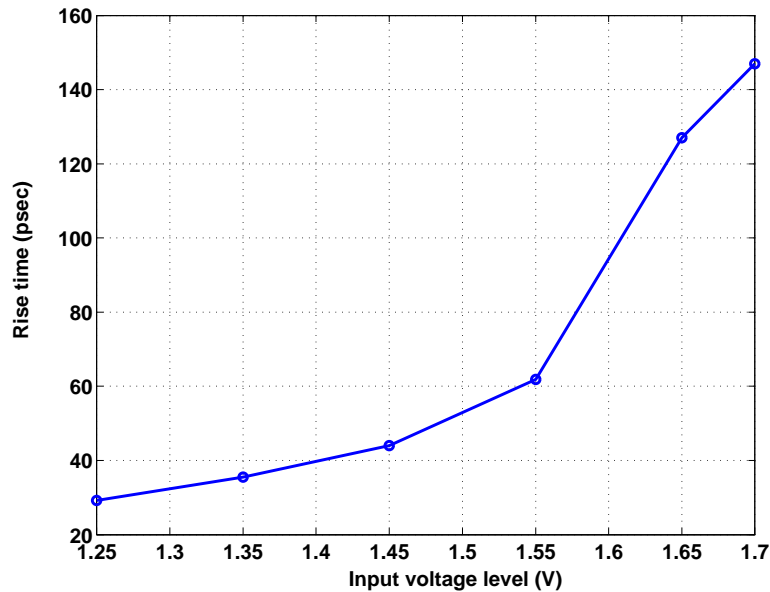


Figure 3.19: Rise-time of the NMOS sampler in response to a 10mV step voltage as a function of input voltage level

The dynamic range of this integrating front-end also depends on the behavior of samplers and comparator. Depending on the data pattern, the common-mode value for the two samples changes over time. Figure 3.20 shows the output voltage swing of a sampler in response to a 10mV step, in a $0.25\mu\text{m}$ CMOS technology with 2.5V supply, as the common-mode voltage increases. The threshold voltage of NMOS pass-transistor in this technology is about 0.6V. This figure shows that in order to avoid large attenuation on sampled signal the common-mode level of input voltage needs to be less than 1.6V.

The sense-amplifiers in Figure 3.18 need to resolve very small voltage differences and consume very low power. To achieve very high comparison precision, which directly improves the sensitivity of the receiver, the comparator offset and noise should be very small. The offset is dominated by the mismatches between the two input transistors and there is a direct trade off between the size of these transistors and the input-referred offset voltage [83]. For this receiver we use an offset compensation

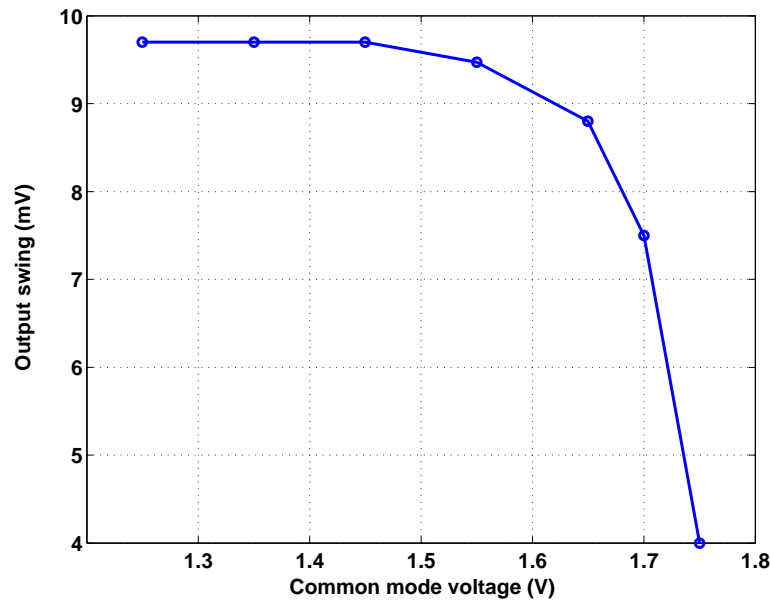


Figure 3.20: Output voltage swing of the sampler with a 10mV step voltage at the input as a function of input voltage level

technique to break the dependence of offset voltage on comparator transistor size, allowing smaller input devices and therefore lower power and area. Offset compensation is done by digitally adjusting the number of small capacitors added to the internal nodes A and B in Figure 3.18 [84]. Process mismatches between the two branches of the sense-amp, and also mismatches between the two branches of the sampler or between $\phi[1]$ and $\phi[2]$, are compensated at this stage. The offset correction capacitors are implemented by small PMOS transistors shown in Figure 3.21. Since nodes A and B are initially pre-charged to V_{dd} , by switching the *ctrl* signals in Figure 3.21 between V_{dd} and Gnd , each capacitor value is changed approximately between C_{ov} and $C_{ox} + C_{ov}$, the digital control number changes between ± 15 . Figure 3.22 shows the input-referred voltage differences for the comparator as we switch the control signals. The maximum voltage step is about 4mV and therefore the offset compensation precision is $\pm 2\text{mV}$. The maximum offset compensation range is about $\pm 60\text{mV}$, which covers the $\pm 6\sigma$ of the offset. Figure 3.22 also shows how input-referred offset changes with the common-mode voltage level. This behavior can affect the sensitivity of the

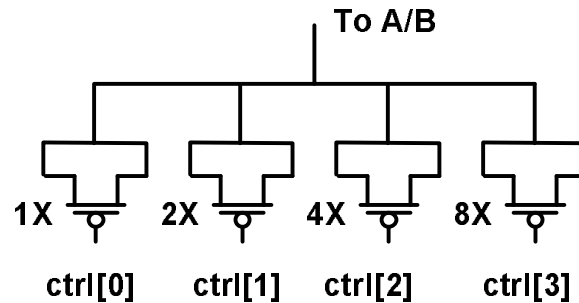


Figure 3.21: PMOS adjustable capacitors for offset compensation

receiver since the common-mode voltage at the input node can change due to the integrating nature of the front-end receiver. The change in offset is worst when the offset is maximum ($ctrl=15$ or -15). If the offset is canceled for $V_{cm} = 1.2V$, for a $\pm 200mV$ voltage swing, the offset can change $\pm 10mV$. The offset compensation also depends on the magnitude of V_{dd} , and the power supply noise can change the residual offset after the compensation. Our simulation results shows that these changes are relatively small but higher for larger initial offsets. For a $200mV$ change in the supply voltage, the maximum of offset compensation ($\simeq 60mV$) changes by less than $4mV$. This offset cancelation technique has a very small dependence on temperature. The offset changes less than $1mV$ for temperatures between $25^{\circ}C$ and $100^{\circ}C$.

The described sampling-comparison scheme is inherently robust against kick-back and charge injection from the comparators (sense-amps) to the high impedance input nodes. The reason is that there are two similar comparators that their inputs are connected to the same nodes of one sampler unit. Soon after one comparator injects some charge to these nodes during the evaluation phase, the other comparator is reset and injects the opposite charge to the same nodes. The total injected charge is zero after one bit-period and the sample is valid for the next comparison. The key point here is that the shape of the voltages of the precharged nodes, A and B are always very similar and therefore kick-back is not significantly data dependent.

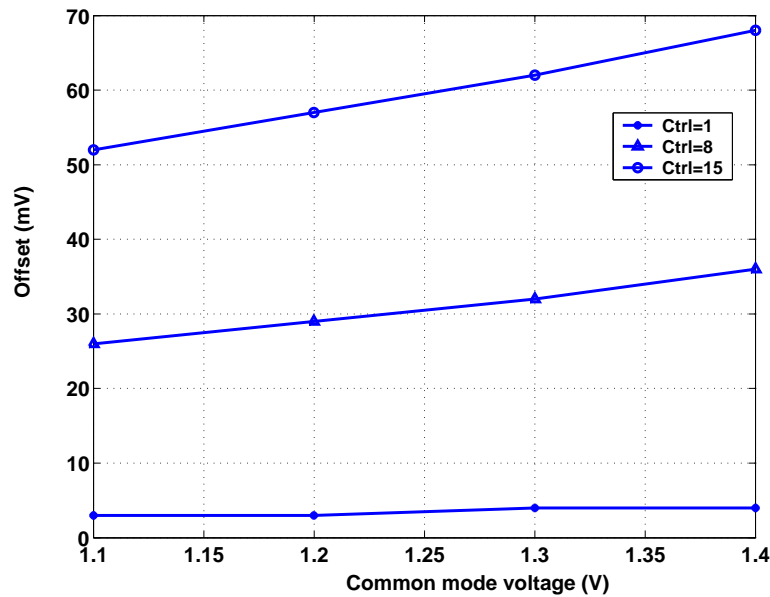


Figure 3.22: Input-referred offset as function control signals and input common-mode voltage

3.3.3 Filter and Current Feedback

Assuming that the stream of incoming data is DC-balanced, the DC voltage of the input node remains constant if I_{DC} , in Figure 3.17, is equal to the average optically generated input current: $I_{DC} = (I_0 + I_1)/2$, where I_0 is the average photocurrent during a "0" bit and I_1 is the photocurrent during a "1" bit. Note that I_0 and I_1 can vary due to variation of the optical input power and characteristics of the photodetector. If I_{DC} is any other value, the DC value of V_{in} will increase or decrease even after equal numbers of "0"s and "1"s. Therefore a feedback loop can be used to adjust I_{DC} by looking at V_{in} . The key is to use a low-pass filtered version of V_{in} to ensure the current does not fluctuate in response to the high frequency changes of V_{in} due to the incoming data. For instance, if we assume that data is DC-balanced within 32 bits, I_{DC} should be fairly constant even if we receive a row of 16 sequential ones or 16 sequential zeros.

The filter should also have a relatively high DC gain to be able to handle wide ranges of I_0 and I_1 values while keeping V_{in} relatively constant, at the best point

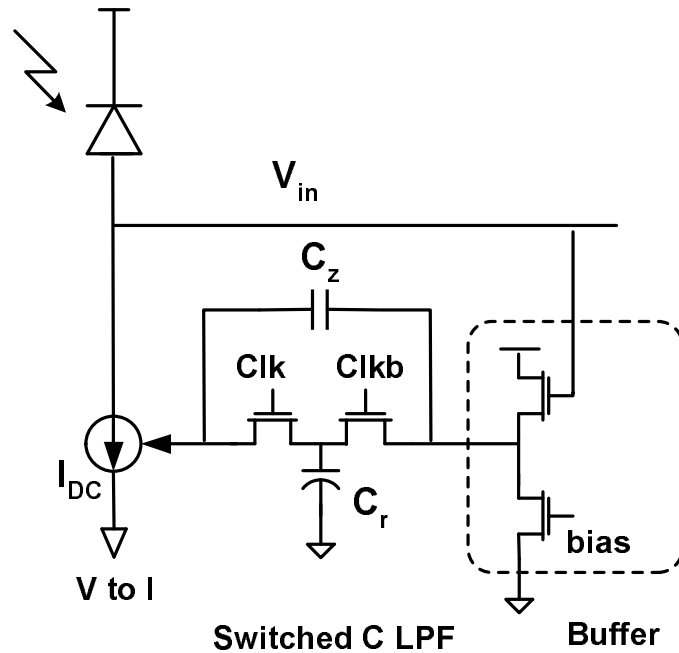


Figure 3.23: Feedback loop with low-pass-filter to adjust I_{DC}

of operation for the sampler and the comparators. The simplest approach to build the needed low-pass filter is a single pole RC circuit, but because of the capacitive nature of the input node, the open loop transfer function of this simple system will have two poles and it will cause an under-damped response in a feedback loop. One way to make the loop stable and increase the phase margin is adding a zero to the loop transfer function. Capacitor C_z in Figure 3.23 is added to the filter for this purpose. Resistor R is implemented by a switched capacitor, $R = 1/fC_r$. Where f is the frequency of non-overlapping clocks, clk and $clkb$.

In this design the data pattern that stresses the filter the most is the one with long sequences of ones or long sequences of zeros. In that case the input voltage has the highest voltage swing and the lowest frequency, thus the output voltage of the low-pass-filter and I_{DC} have the highest swing. Figure 3.24 shows the simulated percentage of peak-to-peak change in the I_{DC} as a function of DC-balance range of input data when the pattern has long sequences of ones followed by long sequences of zeros. The data rate here is 5Gb/s and the voltage swing per bit is 10mV. This graph

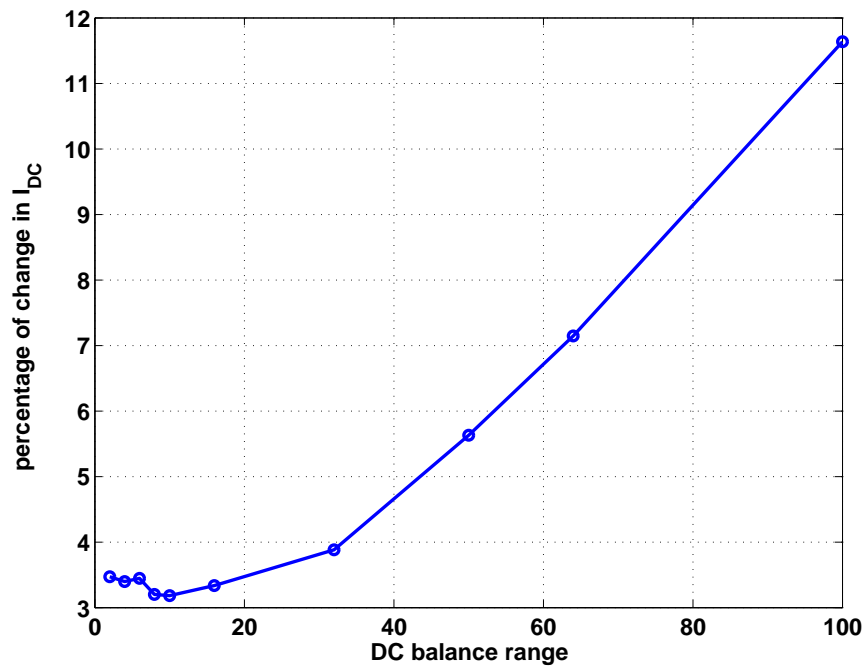


Figure 3.24: I_{DC} variations with the DC-balance range of input data

shows that for a 64 bits DC-balance range corresponding to 320mV input voltage swing at 78.125MHz, the ΔI_{DC} is about 7% and equivalent to 0.7mV voltage error. This error appears as an offset at the input.

It is also essential that I_{DC} can cover a sufficient range and the feedback loop has a relatively high gain to keep the DC value of the input voltage in the middle of the voltage range. The voltage to current converter in Figure 3.25 shows how the DC voltage value can be set initially for a typical I_{DC} . This is done by switching the signals $I_{set}[0, 1]$. The I_{DC} adjustment range is from $2\mu A$ to more than $300\mu A$. The DC voltage of the input node can be set by V_{set-in} since $V_{in} \simeq V_{set-in} + V_{gs1}$.

3.3.4 Supporting Circuits

To avoid hysteresis and to increase the sensitivity and speed of comparison a secondary small sense-amp/latch follows each of the first-stage sense-amps. The output of the latches are negative true pulses, which are converted into levels using a dynamic SR

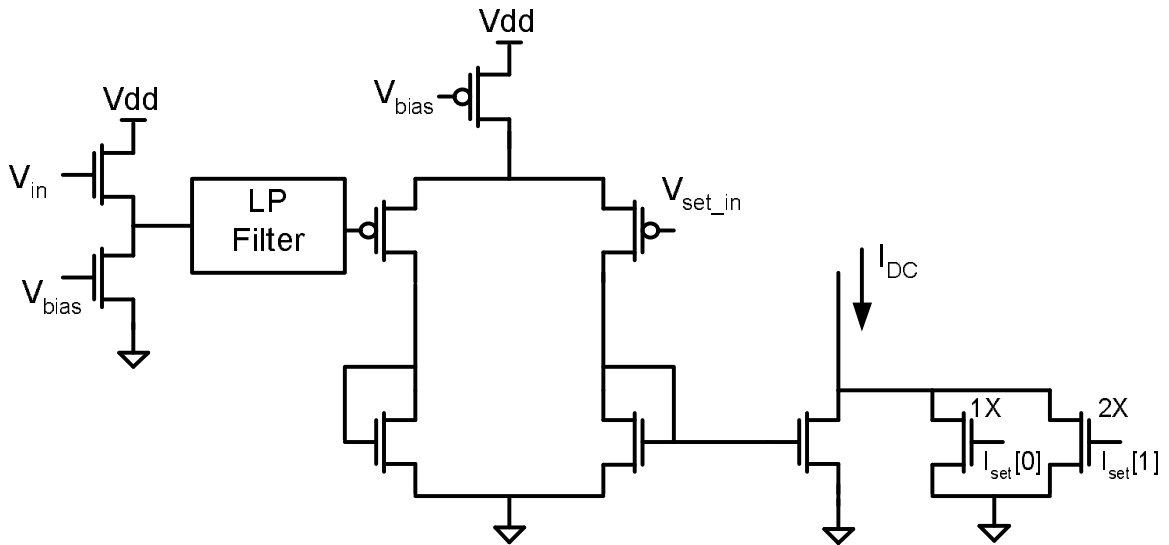
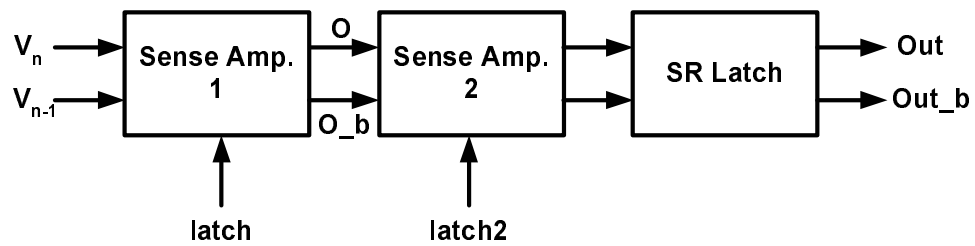
Figure 3.25: Voltage to current conversion for I_{DC} loop

Figure 3.26: Comparator and following stages of sense-amp and SR latch

latch, shown in Figure 3.26. Note that for systems where latency is critical, adding the extra stage increases the input to output delay by half a clock cycle.

In our first design only two clock phases are used for front-end. The sampling clocks need a duty cycle of less than 50% to have enough time for comparison as shown in Figure 3.16. The non-overlapping phases $\phi[0]$ and $\phi[1]$ and the trigger clocks of the comparators, $\text{latch}[0]$ and $\text{latch}[1]$ are all generated from the reference on-chip clock and its inverse, as illustrated in Figure 3.27. $\phi[0]$ and $\phi[1]$ are the chopped versions of Clk_b and Clk and their duty cycle can be adjusted with digitally controlled capacitors, C_{adj} . Clk_b needs to have same rising and falling rates as Clk and very low skew. The rising edge of $\text{latch}[0]$ is delayed by one inverter, therefore

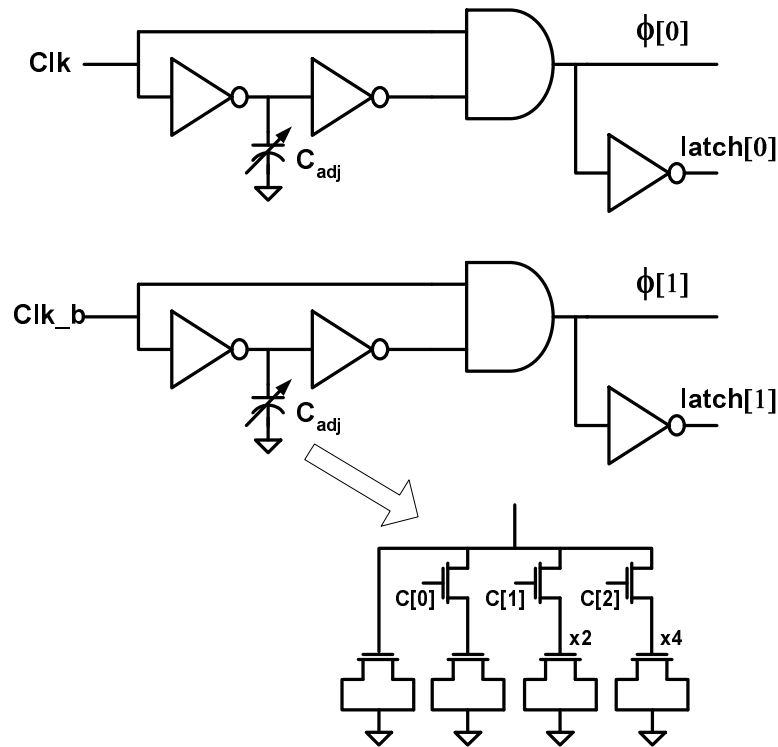


Figure 3.27: Chopping the clock for 1:2 de-multiplexing

right after sampling is done by $\phi[0]$, first comparator starts to evaluate its inputs. The evaluation should be done before the rising edge of $\phi[1]$. This condition is met because $\phi[1]$'s duty cycle is less than 50%. Clock jitter and phase offsets reduce the time left for the comparison and may require unpractical duty cycles. This problem limits the maximum clock frequency and data rate of the design. In the second version of the design, we used multiple clock phases, so the timing of the samplers/comparators can be managed without chopping the clock; see Figure 3.28. Multi-phase clock generation and timing recovery are explored in Chapter 5.

3.3.5 Performance Analysis

In the double sampling/integrating front-end, the data rate is inherently limited by the bandwidth of the samplers. As we mentioned before, the sampling capacitor

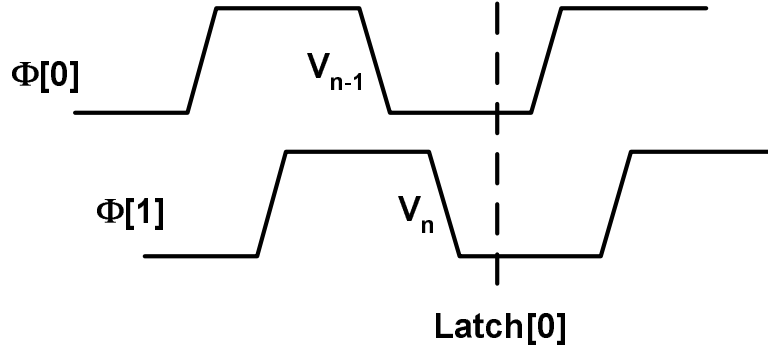


Figure 3.28: Multi-phase clocking, which automatically allows time for comparison

C_s is mainly the input capacitance of the comparators and R_s is the ON resistance of the NMOS sampler in Figure 3.29. In most systems, the possible bandwidth of the sampler is considerably higher than the frequency of on-chip clocks and with a de-multiplexing scheme higher data rates are possible. The comparator speed with a regenerative, positive feedback load, does not limit the overall speed as long as its delay is less than half a clock period. The comparator delay depends on the differential input voltage as well as the sizing of the transistors.

The minimum required optical power for this front-end depends on the photodiode responsivity R , the total parasitic capacitance C_{tot} at the input node, and the minimum required input voltage swing, ΔV_b , i.e.

$$P_{op} = R^{-1} \Delta V_b (C_p + \frac{n}{2} C_s) f \quad (3.4)$$

Here f is the bit rate and n is the multiplexing factor. The division by two is due to the fact that at any time only half of the samplers are ON. The minimum voltage swing per bit required for a certain bit error rate of the integrating receiver is set by the voltage noise and offset of the front-end. A minimum voltage swing V_{min} is also required for the comparator to resolve the output in less than half a clock period. Thus, in order to achieve a certain SNR for a target BER:

$$\Delta V_b = \pm(\sqrt{SNR} \sigma_n + V_{off} + V_{min}) \quad (3.5)$$

dominant capacitances are the two C_i s and the capacitance of the tail node of the comparator. These capacitances are proportional to C_i since the relative sizes of the devices are set by timing constraints. For our design, k_i is around 3. The transistor and capacitor sizing of the sampler and the clocked comparator are very important and set the sensitivity, electrical power and the bandwidth of the receiver. For the front-end analysis, we assume that the voltage noise due to the clock jitter ($\frac{\sigma_j \Delta V_b}{T_b}$) is negligible. Since the comparator acts by switching the internal capacitor C_i through its input transistor, the variance of the voltage noise at node A can be roughly approximated by kT/C_i . Note that the noise at node A affects the accuracy of the comparator the most among other noise sources. Therefore, the input-referred noise of the amplifier is about $A_c^{-2}kT/C_i$, where A_c is the voltage gain from the input (V_s) to point A in Figure 3.29. A_c is between 1 and 3 depending on the common-mode voltage.

With these assumptions, we effectively have a sample and hold circuit with an equivalent capacitance of C_s and $A_c^2 C_i$ in series. The total thermal noise can be expressed as $\sigma_n^2 = kT/C_{eff}$ where $C_{eff} = A_c^2 C_i C_s / (A_c^2 C_i + C_s)$. The required input optical power due to the random noise is then

$$P_{op} = R^{-1} \sqrt{kT \cdot SNR / C_{eff}} (C_p + \frac{n}{2} C_s) f \quad (3.7)$$

Figure 3.30 shows how required optical energy per bit changes as a function of C_s for different values of C_i and C_p . Increasing C_s up to the point that the total input capacitance is not increased significantly can help to reduce the optical energy by decreasing the kT/C noise. For $C_p = 200\text{fF}$, and a multiplexing factor of 5, the optimum value of C_s is around 15fF. As shown in Figure 3.30 the optimum value of C_s is about 2x smaller if C_p is reduced to 50fF. Although larger values of C_i can reduce the noise, smaller C_i is preferred for lower electrical power. In our design, we chose $C_i = 10\text{fF}$ as a compromise between optical and electrical power.

As we mentioned before, in order to generate a full swing NRZ output, a second stage sense-amplifier followed by a RS latch is added to the front-end. The power consumption of these stages plus the power of clock buffers driving both wires and

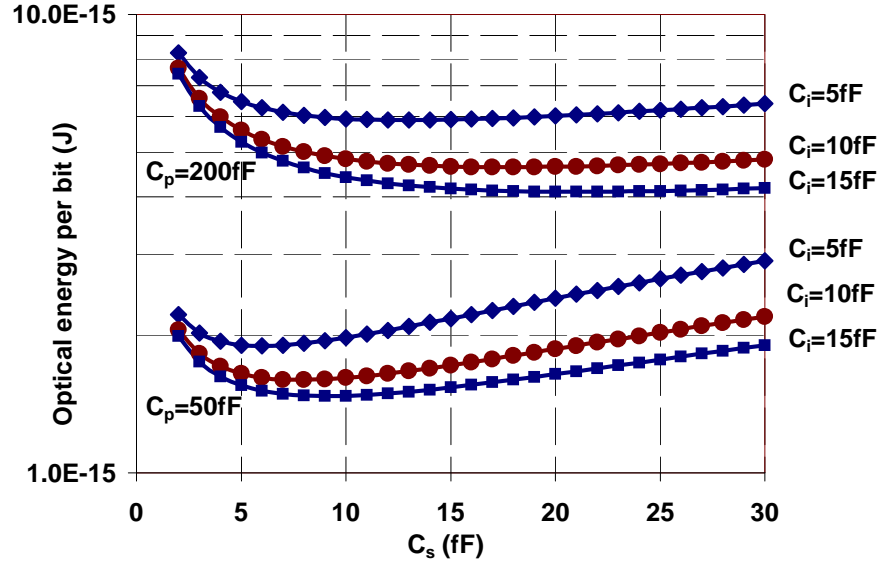


Figure 3.30: Required optical energy per bit versus C_s , C_i and C_p assuming $n = 5$, $R = 0.5\text{A/W}$, $SNR=36$ ($\text{BER} = 10^{-10}$), and $A_c=1$

transistors should be considered for total power of the front-end.

3.4 Receiver Testing

The dynamics of the voltage at the input node of the receiver is observed by adding an analog sampler [85] shown in Figure 3.31, to the this node. The samples of the input waveform generate a proportional current at the output of the sampler. If a periodic input pattern is sent to the receiver and the sampling frequency is not exactly equal to the frequency of the input pattern, over a long time, the sampler sweeps through many points of the input pattern. Therefore it is possible to reconstruct the input waveform, which gives an estimate of voltage swing per bit, ΔV_b , at the input node. Note that frequency of this sampler can be much lower than the data rate. The input periodic pattern is generated by the transmitter, which can be programmed to send a fixed 16-bit pattern or a pseudo random bit sequence (PRBS).

The replica of I_{DC} is also brought to a separate pin to estimate the average input photocurrent. If I_0 is set to be close to zero, the modulation current is also

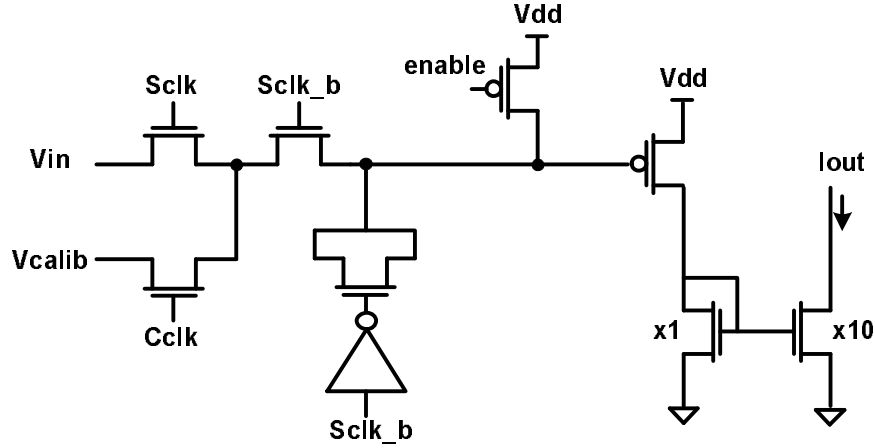


Figure 3.31: Analog sampler for monitoring the input voltage waveform

approximated by $\Delta I = (I_1 - I_0)/2 \simeq I_{DC}$. Having ΔV_b and ΔI one can easily estimate the total input capacitance of the front-end by $C_{tot} = (\Delta I \times T_b)/\Delta V_b$. The photodiode responsivity is also approximated by measuring I_{DC} and the optical beam power.

In this integrating front-end, the voltage swing headroom is limited by the operation range of the NMOS sampling transistors on one side, about 1.6 V in this technology, and the minimum common-mode voltage required by the comparators, about 0.8 V, on the lower side. The initial DC voltage of the input node can be set to about 1.2V to achieve a $\pm 400\text{mV}$ voltage swing due to the input data. Long sequences of ones and zeros in a row not only stress the voltage headroom but also create small ripples in I_{DC} as it is explained in section 3.3.3. Therefore, these patterns are the worst-case data streams for the receiver sensitivity and speed. In order to measure the sensitivity one can reduce the optical input power to hit the desired BER with a DC-balanced input pattern. Alternatively the sensitivity can be estimated by sending an unbalanced data stream using the fixed pattern transmission mode. With unbalanced input data, the I_{DC} is shifted in one direction and the voltage swing ΔV_b will be smaller in that direction, shown in Figure 3.32. Increasing the degree of unbalance up to the point of observing error in the data recovery can also provide an estimate of the receiver sensitivity.

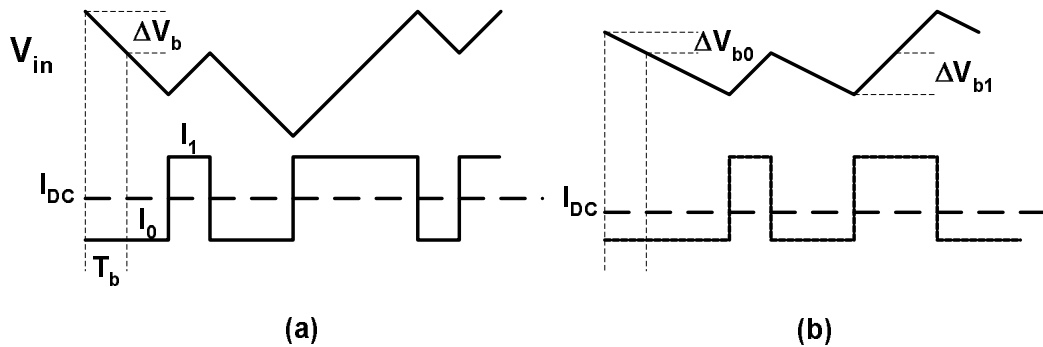


Figure 3.32: Effect of sending unbalanced input data stream to the receiver, (a) balanced input data (b) unbalanced input data, $\Delta V_b = 2\Delta V_{b0}$

3.5 Results and Performance Comparison

The first version of the receiver front-end was implemented in a $0.25\mu\text{m}$ CMOS technology. A block diagram of the top-level design is shown in Figure 3.33. An on-chip delay-locked loop takes a reference clock and generates an internal low-jitter clock for the samplers, comparators and the support circuits. In this initial design, the phase of the sampling clock is set manually by adjusting the phase of the external clock. Before the flip-chip bonding of the $p-i-n$ diodes and optical testing, this receiver is tested with electrical signaling. The electrical testing of the receiver is possible with on-chip current switches, modulated by the data, that mimics the input current from the photodiode. A row of MQW $p-i-n$'s is flip-chip bonded to the CMOS chip shown in Figure 3.34. This receiver was tested with the optical input data and the results are summarized in Table 3.1. The reported total power consumption, 3mW, includes the front-end sampler/comparator, the second stage Strongarm latch, the SR latch and local clock buffers at 1.6Gb/s data rate and with a 2.5V supply. As we mentioned before, in this design, the duty cycle of the sampling clock phases should be less than 50% to provide enough timing margin to trigger the sense-amplifier. The comparison should start after all the charge injections from the samplers are settled, and before the new sampling starts. Clock jitter and skew add to the required non-overlapping region, and call for very low duty-cycles. In this test-chip the data rate was limited by the minimum duty-cycle of a reliable sampling phase that was possible

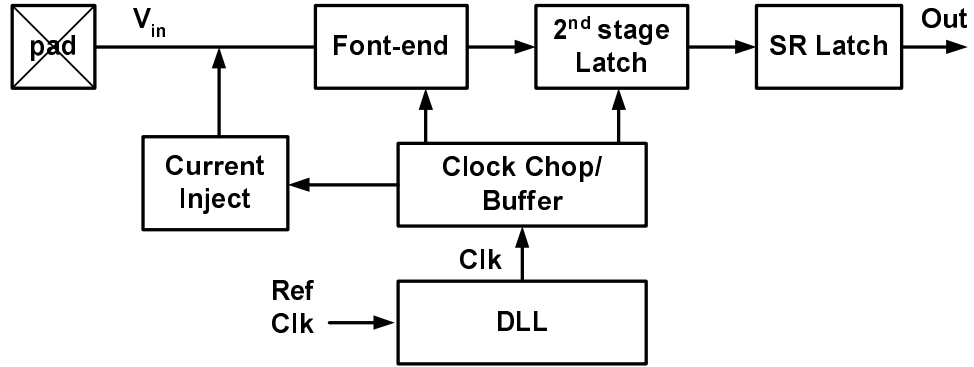


Figure 3.33: Top level block diagram of the first receiver test-chip

on-chip, which allowed about 150psec non-overlapping region and 1.6 Gb/s data rate. As we mentioned in Section 3.3.4, this problem can be removed if multi-phase clocks are used for the sampling/comparison as we did in our second test-chip.

In the second design a double sampling/integrating front-end was fabricated in a similar $0.25\mu\text{m}$ CMOS technology as part of a transceiver test-chip for parallel optical interconnections. Using five 1GHz clock phases, a de-multiplexing factor of $n = 5$ was implemented to achieve a data rate as high as 5Gb/s. The front-end was optimized for $C_p = 200fF$ and therefore the comparators are sized for $C_s = 15fF$ as shown in Figure 3.30, resulting in about 250fF total capacitance at the input node. For this design the input noise amplitude was measured by gradually increasing the offset and looking at the error rate. For a 10^{-10} error rate, the measured noise amplitude is $\pm 6\text{mV}$. The measured residual offset is only 2.5mV. In order to achieve a BER better than 10^{-10} , ΔV_b needs to be $\pm(6\sigma_n + V_{off}) \simeq \pm 9\text{mV}$, which corresponds to 4.5fJ optical energy per bit and $22.5\mu\text{W}$ of optical power for 5Gb/s data rate ($R = 0.5\text{A/W}$). With a 2.5V power supply, the electrical power consumption is 0.5mW for the first stage comparator, 0.4mW in the second stage sense-amplifier and RS latch, and 0.5mW in the clock buffers (about half of it in wires) at 1GHz clock. The samplers' power is in the order of μW and is therefore negligible. This gives a total of 7mW power for the five samplers/comparators needed to support a 5Gb/s data rate.

For a comparison, assume that a TIA is also implemented in a $0.25\mu\text{m}$ standard

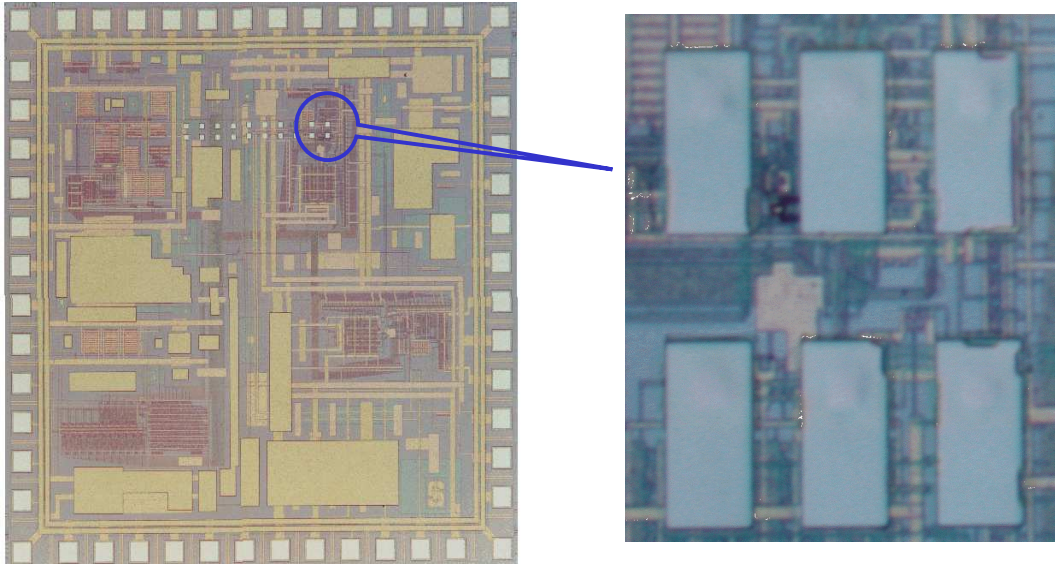


Figure 3.34: Fabricated test-chip with the flip-chip bonded devices

Technology	0.25 μm CMOS
Power Supply / Threshold	2.5 V / 0.55 V (NMOS)
Diode Capacitance	270 fF
Data Rate	1.6 Gb/s
Required Input Photocurrent	11 μA
Voltage Swing per Bit	$\pm 8.5\text{mV}$ (BER < 10^{-8})
Power Consumption	3 mW
Area Consumption	80 μm \times 50 μm

Table 3.1: Chip performance summary

CMOS technology. In order to achieve a 5Gb/s data rate, the overall bandwidth of the TIA (BW) should be at least 2.5GHz to avoid inter-symbol interferences (ISI) at the output. As it is explained in the Appendix A and mentioned in Section 3.2.3, for stability reasons the second pole, which is due to the voltage amplifier itself, needs to be at $3 \times 2.5 = 7.5\text{GHz}$ or higher. Therefore, for a gain of 4 in voltage amplifier, a gain-bandwidth product of around 30GHz is required. The ω_T then needs to be at least $2\pi \times 30\text{GHz}$. The TIA bias current depends on g_m , which is set by ω_T , and C_{gs} , which is close to C_p for an optimum noise performance. This bias current is estimated in the Appendix A, Eq(A.12), to be $\frac{\omega_T^2 C_p L^2}{2\mu}$. Replacing $L = 0.25\mu\text{m}$ and $C_P=200\text{fF}$, results in $I_D \simeq 6\text{mA}$. Therefore, the electrical power consumption with a 2.5V supply will be about 15mW for 5Gb/s data rate. Note that for a simple TIA, this result is optimistic since f_T of 30GHz is not trivial in most $0.25\mu\text{m}$ technologies because at high current densities g_m does not continue to increase with increasing I_D . However, bandwidth enhancement techniques such as inductive peaking can compensate for this. Note that a differential TIA will burn as much as two times the power of a single-ended TIA. Assuming the next stage limiter consumes about the same amount of power, for a 5Gb/s full swing output in a $0.25\mu\text{m}$ CMOS technology the front-end consumes more than 40mW of power. On the other hand, TIA noise analysis (in Appendix) shows that the sensitivity of the TIA can be in fact up to two times better than our double sampling front-end. The double sampling front-end, which requires a very small voltage swing at the input node, has a significantly better sensitivity compared to the other integrating front-ends discussed in Section 3.2.4.

3.6 Summary

The power consumption and area of receiver front-end circuitry is a critical design aspect of parallel optical interconnects for future arrays, where 1000s of beams are sent to a single chip. While transimpedance amplifiers have a relatively high sensitivity, the power consumption and stability issues of the TIAs make them less than optimal for the array applications. Integrating front-ends can reduce the power consumption by avoiding an analog amplifier that runs at the bit-rate.

This chapter described a double sampling integrating front-end that has a low power consumption, high data rate and relatively good sensitivity. The high data-rate is possible by a de-multiplexing scheme where multiple samplers and comparators run in parallel, while using a single photodiode. The double sampling technique provides a self-reference for the data resolution and a negative feedback loop subtracts a DC current from the input that effectively controls the voltage of this node and allows a bipolar voltage change for zero/one bits.

The double sampling/integrating receiver discussed in this chapter was implemented in two different versions, both using National Semiconductor's $0.25\mu\text{m}$ CMOS. In the first design a de-multiplexing factor of two supported a 1.6Gb/s data rate with a 2.5V supply and consumed only 3mW s of power. An on-chip delay-locked loop generated a low jitter clock signal for the receiver, while the clock phase was manually adjusted to for the best sensitivity result. In the second test-chip with the array of the receivers on a single chip, we achieved a data-rate as high as 5Gb/s in a similar $0.25\mu\text{m}$ CMOS technology. The higher data rate is possible by a 1:5 time-division de-multiplexing scheme using five different clock phases. A clock recovery phase-locked loop, described in Chapter 5, generated the synchronous multi-phase clocks.

Chapter 4

Scaling and Common-Mode Control

CMOS has been the leading technology for building integrated circuits for many years [6]. The number of MOS transistors per chip almost doubles and the speed of CMOS gates improves by $\simeq 40\%$ every three years [7]. The shrinking feature sizes and lower supply voltages reduce the power consumption of most processing blocks while they run at a faster speed.

Reduction in power consumption and area of transceivers allow higher numbers of IOs per chip. The combination of higher number of IOs and higher data rates per IO leads to a huge improvement in the overall chip-to-chip bandwidth. This chapter examines how the performance of the optical receivers will scale with advances in CMOS technology. We look into the challenges and problems introduced to the double sampling front-end design in deep submicron CMOS. In order to alleviate some of the problems associated with scaled power supplies, we propose a decision-directed common-mode control technique for this receiver. The scaling of TIAs is also discussed briefly in this chapter for comparison purposes.

4.1 Double Sampling Front-End Scaling

The removal of the high-speed gain stage in the proposed double sampling integrating front-end, makes it a promising candidate for CMOS technologies that commonly have a poor gain-bandwidth product. While the digital behavior of the front-end prompts a good scalability, there are many subtle design issues as we move to advanced technologies. In this section, we examine the scaling behavior of the double sampling/integrating front-end with respect to sensitivity, bandwidth, power consumption and dynamic range.

Sensitivity

Our discussion in Section 3.3.5, and Eq(4.1) show that the required optical power for this receiver, strongly depends on C_p and C_s , the diode capacitance and the sampling capacitance. If these capacitances stay constant, the input optical power increases linearly with the data rate as technology scales. This effect is particularly problematic as the array of transceivers becomes very dense. If we use modulators at the transmitter side, if the optical power per receiver P_{op} increases with the data rate, the power of the external laser should increase even faster than the number of beams on the chip. On the other hand, if we use VCSELs, the current of the laser should increase and the electrical power consumption of the transmitter will not scale as fast. Figure 3.30 shows that the optimum value for C_s reduces only if the diode capacitance is reduced.

$$P_{op} = R^{-1} \sqrt{kT \cdot SNR / (C_s \parallel A^{-2} C_i)} (C_p + \frac{n}{2} C_s) f \quad (4.1)$$

Recent advances in design of photodiodes promise diode capacitances as low as 50fF. This helps to keep optical power constant for few generation of technologies after our 0.25 μm CMOS and C_p of 200fF. As we mentioned before, scaling of diode capacitance is not trivial because of the other trade-offs. The capacitance is proportional to A_d/w , where A_d is the area of the diode and w is the width of i region. Decreasing A_d leads to possible optical power loss due to the imperfect focusing. On

the other hand, increasing w reduces the field in the i region and slows down the photodiode¹. If the scaling of diode capacitance slows down, the sizing of the front-end transistors does not scale as technology scales and P_{op} increases with the data rate.

Data-Rate per Receiver

If the feature sizes in CMOS technology scale down with rate α , the switching speed of transistors increases almost linearly with α^{-1} . The fundamental limit on the bandwidth of the double sampling/integrating receiver is the aperture time of the samplers. For higher data rate the $R_s C_s$ time constant should scale down, where R_s is the ON resistance of the NMOS switch. For the optimum sensitivity we need to keep capacitance C_s constant. Thus, for higher data rates the resistance R_s should scale down. This is possible by keeping the width of the pass-transistor (in microns) constant, while the length is equal to the minimum channel size of the technology. As technology scales, the transistor channel length scales down and so does R_s . Similarly the on-chip clock frequency for generation of multi-phases increases and allows higher data rates.

Power Consumption

The power consumption in high-speed IOs is the most crucial problem today. Increasing the number of IOs per chip is possible only if the power consumption per IO reduces. In digital systems the power consumption is dominated by the dynamic power for switching internal capacitances, $P = CV^2f$ as well as the leakage dissipations. As technology scales, the dynamic power consumption of a similar block reduces as almost α^2 . This is because the capacitances reduce with α , the power supply reduces with α , and the frequency increases with α^{-1} .

For our front-end, the capacitances of the samplers/comparators stop scaling down when the scaling of the diode capacitor stops. Therefore, the power consumption of the front-end scales only as α . However, The power consumption in the following stages (second sense-amplifier and SR latch), wires and clock generation circuits scale

¹However, new device structures are subjects of ongoing research, and may allow capacitances as low as a few fF.

as α^2 . As we will see in Chapter 5, in our $0.25\mu\text{m}$ technology, more than 75% of the total power is consumed in the PLL and clock buffers, resulting a faster scaling than α in the overall power. But eventually the power of the front-end will dominate. A simple scaling to 90nm technology with 1.0V power supply and constant diode capacitance, results in a power break down of 40% in the front-end and 60% in the CDR. Therefore, compared to our $0.25\mu\text{m}$ technology, the feature sizes are scaled by $\alpha = 0.36$ and the power is scaled by $P_{0.25}/P_{90} = 0.2$. With 90nm technology the total power of receiver reduces to about 1.0mW per Gb/s. If scaling continues with the same trend, at 65nm and 0.65V supply the power of front-end and CDR will be equal. After this point, the front-end will burn more power than CDR. It is important to mention that as technology continues to scale, supply scaling will slow down when V_{dd} reaches around 1.0V. When this occurs, front-end power will increase with data rate and CDR power stays almost constant.

Overall Bandwidth

The overall interconnection bandwidth R_{tot} depends on the data-rate per IO, R_0 , and N , the number of IOs per chip, $R_{tot} = N \times R_0$. Some of the limiting factors for the number of IOs per chip are **(i)** power consumption, **(ii)** area, **(iii)** beam density and cross-talk, and **(iv)** minimum pitch and yield in the array of optical devices. The power consumption of the double sampling/integrating front-end reduces as α^k where $1 < k < 2$ and allows a proportional increase in the density of IOs until the device fabrication and focusing problems put a hard limit on the density. Since the data rate per IO increases as α^{-1} , the overall chip-to-chip bandwidth increases as α^{-k-1} . For a 90nm CMOS technology, R_0 can be as high as 15Gb/s and with 1000 IOs per chip we achieve 15Tb/s, while the chip consumes 30W of power (30mW per IO)².

Dynamic Range

The proposed integrating front-end relies on a DC-balanced input data pattern. The balance length N_{DC} is limited by the allowable voltage swing at the input node ΔV_{in} ,

²As a rough approximation we have assumed that the power of the transmitters and receivers are about the same.

as well as ΔV_b , the voltage swing per bit, $N_{DC} = \Delta V_{in}/\Delta V_b$. For a certain ΔV_{in} this requirement puts an upper limit on ΔV_b and therefore on the maximum input optical power. The input power to the receiver can be reduced by sending a feedback signal to the transmitter, requesting a lower transmit power.

As technology scales, the kT/C noise requirements does not allow scaling of ΔV_b to lower values for a constant BER. On the other hand due to the scaling of V_{dd} , ΔV_{in} decreases as technology scales. This effect calls for a lower value for N_{DC} . For a 90nm technology with $V_{dd} = 1V$, ΔV_{in} is less than 400mV and maximum N_{DC} can be as low as 10-20 bits. The overhead percentage of DC-balanced codes increases as the balance range decreases. For instance an 8b/10b code is balanced over 10 bits but has only 8 bits of information with 20% overhead. In the next section we propose a control loop that can reduce the dependency of N_{DC} on V_{dd} .

4.2 Decision Directed Current Control

In our integrating receiver, the voltage of the input node V_{in} goes high upon receiving "ones" and goes down upon receiving "zero" bits. The voltage swing can be controlled if after the data is resolved, an equal amount of charge is added or subtracted from the input node. If the resolved data value is a "one", a positive current should be subtracted, and if it is a "zero", a positive current should be added to exactly compensate for the input photocurrent. These feedback currents are called I_{f1} and I_{f0} in Figure 4.1. The loop delay for subtracting these currents is m bits in this figure. Assume that the input photocurrent is I_1 or I_0 upon receiving a one or a zero. If I_{DC} is set to the average of these currents and the two correction currents are set to be $I_{f1} = I_{f0} = (I_1 - I_0)/2$, the voltage waveform at the input node would look like the graph shown in Figure 4.2, where m is predicted to be 4.

In this design, if the received bit at time n is equal to the bit at time $n - m$ the new voltage swing per bit is zero, while if they are not the same the voltage swing per bit will be two times the voltage swing per bit with no correction loop, $2\Delta V_b$. The maximum voltage swing at the input node is now $\Delta V_{in} = 2(m - 1)\Delta V_b$. The swing does not depend on N_{DC} anymore and instead it depends on the delay of the

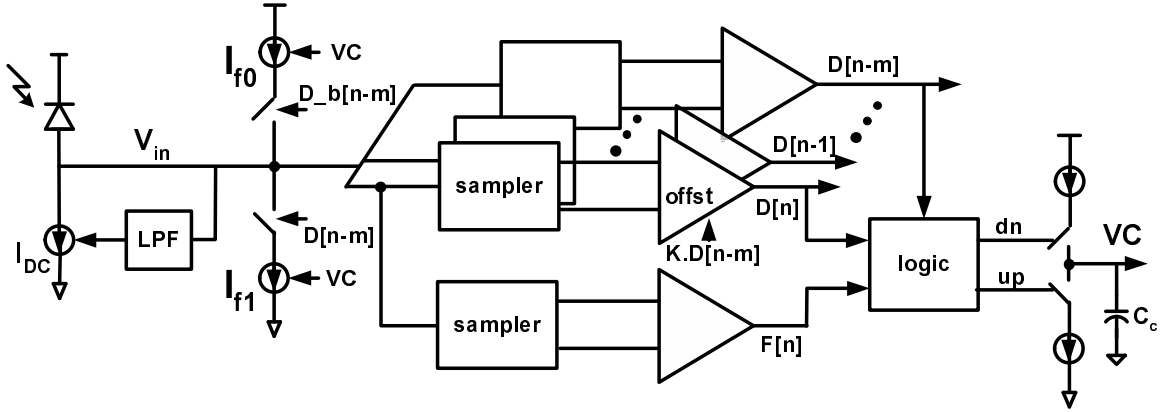


Figure 4.1: Block diagram of decision-directed common-mode control

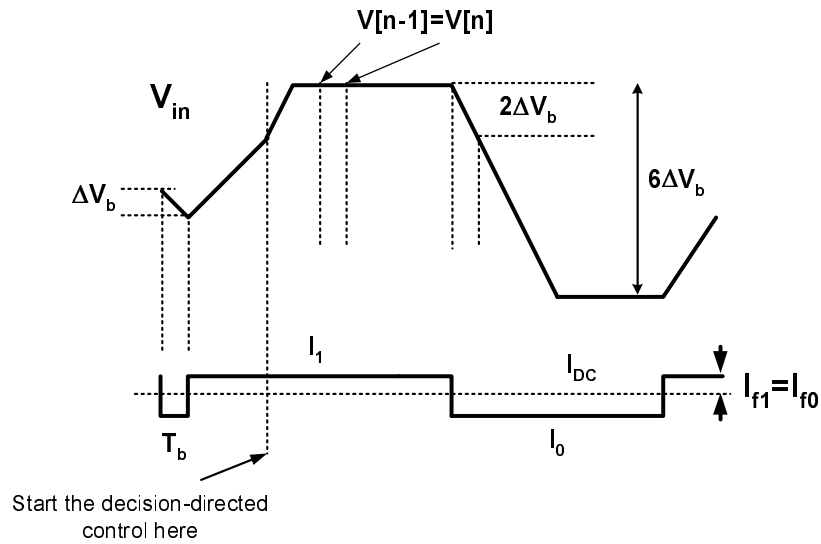


Figure 4.2: Input voltage waveform with the decision-directed current control

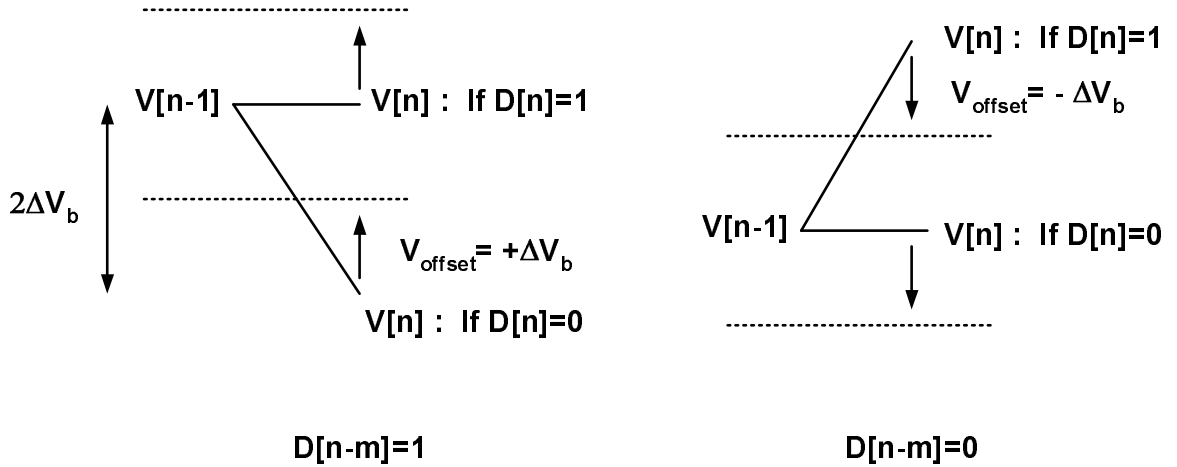


Figure 4.3: Data resolution for DDCC by adding offset, four possible cases for different values of $D[n]$ and $D[n-m]$ are shown

correction loop.

With this common-mode control method, the double sampling data resolution scheme needs a slight modification to work. As shown in Figure 4.2, if the current added to the input node is equal to the current subtracted, the two consecutive samples $V[n]$ and $V[n-1]$, are equal. In this situation the data resolution for $D[n]$ is possible only if an offset voltage equal to ΔV_b is induced to the sense-amplifier comparator, in a direction that is determined by $D[n-m]$. The detail of data resolution with the induced offset is shown in Figure 4.3.

The two current sources I_{f1} and I_{f0} need to be dynamically adjusted and track the incoming optical power. We can employ a feedback loop to control the correction currents. Whenever $D[n] \oplus D[n-m] = 1$, if the currents are set correctly, the two consecutive samples $V[n]$ and $V[n-1]$ should be equal. The difference between these values can be used as an error information for the current adjustment loop. By adding an extra branch of double sampler/comparator to the front-end with a constant zero offset, we can compare these two consecutive samples. The output $F[n]$ is used whenever the condition of $D[n] \oplus D[n-m] = 1$ is satisfied to drive a bangbang-controlled loop as shown in Figure 4.1. A simple logic block determines the direction of up/dn command by looking at $D[n]$, $D[n-m]$ and $F[n]$. The loop filter

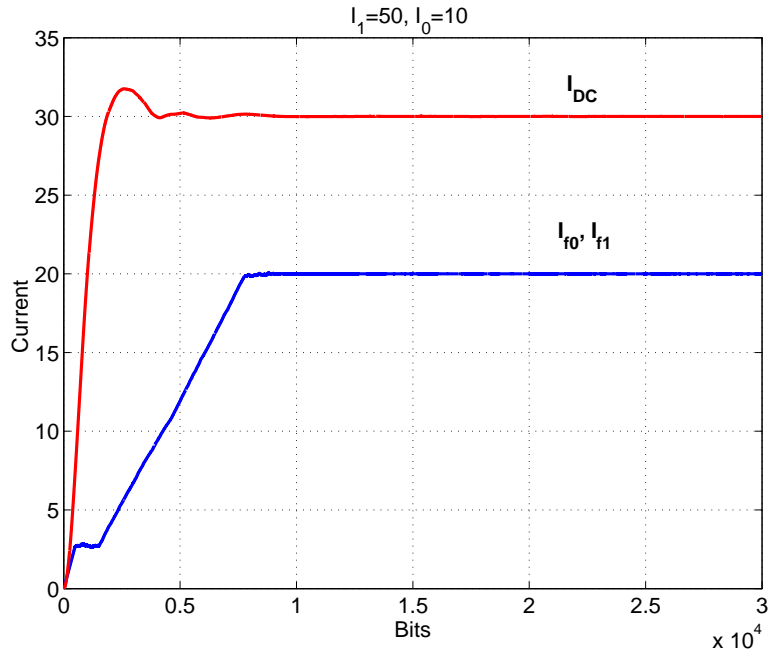


Figure 4.4: Simulated loop dynamic for DDCC

can be a charge-pump with a capacitor C_c creating a single pole/integrating transfer function. The output of the charge-pump is then used to control the two correction currents.

One interesting issue is the interaction of the common-mode control loop and the loop that adjusts I_{DC} . Figure 4.4 shows the simulated dynamics of the two loops when the decision-directed common-mode control (DDCC) loop is set to be slower than the I_{DC} control loop. As we mentioned before, the I_{DC} control loop is of second order and has a compensation zero. In Figure 4.4 the input photocurrents are set to $I_1 = 50\mu\text{A}$ and $I_0 = 10\mu\text{A}$. The simulation results confirms that the maximum voltage swing is confined by $2(m-1)\Delta V_b$.

In this design both I_{f1} and I_{f0} are controlled by the same voltage VC . Since we have two different loops adjusting the overall currents at the input node, any offset between I_{f1} and I_{f0} is compensated by I_{DC} control loop. In fact we can simply remove I_{f1} which will be incorporated into I_{DC} automatically.

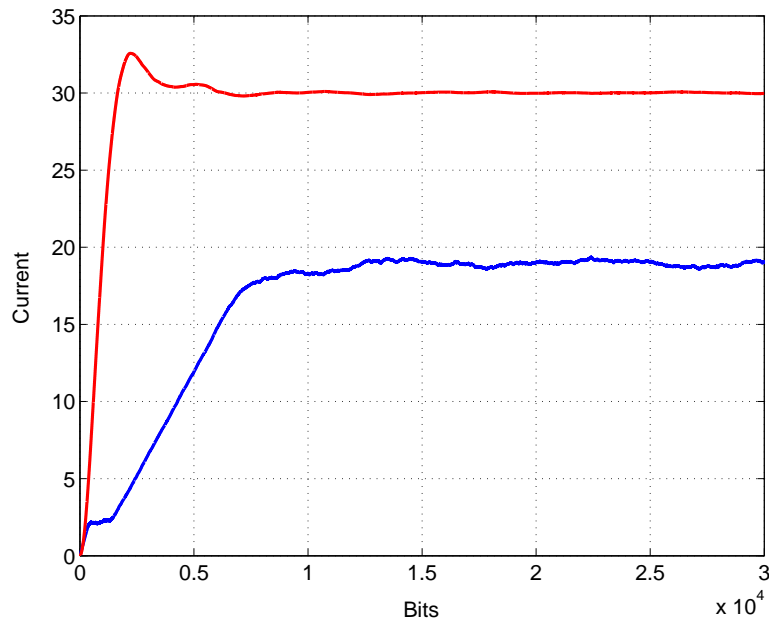


Figure 4.5: Simulated loop dynamic for DDCC with comparator noise and offset, $\sigma_n = 1$, offset=3

It is also important to investigate the effects of the noise and offset of the comparator in the DDCC loop. Figure 4.5 illustrates the performance of the loop in presence of noise and offset in the comparator, where they create small ripples in the correction currents. Heavy filtering in the loop can reduce this problem. However, if the offset is relatively large compared to the random noise, as is the case in Figure 4.5, a dead-zone is created and the correction current will not reach to the final value. The DC error in the size of the currents leads to larger overall voltage swing at the input, which depends on the input pattern and N_{DC} . Therefore, it is important to reduce the offset of the comparator below the noise level.

Among the challenges of this design is the introduction of the correct offset into the data comparators shown as $kD[n-m]$ in Figure 4.1. In this design we assume that we can calibrate the offset using a training signal at the beginning of the actual data transmission, for the best BER. It is also possible to set the induced-offset dynamically by adding an extra comparator and a secondary loop for offset adjustment. However, this solution adds to the complexity and area of the front-end.

In our double sampling/integrating front-end, for small values of ΔV_b , precise offset cancelation is essential to achieve a low BER. As we mentioned in Chapter 3, the residual offset of the sense-amps has an unwanted dependency on the common-mode voltage. The proposed control loop can reduce the overall voltage swing at the input node when the input optical power is close to the receiver sensitivity. Therefore the problem of offset variation reduces by this technique and sensitivity improves. At large input optical powers, ΔV_b is large and offset cancelation is not as crucial for a correct data resolution.

4.3 TIA Scaling

For comparison reason, it is interesting to investigate the performance scaling of a TIA as technology scales, as well as its dependence on the characteristics of photodetector. Most of the discussions in this chapter are based on the shunt-shunt TIA performance analyzed in detail in Appendix A.

As we mentioned in Chapter 3, to ensure stability the bandwidth of the voltage amplifier needs to be at least $\alpha = 3 - 4$ times higher than the bandwidth of the TIA (BW). The TIA bandwidth is set by the pole at the input node and defines the maximum data rate of the receiver, $R_{max} \simeq 0.5BW$. With constant diode capacitance, in order to achieve higher data rates, the feedback resistor R_f needs to scale down, and the bandwidth of the voltage amplifier should scale up. Reducing R_f creates a linear increase in its thermal current noise power density. Moreover at higher frequencies the input-referred voltage noise of the amplifier translates to a proportionally higher current going onto the parasitic capacitance at the input. The total noise at the input, which is the sum of the amplifier noise and noise of R_f , can be estimated by integrating the current noise power density over frequency. Since the voltage amplifier bandwidth is proportional to BW , we can assume that the noise bandwidth is $F_n \times BW$, where F_n is commonly around 4. With the noise power density and its bandwidth being both proportional to BW , after the integration the total noise power is proportional to BW^2 .

The sizing of the input stage of TIA can be optimized for the lowest input-referred

noise [72]. In the optimized design the C_{gs} of the input common-source transistor turns out to be 0.5-1 times the diode capacitor C_p . The total current noise at the input of TIA is expressed in Eq(4.2) for $C_{gs} \simeq C_p$ from Appendix A.

$$I_n \simeq BW \sqrt{\frac{2kTC_p}{A} \left(\frac{8\gamma\alpha}{3} F_n^3 + 4F_n \right)} \quad (4.2)$$

γ here is the excess noise coefficient of the input transistor and for $0.25\mu\text{m}$ technology is about 2.5. This equation shows that while I_n is reduced by the gain of the amplifier A , it increases linearly with the data rate and bandwidth, BW . The required optical power is proportional to this current noise, $P_{op} = \sqrt{SNR} I_n / R$, where R is diode responsivity and SNR is the target signal-to-noise ratio. Therefore, with a constant C_p the required input optical power, P_{op} increases linearly with the data rate. Note that for such a design the noise and thus the minimum optical power is proportional to the square root of C_p .

As we discussed in Chapter 3 and Appendix A, g_m of the input TIA transistor depends on target ω_T for achieving the required gain-bandwidth product, as well as the diode capacitance C_p for an optimum noise performance. This g_m sets the required bias current of the TIA, I_D . The bias current of the TIA is roughly derived in the Appendix A, Eq(A.12), and shows that I_D is proportional to the diode capacitance, $I_D \propto \omega_T^2 C_p L^2$, where L is the transistor channel length. In order to increase the data rate with the same rate that dimensions shrink (α) while keeping A constant, ω_T needs to increase as α^{-1} . Therefore, as technology scales, with a constant C_p , I_D stays constant. Since the electrical power consumption is the product of the current and power supply voltage Vdd , the TIA power consumption scales down with the same rate as Vdd in future technologies.

The above power analysis is optimistic since I_D is estimated without including the short-channel effects in CMOS technologies. Achieving high gains in deep submicron technologies is in fact very challenging and requires higher bias currents, unless bandwidth enhancing techniques are employed.

4.4 Summary

Advanced CMOS technologies are promising for building thousands of optical IOs per chip that run at multi-Giga bit per second data rates. For building dense arrays of receivers on chip, scaling of required electrical power and optical power per IO, are critical. In this chapter we examined different aspects of double sampling integrating front-end and TIAs as CMOS technology scales. While the proposed integrating front-end is more like a digital block and many of its peripheral circuitries scale very well, the noise and sensitivity requirements restrict the scaling of the front-end sampler and comparator. Thus the front-end power scales only with same rate as V_{dd} . Interestingly, the power consumption of TIAs scales at most as fast as our receiver. The power consumption and sensitivity of both receivers depend strongly on the diode capacitance. However, the TIA sensitivity depends on the square root of C_p while the integrating front-end sensitivity linearly depends on C_p . Scaling of the power supply and C_p will slow down due to physical limitations. Therefore, both the electrical power and the optical power of IOs put a hard limit on the number of IOs possible on-chip.

TIAs and the integrating front-end both suffer from the reduced voltage headroom of scaled technologies. The reduced dynamic range at the integrating node limits the maximum input optical power and the DC-balance range of data. In this chapter we proposed a decision directed common-mode control technique to solve this problem. By adding one extra comparator, a feedback loop adjusts and switches an extra current onto the integrating node. The direction of the current depends on the value of resolved bits. Thus this current compensates for the voltage swings of previous bits and removes the dependence of overall voltage swing on the DC-balanced range. Instead, the voltage swing depends on the delay of the loop.

Chapter 5

Clock Generation and Timing Recovery

The integrating front-end described in Chapter 3, directly samples the voltage of the input node to resolve the data. In order to maximize the signal amplitude and sensitivity of the front-end, the sampling must occur precisely at the end of a bit time to make the integration period exactly equal to the bit period. As we mentioned in Chapter 3, the high data-rate in this design is feasible by a time-division demultiplexing scheme. The timing to trigger each branch of the receiver in sequence is governed by multiple clocks with equal phase spacings. The average clock frequency times the number of phases should be equal to the data rate. In this chapter we explore techniques to accurately generate and align the synchronous multiphase clocks for the receiver, with focus on low power techniques. While the clock recovery for this receiver has many similarities with the electrical link CDRs, the integrating nature of the high impedance input node and small voltage swings, differentiate it from the previous solutions. Unlike long-haul optical communication, in this work the clock recovery is performed right at the front, with no initial signal amplification. We start the chapter by exploring the possibility of extending the standard 2x-oversampled clock recovery technique, developed for the electrical links and long-haul optical links, to this front-end.

In most designs a large percentage of receiver power is consumed in the clock

recovery circuits and clock buffers. This is particularly true for our low-power front-end design. In order to reduce the overall power consumption of the receiver a new baud rate clock recovery technique is proposed and analyzed next.

For validation of these ideas we built a transceiver test-chip in $0.25\mu\text{m}$ standard CMOS. This chip consists of 2D arrays of transmitters and receivers with the proposed clock and data recovery techniques. For comparison reasons, half of the receivers have the 2x-oversampled CDR and half the baud rate CDR. The second part of this chapter is focused on the design and implementation of this test-chip. As we will see, our measurement results indicate that the noisy phase information of these techniques require CDR loops with long integration periods for a low-jitter performance. Possible improvements in the design and loop architecture are proposed at the final section of this chapter.

5.1 2X OverSampled Clock Recovery

The most common CDR technique for electrical signaling over $50\ \Omega$ wires is the 2x-oversampled CDR described in Chapter 2. A similar technique can be used for our integrating front-end. Figure 5.1 shows the input voltage waveform upon receiving a one-zero transition. The front-end samples the input waveform at the end of each bit-period. The data $D[n]$ is then resolved by comparing the present sample with the previous sample taken one bit-period before that. At any transition, if the clock is in-phase with data, the two samples taken at the middle of these consecutive non-equal bits, $V_{m_{n-1}}$ and V_{m_n} in Figure 5.1, are expected to be equal. Any phase error would cause these two voltages to be different. Therefore if these consecutive middle samples which are similarly one bit-period apart, are compared at any transition, the difference between the two values provides complete information about the magnitude and direction of the phase error. This difference can be used as an error signal in a PLL or DLL to adjust the frequency and phase of the sampling clocks. In a bangbang-controlled loop only the sign of the error signal with respect to the transition (10 or 01) is used to correct the phase and frequency with constant steps. This is done by generating the corresponding up/dn commands. In order to implement the clock

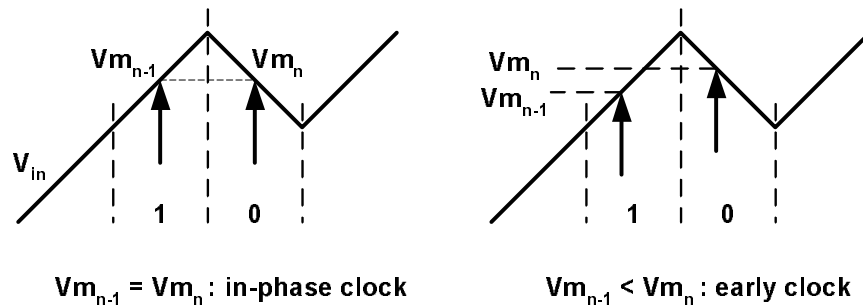


Figure 5.1: 2x-oversampled phase detection for the integrating front-end

recovery loop, one can simply duplicate the samplers/comparator part of the double sampling/integrating front-end. This new set of samplers/comparators needs to be clocked with an extra clock phase, shifted by half a bit-period. The phase correction up/dn commands can be easily generated by using the resolved values of the data and phase resolution blocks. The control loop then adjusts the clock phase by trying to equalize the consecutive middle samples, Vm_n and Vm_{n-1} in Figure 5.1, at any transition.

The 2x-oversampled technique has the advantage of providing phase correction at any data transition. However, it requires extra sets of samplers and clock phases that add to the area, power consumption and design complexity, particularly for multiplexing schemes with many phases.

5.2 Baud Rate Clock Recovery

Removing the extra phases required for oversampled CDR can help to reduce the power consumption in the oscillator and clock buffers and relax the difficulties of phase spacing control. As we discussed in Chapter 2, there are a number of techniques proposed to avoid the extra sample and to perform baud rate clock and data recovery in digital communication systems [43]. The requirement of analog and digital signal processing in such techniques adds to the complexity and can increase the delay of the loop. The double sampling/integrating front-end has unique properties that allow a baud rate CDR with an efficient and simple scheme. In this section, we discuss

the basic principles of this CDR technique, followed by performance analysis and comparison.

5.2.1 Principles of Baud Rate Clock Recovery

The 2x-oversampled CDR can be slightly modified to enable a phase recovery scheme that only needs the same voltage samples used for data recovery. Figure 5.2 shows the input voltage waveform of the integrating front-end upon receiving the data. As an example we choose a "1100" data pattern. It is clear from the figure that for this particular pattern, if the sampling clock is in-phase with the incoming data, the two samples V_n and V_{n-2} will be equal. Any error in the sampling clock phase would lead to nonequal V_n and V_{n-2} . The phase error direction, early or late clock, determines the sign of the error difference between the two samples for the "1100" pattern. Therefore, if each data sample is compared with its two-bit older sample V_{n-2} , the result information can be used for phase recovery. The operation is similar to normal data resolution where we compare each sample V_n with a one-bit older sample V_{n-1} . The P comparators in Figure 5.3 are added to the front-end for this purpose.

The error information for the CDR loop is now the difference in the two samples and the 4-bit pattern that corresponds to samples V_{n-3} to V_{n+1} . Not all 4-bit patterns provide phase information for the clock recovery. The valid patterns for phase corrections are those that give equal V_n and V_{n-2} samples when the clock is synchronized with the incoming data. "0011" and "1100" are patterns that have complete early/late phase information. Most other patterns have conditional phase information only in one direction. For instance "1101" only gives robust results when the input leads the clock as shown in Figure 5.2. Table 5.1 lists valid patterns with the corresponding condition for a meaningful result. Out of 16 possible 4-bit patterns, two give full phase information and four give half phase information. Thus the effective probability of getting phase information from a random input data is 0.25. This is two times smaller compared to 2x-oversampled CDR where phase information is provided at any transition ($P_r = 0.5$).

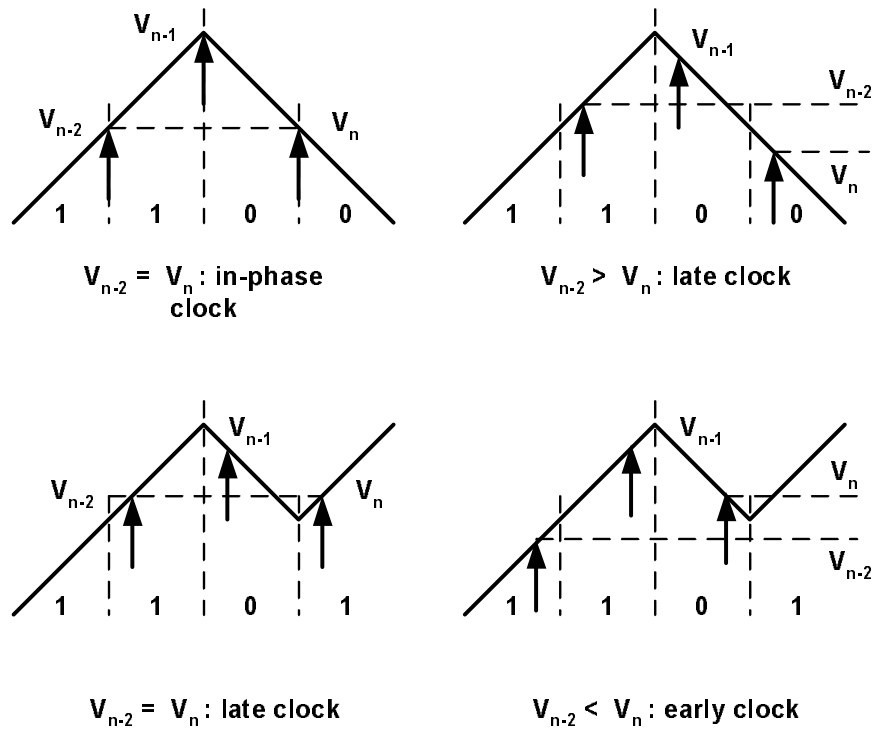


Figure 5.2: Integrating input waveform and baud rate phase detection

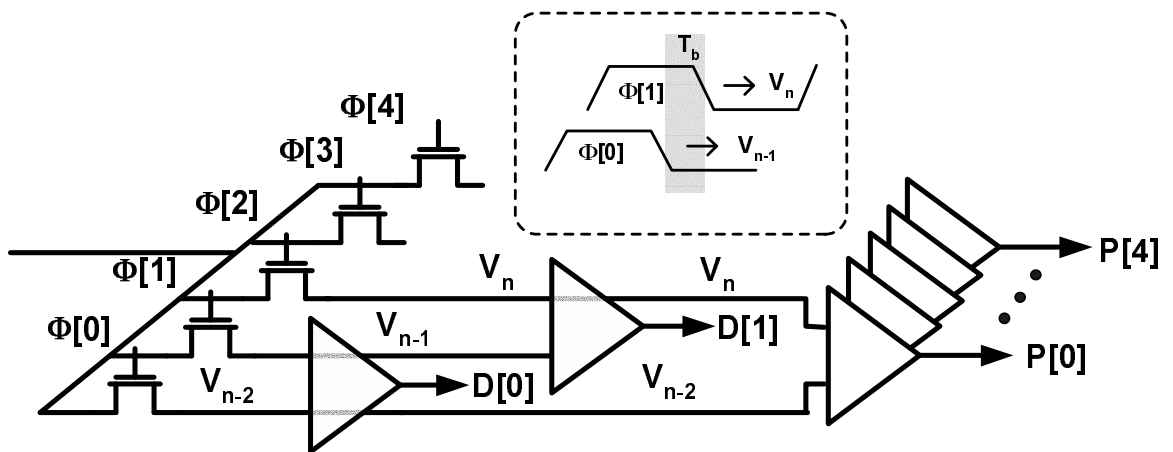


Figure 5.3: Samplers and comparators for baud rate CDR

pattern	<i>late</i>	<i>early</i>
0010	No	Yes
0011	Yes	Yes
0100	Yes	No
1011	Yes	No
1100	Yes	Yes
1101	No	Yes

Table 5.1: 4-bit patterns with phase information for baud rate CDR

In this CDR technique the storage of each data sample for at least two bit-periods is the main requirement. In order to perform the sampling/comparison with a sample kept for two bit periods, at least four different clock edges are required to define the sample, hold and comparison timings. In a de-multiplexing scheme described in Chapter 3, multiphase clocks used to increase the data rate can satisfy this requirement as well. Four or more clock phases allow keeping each data sample for more than two bit-periods. In the next section the performance of described baud rate CDR is discussed in more details.

5.2.2 Phase Detector Performance Analysis

The clock recovery techniques discussed in the previous sections resolve the phase information with schemes similar to the data resolution. It is important to evaluate and compare the phase detection behavior and the output clock jitter, with small voltage swings at the input node. A typical block diagram of a bangbang-controlled clock recovery loop is shown in Figure 5.4. From the incoming data two sets of samplers and comparators resolve the data, D signals, and raw phase information, P signals. A pattern and phase detector module then generates the up/dn phase correction commands for the loop. The clocking of the P comparators and the phase detector design and logic depend on the clock recovery algorithm. The up/dn commands are filtered and used to adjust the frequency and phase of a controlled oscillator, which generates the receiver sampling clock. In this loop the frequency adjustments step or the integral gain is δ_f , and the phase adjustment step or the proportional gain is δ_ϕ .

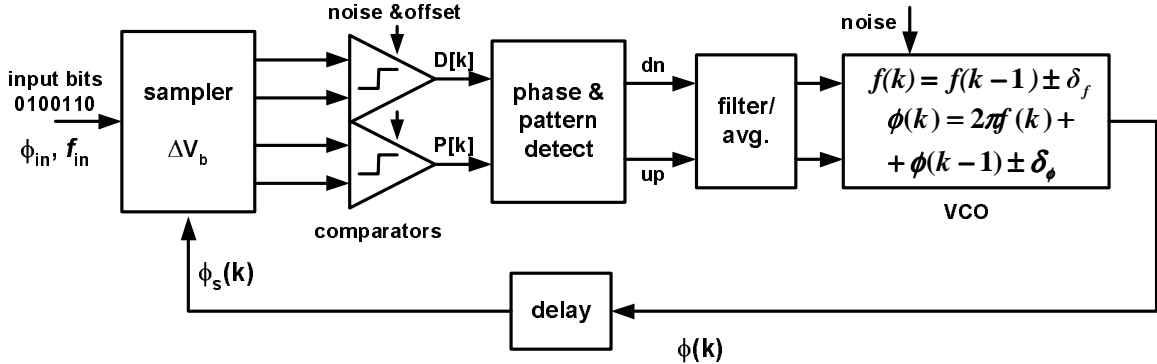


Figure 5.4: Bangbang CDR loop architecture used for the performance analysis

In practice, the comparator, phase detector and filter introduce extra delays in the loop which is included as a delay block.

We use this architecture to simulate and evaluate the performance of the two phase detection techniques for our front-end. The input stream of data creates a voltage waveform similar to the integrating front-end with a voltage swing per bit of ΔV_b and has a certain phase ϕ_{in} and frequency f_{in} . For a general analysis, the input frequency is scaled to unity, and each bit-period is assumed to be 360 degrees. The timing of the samplers which generate the inputs for the comparators, is set by the recovered clock ϕ_s .

The phase detector performance is examined while the loop is open and intentional phase misalignments are introduced between ϕ_s and ϕ_{in} . The graph in Figure 5.5 compares the performance of the two phase detection techniques, 2x and baud rate sampling, by looking at the probability of up/dn commands versus phase misalignment for a random input data. We included the noise and offset of the samplers and comparators in these simulations. The lower phase detection gain for the baud rate CDR is expected since the overall probability of occurrence of patterns with phase information in baud rate CDR is 0.25, while it is 0.5 for a standard 2x-oversampled system. Unlike 2x-oversampled technique, the baud rate CDR, has wrong up/dn decisions even at high phase offsets or with no noise and offset in the comparators. This is because of the fact that each pattern with partial phase information has a

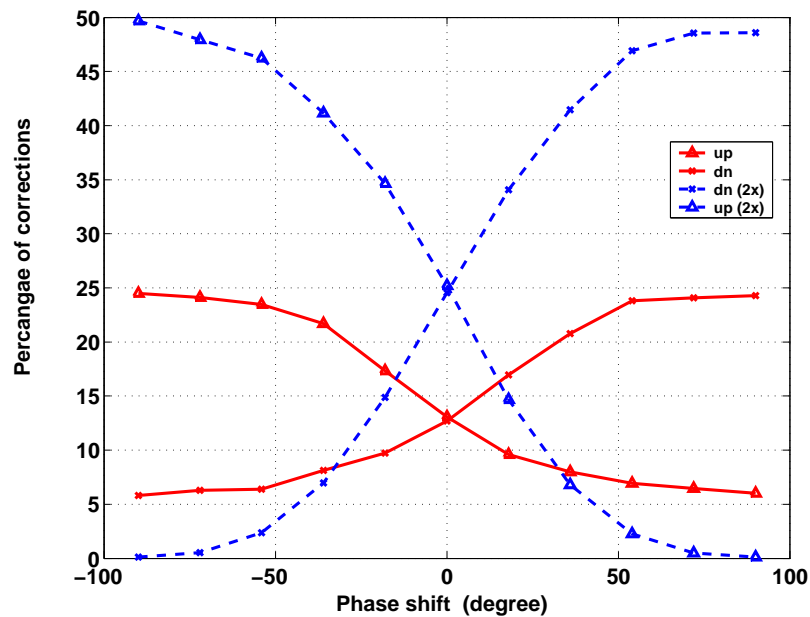


Figure 5.5: Percentage of phase correction commands vs. phase misalignment for the 2x-oversampled and baud rate techniques

25% chance of wrong phase resolution. These patterns happen with probability of $4/16=0.25$. Therefore the overall probability of wrong phase decision is $1/16$ (6.25%) even with no noise and offset in the system. This effect can be reduced by filtering the up/dn commands. With a majority vote filter, the phase detection gain of the two techniques get closer, as shown in Figure 5.6. The reductions in the effective gain of the phase detector and the probability of phase corrections are the main trade-offs for using baud rate clock recovery.

5.3 Optical Transceiver Test-Chip

The clock recovery techniques discussed in the previous sections were implemented in a CMOS test-chip. This chip consists of arrays of transceivers for parallel optical interconnection. The 2D array of receivers uses the double sampling/integrating front-end with clock recovery per receiver. In this section we discuss the architecture

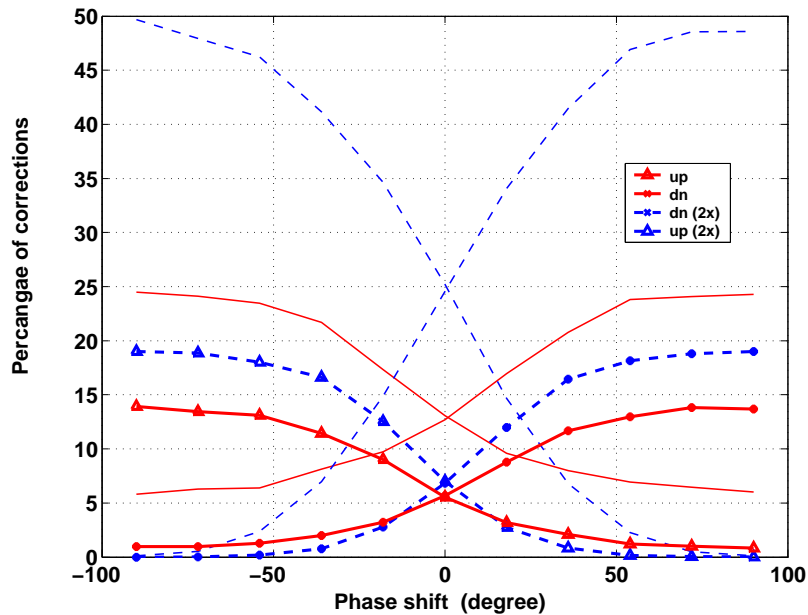


Figure 5.6: Percentage of phase correction commands after a majority filtering over every 5 bits

and design of this chip with focus on clock generation and timing recovery circuits. The measurement results are discussed and evaluated next. As we will see some of the measured results are less than optimum. In order to explain these results, we revisit our simulations, which helps us to propose improvements for enhancing the performance of the CDR.

5.3.1 Clocking High-Level Architecture

Power consumption and jitter performance are the most important considerations for choosing the overall clocking architecture for dense high-speed IOs. For our system we need to generate synchronous multiphase clocks for the multiplexed transmitters and integrating front-end receivers.

In a 2D array of transmitters on chip, we can either distribute the multiphase clocks from a shared PLL or generate them at each transmitter locally. In both cases, precisely matching the delays of the multiple clock paths is important to maintain

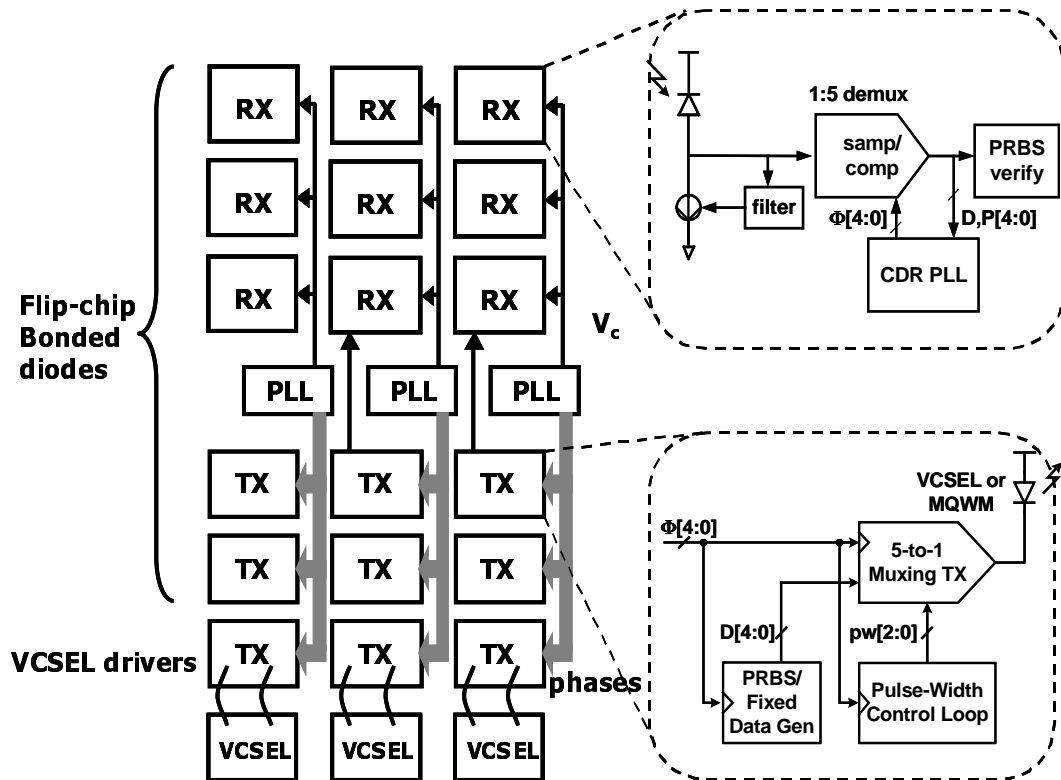


Figure 5.7: Optical transceiver chip block diagram

equal spacing between the clock phases. Even with identical layout, the distribution paths are subject to random mismatches and noise, causing both jitter and phase offset. Longer path delays create larger mismatches [86] [87]. The mismatches cause non-equal bit-times and degrade the SNR. For this reason, distributing the centrally-generated multiphase clocks to all transmitters is usually impractical. However, by sharing clocking components and blocks, we can reduce the overall power consumption and area. Figure 5.7 shows a top-level block diagram of the transceiver array. A compromise was used in this design where each column of the transmitter array shares the same global PLL and multi-phases are distributed vertically to all transmitters in the same column. At the receiver side, each receive channel has its dedicated local CDR PLL that generates the multiphases.

Figure 5.8 illustrates the high-level architecture of the transmitter and receiver

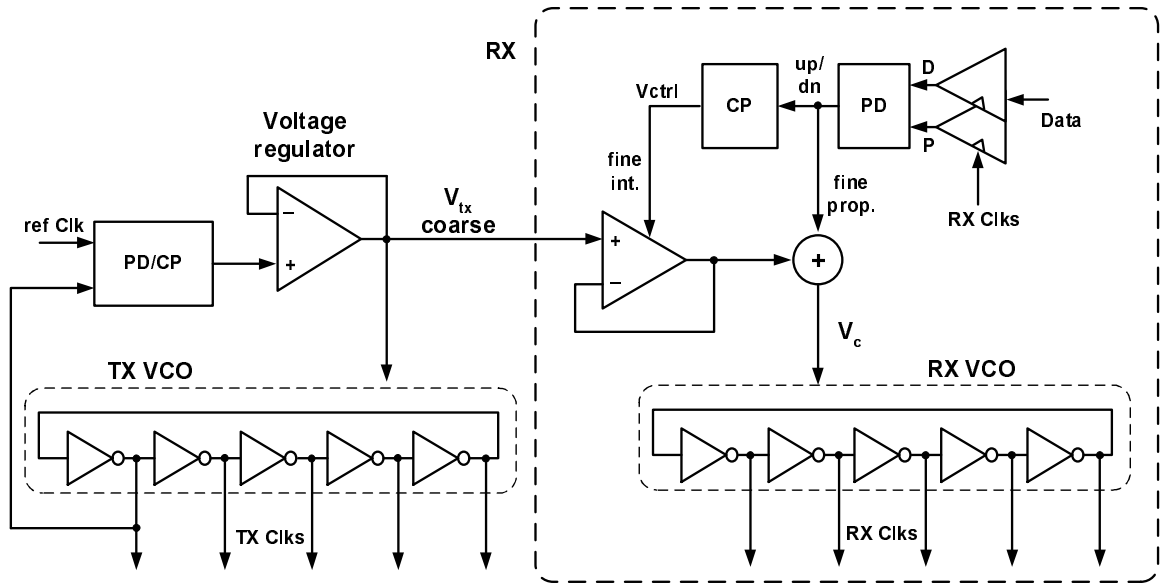


Figure 5.8: Multiphase clock generation for the transmitter and clocked integrating front-end

clocking. The multi-phase clocks are generated by tapping equally spaced outputs of a ring VCO. The control voltage from the transmitter's VCO, V_{tx} is used to set the coarse frequency level of a similar VCO at the receiver, shown in Figure 5.8. The phase correction signals from the CDR then drive the fine control loop of the receiver VCO.

5.3.2 CDR Building Blocks

As illustrated in Figure 5.8, in our PLL design, the output of the phase detector sets the control voltage of the VCO through a charge pump filter. The control voltage is sent to a voltage regulator that controls and suppresses the noise of the VCO supply to improve the jitter performance. The following sections briefly discuss the design of the loop components.

Voltage-Controlled Oscillator

The core of the transmitter and receiver PLL is the VCO with multiphase clocks. Ring oscillators consist of adjustable delay elements and can facilitate generation of multiphase clocks at a variable frequency. CMOS inverters with regulated supply voltage have been used widely as delay elements of ring oscillators [42] and the noise performance of such VCOs has been investigated in [40] [88] [89]. In a standard 2x-oversampling CDR, extra clock phases with half a bit-period shift are required. Generation of main phases and middle phases with single-ended inverters using a single ring is not possible. For such systems, differential delay elements [41] or coupled ring oscillators [90] are required to generate both true and complementary clocks. However, the clocking power consumption increases significantly in these designs. Figure 5.9 illustrates the coupled ring VCO described by Kim in [1] and used in our test-chip. For this design in a $0.25\mu\text{m}$ CMOS, the VCO gain is about $K_{vco} = 600\text{MHz/V}$.

Phase Detector and Charge Pump for Global PLL

The global/transmitter PLL locks to an external reference clock. In such PLL designs, a wide VCO frequency and phase adjustment range is required and frequency-phase detectors (PFD) are used for this purpose. The PFDs output is ideally linear for the entire range of input phase differences from 2π to -2π . The performance of a number of different PFDs are described by Mansuri et al. in [2]. The phase-frequency detector shown in Figure 5.10 with pulsed latches [91] is used for our global PLLs. This PFD has a high operating range and low power consumption, analyzed in [2].

A low-pass loop filter follows the PFD to convert the output up/dn signals to the control voltage of the VCO. In this PLL design the filter is implemented by a charge-pump (CP) which sources an average current proportional to the phase misalignment onto a series capacitor and resistor. The up/dn signals turn on small current sources to add/subtract charges to/from the CP capacitor at each cycle. A simplified schematic of the CP is illustrated in Figure 5.11. As we mentioned in Chapter 2, a compensation zero in this second order loop is needed to insure stability. The zero is implemented

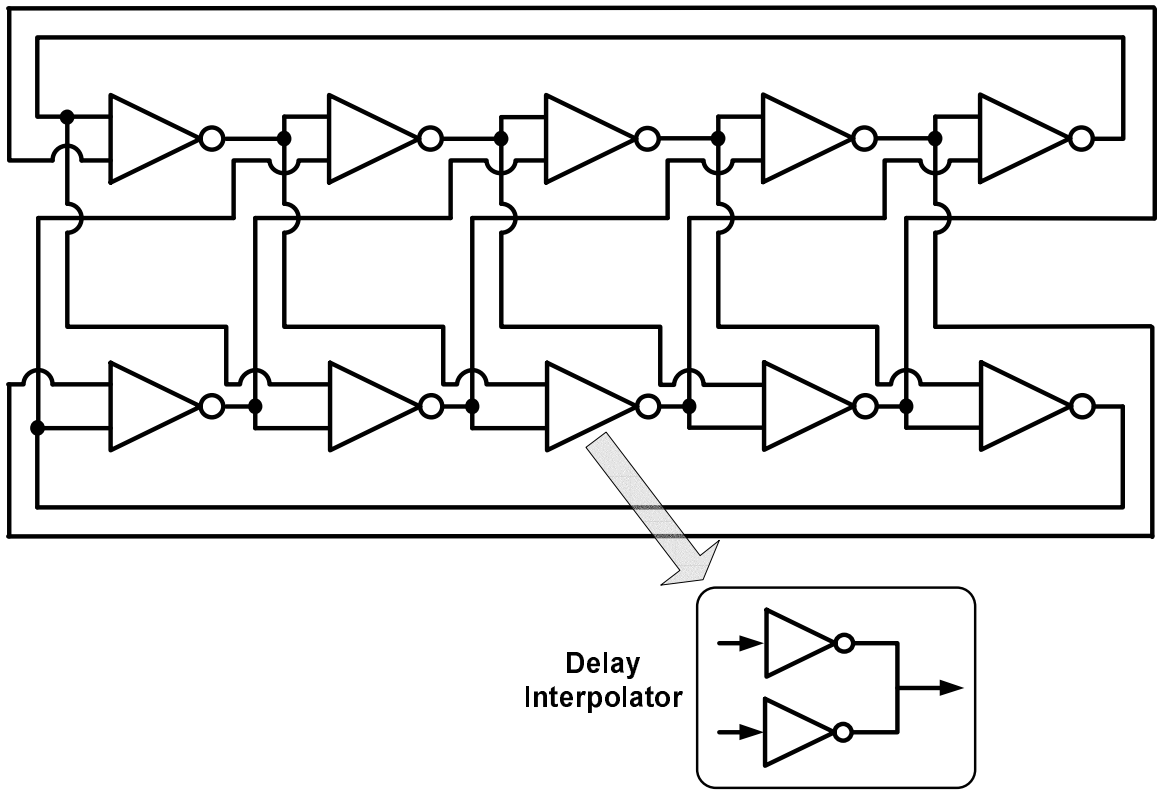


Figure 5.9: Coupled ring-oscillator [1]

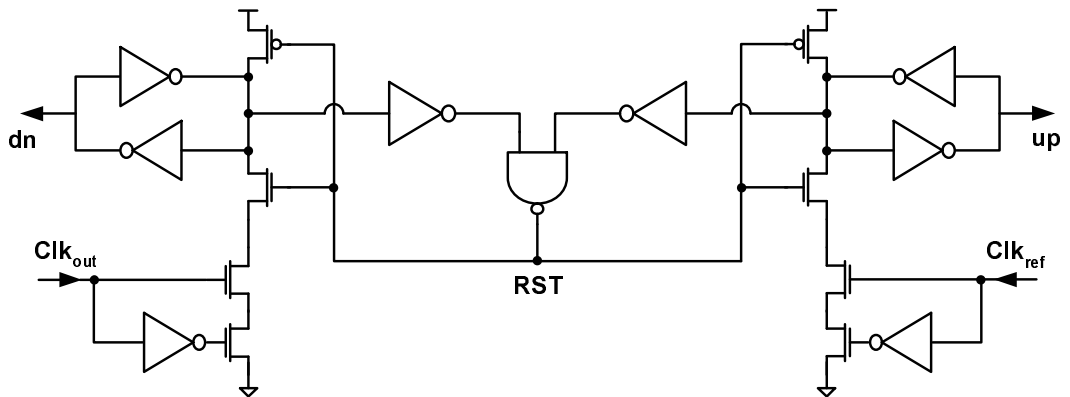


Figure 5.10: Phase-frequency detector for the global PLL [2]

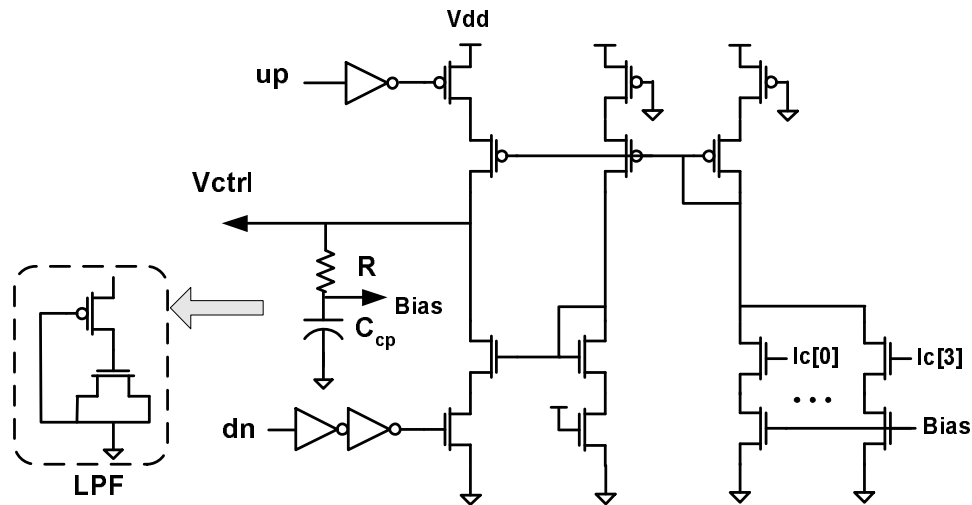


Figure 5.11: Charge pump and loop filter

by a series resistance which is a PMOS transistor with its gate terminal connected to ground. The output voltage of the CP, V_{ctrl} , is buffered by a unity-gain voltage regulator described in the next section.

Voltage Regulator

In order to reduce the clock jitter the VCO control voltage noise should be suppressed to very small values. Since the output of the loop filter by itself is not immune to noise and can not drive the switching current of the VCO inverters, a linear voltage regulator that tracks the average of the loop filter output is necessary to provide a stable and low noise supply for the VCO.

The linear regulator is a differential amplifier driving a PMOS current source with a unity-gain feedback; see Figure 5.12. In order to filter out the high frequency noise generated by the switching of the VCO inverter, a large load capacitor (about 2pF) is used at the output of the regulator. This high capacitive load requires compensation to insure the stability of the regulator. Supply noise rejection and area are the main considerations for the compensation techniques.

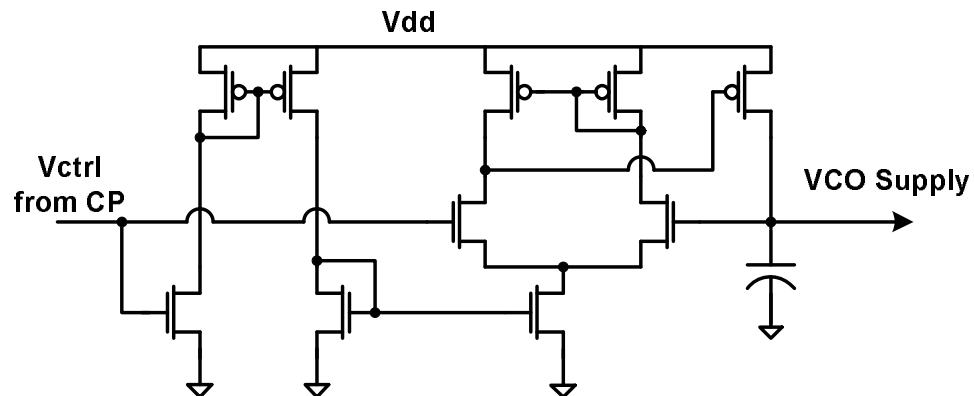


Figure 5.12: Linear voltage regulator for low noise VCO voltage control

Phase Detector for Baud Rate CDR

As we discussed in Section 5.2.1, the phase detection for the baud rate CDR is specific and is based on the output of phase resolution comparator P , as well as the four data values adjacent to it. Having a fast and simple phase detector is essential in the CDR loop design.

The phase/pattern detector generates the up/dn commands for correction of early/late output clock. As shown in Figure 5.3 an extra set of sense-amplifiers that compare the V_n and V_{n-2} for baud rate CDR, and $V_{m_{n-1}}$ and V_{m_n} for 2x-oversampled CDR, generate the P signals and the normal data comparators resolve D signals. The phase detector logic for baud rate CDR is shown in Figure 5.13. This block looks for the patterns in Table 5.1 and decide to activate one of the up or dn outputs. The delay of this block directly adds to the overall loop delay and should be minimized. The CP/loop filter adjusts the control voltage of the VCO based on these correction signals.

BangBang Integral and Proportional Control

The receiver PLL uses a ring oscillator similar to the one used for the global transmitter PLL, Figure 5.9. The control voltage of the transmitter PLL, V_{tx} is used for coarse frequency adjustment of local receiver PLL. This voltage is filtered and sent to an

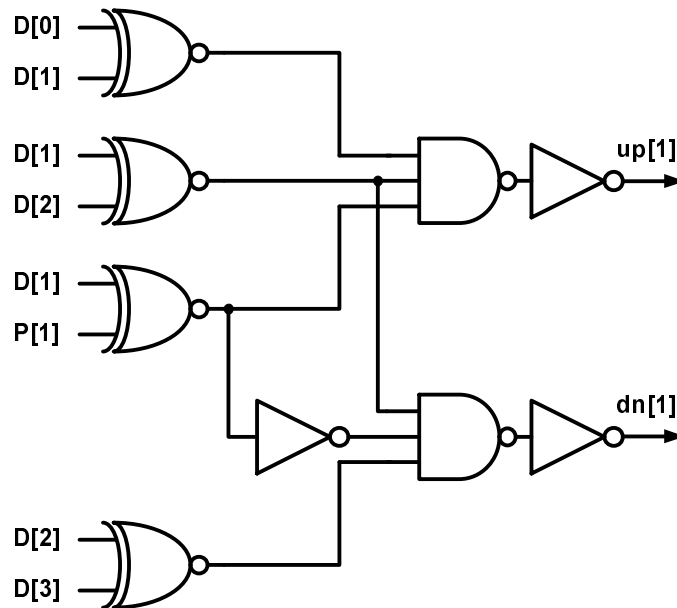


Figure 5.13: Phase and pattern detector for baud rate CDR

additive voltage regulator shown in Figure 5.14. Implementation of the integral and proportional gains is similar to the techniques described in [1] [32]. The output of the charge-pump, V_{ctrl} is attenuated and added to V_{tx} to fine tune the frequency of the RX VCO, acting as an integral loop gain. V_{ctrl} changes the resistance of M_3 transistors in the resistive voltage divider in the feedback path and effectively changes the output voltage. The initial input voltage to the amplifier is reduced by using a voltage divider ($M_{5,6}$) to enable output voltage adjustment in both directions, smaller and higher than V_{tx} . A charge-pump shown in Figure 5.15 generates the V_{ctrl} . For this bangbang-controlled loop the up/dn commands turn on small constant currents onto the CP capacitor at each cycle. The effective voltage change seen at the output of the regulator after each correction can be adjusted and is in the order of $\pm 0.3\text{mV}$. The integral path compensates the frequency mismatches between the receiver VCO and incoming data from another chip. To ensure a stable and non-ringing loop behavior, a proportional gain acting as a zero in the transfer function is required.

The proportional path creates momentary changes to VCO control voltage V_c , and only adjusts the phase of the clock. This behavior is implemented by adding two

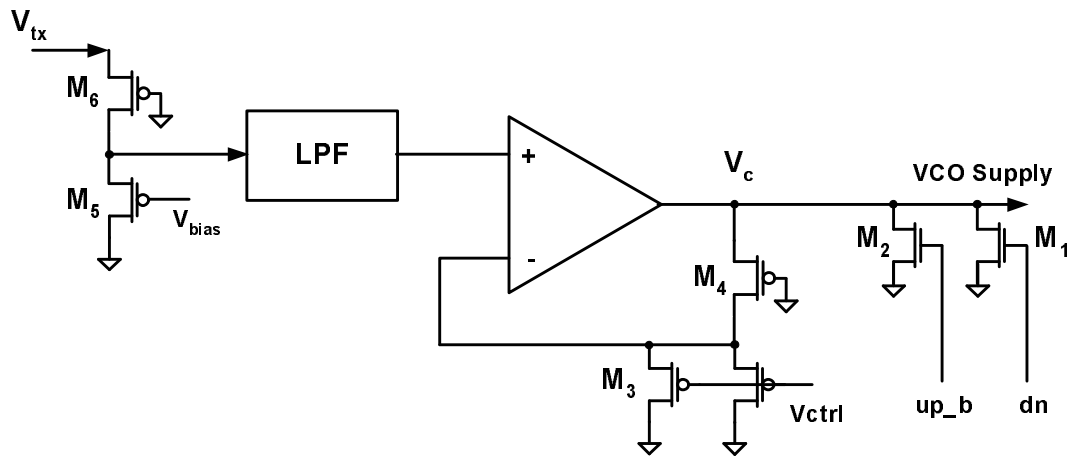


Figure 5.14: Receiver VCO fine control, integral and proportional gains

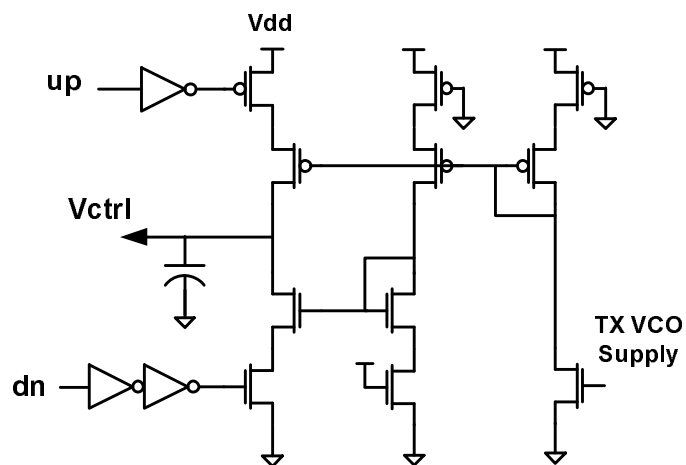


Figure 5.15: Receiver PLL charge-pump

small transistor M_1 and M_2 to the output of voltage regulator. Upon receiving the up/dn commands the transistors will be both ON or both OFF, while in the case of no correction only one of them is ON. Switching of these transistors creates small voltage ripples of around $\pm 30\text{mV}$ on the VCO control voltage at around 1.0GHz.

In a bangbang-controlled loop, the correction steps through the proportional and integral path determine the dynamics of the loop. The bandwidth and locations of dominant poles and zeros automatically track the frequency of operation since both the charge pump bias voltage and the gate voltage of M_1 and M_2 are driven from the V_c itself. Therefore the correction steps are proportional to the VCO frequency and are scaled accordingly. The dynamics of a bangbang-controlled loop are analyzed by Walker in [92] and Kim in [1].

5.4 Experimental Results

The test-chip, with the optical transceiver array and per-channel clock and data recovery is implemented in a $0.25\mu\text{m}$ CMOS technology. A die photo of the transceiver test-chip is shown in Figure 5.16. This chip includes a 3×3 array of receivers and a 3×3 array of transmitters, with 3 global PLLs for each column of the array. Four of the receivers use the 2x-oversampled clock recovery and five of them use the baud rate clock recovery technique.

This chip was designed to be flip-chip-bonded with a 2D array of GaAlAs MQW $p-i-n$ diodes. The receiver array, as well as the first two rows of the transmitters, can interface with these devices through the flip-chip bonding pads. These devices act as detectors at the receiver side and as modulators at the transmitter side. The last row of transmitters are the VCSEL drivers with devices connected to the CMOS chip using short wire-bonds. Unfortunately the flip-chip bonding of the GaAs optical chip to the CMOS transceiver chip faced a number of problems and the optical testing of the receivers and modulator drivers has not been yet possible. The receivers were thus tested by on-chip electrical signaling as well as direct driving of the input node. Table 5.2 summarizes the performance of the receiver and the CDR.

The jitter and sensitivity performance of the two CDR techniques are about the

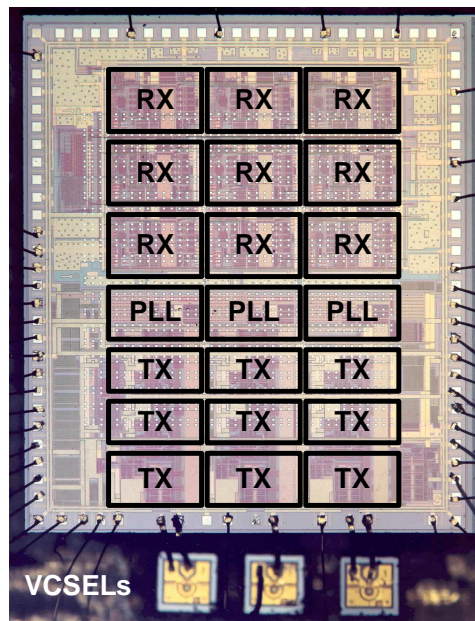


Figure 5.16: Transceiver test-chip for parallel optical interconnection

Technology	0.25 μ m CMOS
Power Supply / Threshold	2.5 V / 0.55 V (NMOS)
Data Rate	5.0 Gb/s
RX Clock Jitter	4.7 psec RMS @ 1.0 GHz
Voltage Swing per Bit	\pm 40mV (BER < 10 ⁻¹⁰)
Power Consumption (2x-oversampled CDR)	87 mW
Power Consumption (baud rate CDR)	75 mW
Area Consumption	0.15 mm ²

Table 5.2: Chip performance summary

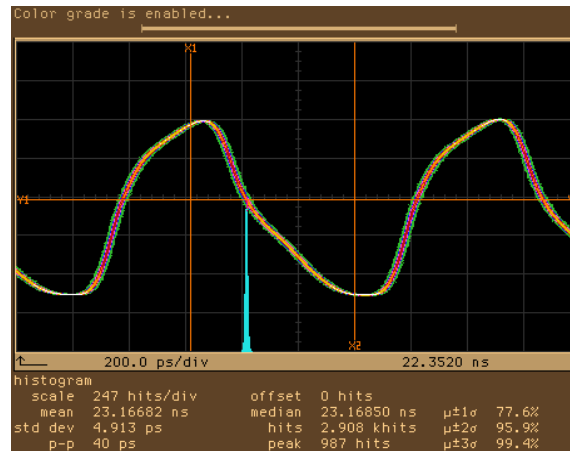


Figure 5.17: Receiver recovered clock signal

same, but the baud rate CDR has about 20% less power consumption due to the removal of middle sampling phases. About 75% of the total power is consumed in the PLL and clock buffers and only 7mW is burnt in each set of comparators for data and phase resolution.

In order to lock the on-chip clock to the incoming data and achieve a low BER, a voltage swing per bit of 40mV was necessary. Figure 5.17 illustrates the waveform of the recovered 1.0GHz clock for a 5.0Gb/s data rate at the receiver. Unfortunately the measured clock jitter and required voltage swing per bit are higher than what we expected. Our secondary simulations and analysis showed that our loop design was not optimum. In the next section we analyze and explain these results and propose techniques for enhancing the performance.

5.5 Design Improvements

The loop architecture of Figure 5.4 can help us to investigate and simulate the closed loop performance of the two CDR techniques for the integrating front-end. These simulations help us to explain the higher jitter and required voltage swing.

The dynamics of the bangbang-controlled loop greatly depends on the integral and proportional gain of the PLL, δ_ϕ and δ_f , as well as input SNR, VCO noise and loop

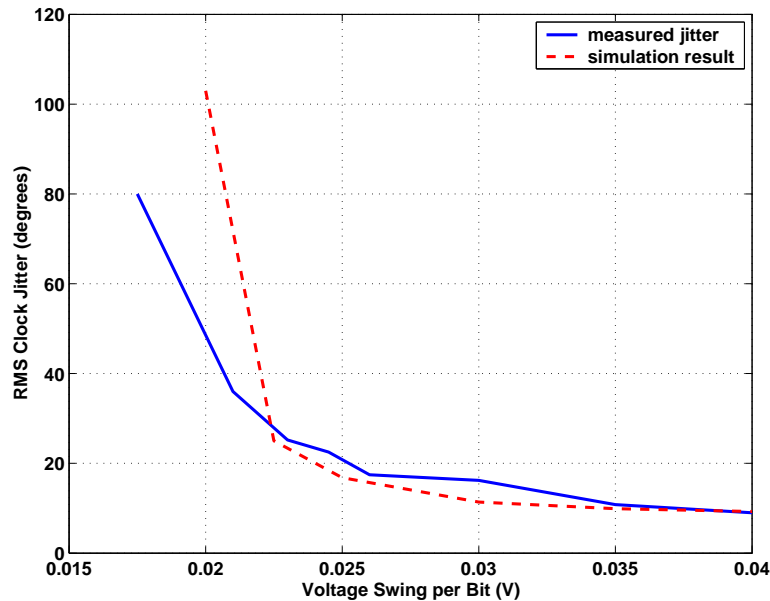


Figure 5.18: Simulated jitter and measured jitter vs. input voltage

delay. In most CDR loops, since the input SNR of practical systems are relatively low, a very low-bandwidth PLL is required. The phase information from our front-end is particularly noisy due to the offset, noise and small voltage swings at the input node. Our simulations show that the voltage swing per bit required for getting the loop to lock, strongly depends on the offset of the comparators, and we need at least 3-4 times more signal swing compared to the offset for a low BER.

The measured RMS jitter in degrees versus input swing per bit, ΔV_b , is illustrated in Figure 5.18. The input frequency is scaled to one and each bit period is assumed to be 360 degrees. This figure also shows our simulation results when offset and noise are 10mV and $\sigma_n = 2\text{mV}$ respectively, with $\delta_\phi/2\pi = 5 \times 10^{-3}$ and $\delta_f = 5 \times 10^{-5}$. The gain values are very close to the SPICE simulation results for our CDR circuits. In these graphs, the jitter has a sharp drop as ΔV_b is increased from low values and then reaches a noise floor. Our simulation indicates that large offsets and high loop gains are the main reasons behind increased ΔV_b and noise floor. In our test measurements, many of the comparators had a large initial offset and needed high offset compensation factors. As we discussed in Chapter 3, larger initial offsets correspond to larger

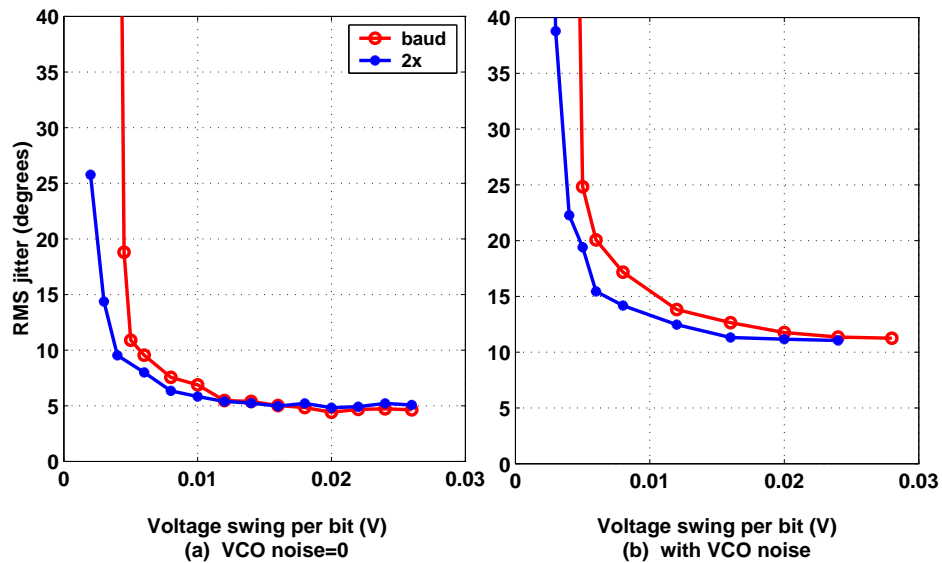


Figure 5.19: Jitter vs. input voltage swing with low loop gain and offset

residual offset variations with the common-mode voltage, after the compensation.

Figure 5.19(a) illustrates the simulated jitter performance for 2x-oversampled and baud rate CDRs, when the loop gains and offset are reduced. In these simulations, we set the proportional and integral gains to $\delta_\phi/2\pi = 3 \times 10^{-3}$ and $\delta_f = 3 \times 10^{-5}$ and offset to 2mV. The jitter and required voltage swing per bit are significantly reduced, which corresponds to better sensitivity and lower BER. However, in this simulation the VCO phase noise is ignored. Although a small loop gain helps to heavily filter the noisy phase measurements, the VCO noise is filtered less effectively and directly appears at the output as shown in Figure 5.19(b). In this simulation the VCO phase noise has a RMS of $\sigma(\Delta f/f) = 0.5\%$. This results confirm that for a low-bandwidth CDR loop, it is essential to keep the VCO noise as low as possible. The phase noise of the inverter-based ring oscillator though, is relatively high [40] [88] for such application.

A dual loop architecture can break the direct trade-offs between the loop gain and VCO noise [93]. A typical block diagram of this PLL is shown in Figure 5.20. The outer loop locks to a clean reference clock with an optimized bandwidth to filter out

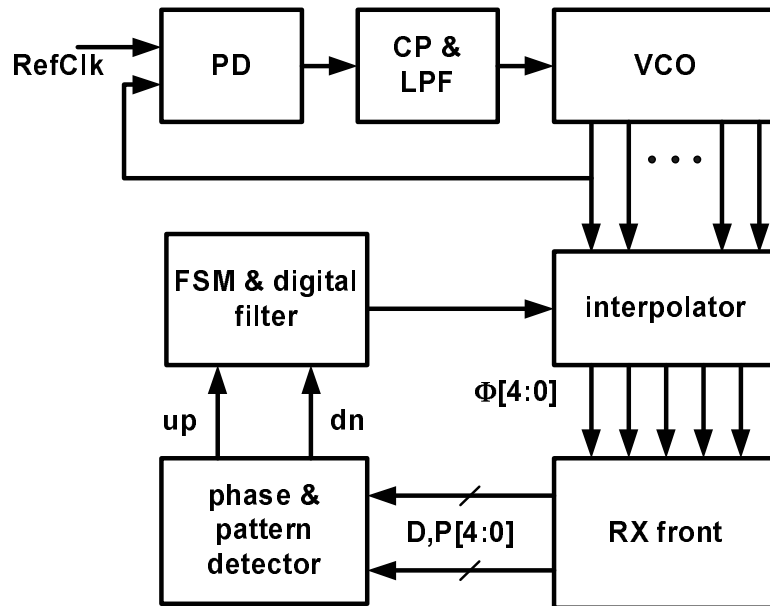


Figure 5.20: Dual Loop Clock and Data Recovery Loop

both VCO phase noise and reference jitter. The digital inner loop mixes the multiple clock phases from the outer loop using a phase interpolator. The interpolation is controlled by a finite-state machine (FSM), driven by the CDR loop with very low bandwidth and heavy filtering. The interpolators in this technique need to generate very accurate phases with fine steps. Note that adding the phase interpolators and the digital loop increases the power consumption and area of the design.

A low bandwidth CDR can help to reduce the clock jitter, but it also reduces the frequency tracking range of the loop. Recently a number of researchers have proposed building a second order phase tracking loop [94] to improve the frequency tracking range.

As we mentioned in Chapter 4, using the decision-directed common-mode control technique described in Chapter 4 can help to reduce the maximum offset by limiting the common-mode variations. However, in future technologies the common-mode change, even with a control loop, will be a large percentage of the voltage headroom. As we mentioned in Chapter 4, the kT/C noise does not allow the scaling of the input transistors and ΔV_b . This is in fact helpful in regard to offset since larger

transistors correspond to smaller mismatches [83] and with scaling, residual offset will be a smaller percentage of the ΔV_b .

5.6 Summary

In this chapter we discussed two different clock recovery techniques suitable for the double sampling/integrating front-end. These techniques use the same decision scheme proposed for the data resolution and do not require any signal amplification for phase recovery. We showed that the standard 2x-oversampled CDR can be extended to this front-end. Since the power consumption is the major limiting factor for the number of IOs per chip and a large percentage of power is consumed in the clock generation and buffering, developing low-power CDRs is crucial.

The integrating nature of the front-end, the double sampling scheme and having multiple clock phases allow us to perform a simple and efficient baud rate CDR. This technique removes the need for the extra clock phases and middle samples which facilitates lower power consumption and less complexity. The baud rate CDR technique requires specific four-bit patterns for phase detection. In a random input data the overall probability of receiving phase information embedded in the data stream is 0.25 compared to 0.5 in 2x-oversampled technique.

The implementation of clock recovery loops using these phase detection techniques was described in this chapter. Our measurement results and secondary simulations showed that our loop architecture limits the jitter performance of the CDR. With the small swings at the input node and the large offset and noise of the samplers and comparators, the output of the phase detector is noisy and requires heavy filtering for a low-jitter performance. However, the phase noise of the ring VCO grows in a low-bandwidth PLL. In our design the combination of limited filtering ability of the loop, high offset and phase noise of the ring oscillator caused a sensitivity and jitter performance which are less than optimum. A dual loop architecture that provides low-jitter clock phases for the inner CDR loop with heavy filtering, can reduce these problems. Our analysis shows that for a good sensitivity offset of the comparators should be very small, thus a dynamic offset cancelation technique or larger transistor

sizes might be inevitable.

Chapter 6

Conclusions

For practical optical chip IOs, simple, low-power CMOS interface circuits are needed. In this thesis we proposed a new double sampling/integrating receiver that looks promising for these applications. By trading off a little sensitivity, it removes the need for gain at the bit rate, making it suitable for low-power CMOS implementation. The double sampling scheme facilitates high data rates and improved sensitivities compared to other integrating receivers.

In this receiver, the optically generated current is integrated onto the parasitic capacitor of the input node and voltage samples at the end of two consecutive bit-times are compared for data recovery. The input node is effectively AC-coupled using a negative feedback loop that subtracts a DC current equal to the average optical current. The data rate for this front-end is limited by the aperture time of the NMOS samplers. Thus by using de-multiplexing and parallelism it can support very high data rates while the timing constraints on the following comparators are relaxed. For this receiver we could achieve bit-times less than 2 FO4 inverter delay, or 5.0Gb/s in our 0.25 μ m test-chip.

The sensitivity of this receiver is determined by the total capacitance at the input node and noise and offset of the samplers and comparators. For a certain photodiode capacitance and de-multiplexing factor, we can optimize the sizing of the front-end for the best sensitivity. The optimum sampling capacitance is a compromise between the kT/C noise and total capacitance of the input node. Further improvement of receiver

sensitivity is possible only if the photodiode capacitance is reduced. The offset of the comparator also directly degrades the sensitivity and should be minimized. In our design an offset compensation technique is used to allow smaller transistors and lower power consumption, though the precision and stability of the offset set the residual offset.

The low power consumption of this receiver makes it an excellent choice for dense array of receivers on-chip. The scaling behavior of this receiver for the purpose of bringing 1000's of high data rate beams to one die is discussed in Chapter 4. The data rate per receiver increases with the same rate as feature sizes scale down. Since the sizings of the samplers and comparators are dictated by the noise, the power of the front-end scales slower than the digital and clocking circuits. However, in our $0.25\mu\text{m}$ CMOS technology, the power of the receiver is dominated by the clock buffers and PLL, thus the overall power consumption initially scales rapidly with the scaled technologies. For a 90nm CMOS technology and a 1.0V supply, $\simeq 1\text{mW}$ receiver power per Gb/s is possible, which allows more than 10Tb/s chip-to-chip bandwidth. At this point, still 60% of the power is consumed in the other parts rather than the front-end. However, since below 90nm the scaling of the power supplies slows down, we predict that the power scaling will not continue for technologies below 65nm.

A major design challenge with advanced technologies is the reduced power supply voltage. For the integrating front-end, this corresponds to limitations on the maximum optical power. Moreover, since the required voltage swing per bit does not scale as fast, the maximum DC-balance range of input data will be tightly restricted. A decision-directed common-mode control scheme proposed in Chapter 4 relaxes the dependence of dynamic range at the input node on the scaled power supply voltages. In this scheme after each data resolution an extra current equal to the photocurrent that corresponds to the received bit is subtracted from the input node. This current compensates for the voltage swing of resolved bits. The size of the compensation current should be adjusted according to the input photocurrent. The adjustment of the current is done with a feedback loop, by adding one extra set of sampler/comparator. This comparator needs to have precise offset compensation to avoid a dead-zone in the current loop. The current subtraction loop has a non-zero delay and the overall

voltage swing at the input node depends on this delay, not the DC-balance range of input data.

The proposed front-end needs a synchronous clock signal to perform the double sampling and comparison. In Chapter 5 we explored the possibility of a standard 2x oversampled CDR technique for our receiver. This technique needs extra clock phases and middle samples which adds to the receiver power consumption. The integrating nature of the input allows us to build a baud rate clock recovery by looking at voltage samples that are separated by two bits. Although this phase measurement is noisy and has low gain, it is effective in a low-bandwidth CDR loop. The principles and performance of this new CDR is also discussed in Chapter 5. A transceiver test-chip with 2D arrays of optical receivers and transmitters was built to support these ideas. This chip supported data rates as high as 5.0Gb/s with 75mW of power per receiver. However, the sensitivity and jitter performance of this design were less than optimum, requiring 40mV voltage swing per bit. Our secondary simulations and analysis showed that due to offset and noise, the loop phase resolutions are very noisy. Therefore, a low-bandwidth loop with longer integration periods and more filtering, is needed. However, a small loop-bandwidth requires VCOs with much lower jitter compared to our ring oscillator. Changing the current architecture to a dual-loop CDR with heavy filtering will help to improve the performance of the receiver.

Appendix A

TIA Analysis

A typical model for a shunt-shunt resistive feedback TIA with a voltage amplifier is shown in Figure A.1. Assuming that the voltage amplifier gain is expressed by Eq(A.1), the overall closed loop gain for this TIA can be approximated by Eq(A.2). This transimpedance has two poles. The first pole is related to the input node with total capacitance of $C_{in} + C_p$, where C_p is the photodiode capacitance and C_{in} is the input capacitance of the voltage amplifier. The second pole P_a is due to amplifier itself, which is usually located at the output node.

$$A(s) \simeq \frac{-A}{1 + s/P_a} \quad (\text{A.1})$$

$$Z \simeq \frac{R_f}{(1 + s(C_{in} + C_p)R_f/A)(1 + s/P_a)} \quad (\text{A.2})$$

In most designs, the dominant pole is the first one at the input node, and to ensure stability the second pole P_a , needs to be at much higher frequencies, at least 3-4 times higher. The 3-dB bandwidth of the TIA is then:

$$BW \simeq \frac{A}{R_f(C_{in} + C_p)} \quad (\text{A.3})$$

The bandwidth of the voltage amplifier BW_a which is set by P_a is then 3-4 times higher than the TIA bandwidth BW . For a target data rate of r_0 , in order to avoid

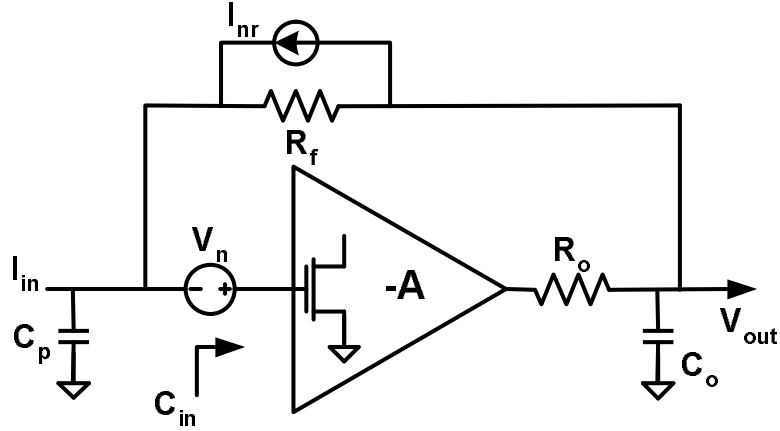


Figure A.1: Shunt-shunt resistive feedback TIA model

dispersion in the output of TIA, the overall bandwidth of the TIA should be at least $BW = 0.5 \times r_0$.

The equivalent noise current power spectral density, \bar{i}_n^2 at the input node is mainly due to the thermal noise of the feedback resistor¹, R_f and the input referred noise of the amplifier v_{nA} :

$$\bar{i}_n^2 \simeq \bar{i}_{nr}^2 + \bar{i}_{nA}^2 = \bar{v}_{nA}^2 \left| \frac{1}{R_f} + (C_{in} + C_p)s \right|^2 + \frac{4kT}{R_f} \quad (\text{A.4})$$

The amplifier noise is dominated by the noise of input common-source input stage and can be estimated with $4kT\gamma/g_m$. γ is the excess noise coefficient of the input transistor and for $0.25\mu\text{m}$ technology is about 2.5. Replacing v_{nA} and also C_{in} with the C_{gs} of the first stage, the noise power density can be expressed by Eq(A.5), assuming that $A \gg 1$.

$$\bar{i}_n^2 \simeq \frac{4\gamma kT\omega^2 (C_p + C_{gs})^2}{g_m} + \frac{4kT}{R_f} \quad (\text{A.5})$$

The total input current noise can be estimated by taking the integral of Eq(A.5) over all frequencies. As we mentioned the amplifier bandwidth is 3-4 times higher

¹ k is the Boltzmann constant and equal to $1.3806503 \times 10^{23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}$ and T is the temperature in Kelvin

than the TIA bandwidth, therefore we can assume that the noise bandwidth is also proportional to BW with a factor F_n which is usually around 4. Now, in order to find the total noise we can simply integrate Eq(A.5) up to the frequency $F_n BW$ [72]. Replacing R_f from Eq(A.3) results:

$$I_n^2 \simeq \frac{4\gamma}{3} \frac{kT(C_p + C_{gs})^2}{g_m} F_n^3 BW^3 + 4 \frac{kT(C_p + C_{gs})}{A} F_n BW^2 \quad (\text{A.6})$$

A popular figure of merit for high frequency performance of transistors is ω_T , a frequency at which the current gain of a common-source amplifier falls to unity. The most common expression assumes that the drain node is shorted and results in $\omega_T \simeq g_m/C_{gs}$. On the other hand the gain-bandwidth product of a common-source amplifier ($A \times BW_a$) with a load capacitance of C_L is equal to g_m/C_L . Since C_L is related to the sizing of common-source transistor (C_{gs}), we can claim that the gain-bandwidth product is approximately proportional to ω_T . As we mentioned before BW_a is larger and proportional to the TIA bandwidth BW . Therefore, we can simply say:

$$g_m/C_{gs} \simeq \alpha \times BW \times A \quad (\text{A.7})$$

where α is at least 3-4. In [72], it has been shown that the optimum value of C_{gs} for minimum noise is 0.5-1 times C_p . Replacing g_m from Eq(A.7) into Eq(A.6) and assuming $C_{gs} \simeq C_p$ results:

$$I_n^2 \simeq \frac{8\gamma\alpha}{3} \frac{kT(C_p + C_{gs})}{A} F_n^3 BW^2 + 4 \frac{kT(C_p + C_{gs})}{A} F_n BW^2 \quad (\text{A.8})$$

From Eq(A.8) we can estimate the minimum required input optical power for a certain bit error rate $P_{op} = R\sqrt{SNR}I_n$, where SNR is the signal-noise-ratio.

We can also estimate the voltage swing at the output of the TIA by $V_{out} = R_f I_{in} = \sqrt{SNR} R_f I_n$. Replacing R_f from Eq(A.3):

$$V_{out} \simeq \sqrt{SNR \frac{AkT}{(C_p + C_{gs})} \left(\frac{8\gamma\alpha}{3} F_n^3 + 4F_n \right)} \quad (\text{A.9})$$

In order to estimate the electrical power consumption of this TIA we need to

relate the bias current of the amplifier to the previous parameters. Transistor drain current is expressed by Eq(A.10)², where for a short channel device κ is less than 2.

$$I_D = \mu C_{ox} \frac{W}{L} (V_{gs} - V_{th})^\kappa \quad (\text{A.10})$$

Defining $\beta = \mu C_{ox} \frac{W}{L} = \mu \frac{C_{gs}}{L^2}$, g_m can be estimated by Eq(A.11).

$$g_m = \kappa \mu C_{ox} \frac{W}{L} (V_{gs} - V_{th})^{\kappa-1} = \kappa \beta^{\frac{\kappa}{\kappa-1}} \sqrt{\frac{I_D}{\beta}} \quad (\text{A.11})$$

Replacing $g_m = \omega_T C_{gs} = \omega_T C_p$ in Eq(A.11) and for $\kappa = 2$:

$$I_D \simeq \frac{\omega_T^2 C_{gs} L^2}{2\mu} \simeq \frac{\omega_T^2 C_p L^2}{2\mu} \quad (\text{A.12})$$

²MOS transistor parameters: μ is the carriers' mobility, C_{ox} is the gate oxide capacitance per unit of area, V_{th} is the transistor threshold voltage, W and L are transistor width and length.

Bibliography

- [1] J. Kim. *Design of CMOS Adaptive-Supply Serial Links*. Phd dissertation, Stanford University, Stanford, CA, 2003.
- [2] M. Mansuri et al. Fast Frequency Acquisition Phase-Frequency Detectors for GSamples/s Phase-Locked Loops. *IEEE Journal of Solid-State Circuits*, 37(10):1331–1334, Oct 2002.
- [3] R. Nair et al. A 28.5GB/s CMOS Non-Blocking Router for Terabits/s Connectivity between Multiple Processors and Peripheral I/O Nodes. *EEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, pages 224–225, February 2001.
- [4] P. Landman et al. A 62Gb/s Backplane Interconnect ASIC Based on 3.1Gb/s Serial-Link Technology. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, pages 72–73, February 2002.
- [5] B. Landman and R. L. Russo. On a Pin vs. Block Relationship for Partitioning of Logic Graphs. *IEEE Transactions on Computers*, 12:14691479, December 1971.
- [6] G. E. Moore. Cramming more components onto integrated circuits. *Electronics*, April 1965.
- [7] W. Dally and J. Poulton. *Digital Systems Engineering*. Cambridge University Press, New York, 1998.
- [8] V. Stojanovic and M. Horowitz. Modeling and Analysis of High-Speed Links. *Custom Integrated Circuits Conference*, September 2003.

- [9] J. W. Goodman et al. Optical interconnections for VLSI systems. *Proc. IEEE*, 72:850–866, 1984.
- [10] M. R. Feldman et al. Optical interconnections for VLSI systems. *Appl. Optics*, 27:1742–1751, 1988.
- [11] D. A. B. Miller. Physical Reasons for Optical Interconnection. *Special Issue on Smart Pixels, Int'l J. Optoelectronics*, 3:155–168, 1997.
- [12] T. Hayashi. An innovative bonding technique for optical chips using solder bumps that eliminate chip positioning adjustments. *IEEE Trans. Components, Hybrids, and Manufacturing Technology*, 15:225–230, 1992.
- [13] J. H. Lau. *Flip Chip Technologies*. McGraw-Hill, New York, 1996.
- [14] K. W. Goossen et al. GaAs MQW modulators integrated with silicon CMOS. *IEEE Photonics Technol. Lett.*, 7(4):360–362, 1995.
- [15] A. V. Krishnamoorthy and K. W. Goossen. Optoelectronic-VLSI: Photonics Integrated with VLSI Circuits. *IEEE J. Selected Topics in Quantum Electronics*, pages 899–912, Nov 1998.
- [16] A. V. Krishnamoorthy and D. A. B. Miller. Scaling optoelectronics-VLSI circuits into the 21st century: A technology roadmap. *IEEE J. Selected Topics in Quantum Electronics*, 2:55–76, Apr 1996.
- [17] A. Emami-Neyestanak et al. A 1.6Gb/s, 3mW CMOS Receiver for Optical Communication. *Proc. IEEE VLSI Symp on Circuit*, pages 84–87, Jun 2002.
- [18] A. Emami-Neyestanak et al. CMOS Transceiver with Baud Rate Clock Recovery for Optical Interconnects. *Proc. IEEE VLSI Symp on Circuit*, Jun 2004.
- [19] E. Reese et al. A Phase-Tolerant 3.8 GB/s Data-Communication Router for Multi- Processor Super Computer Backplane. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, page 296297, February 1994.

- [20] R. Gu et al. 0.5-3.5Gb/s low-power low-jitter serial data CMOS transceiver. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, pages 260–261, Feb 2000.
- [21] M. Fukaishi et al. A 20Gb/s CMOS multi-channel transmitter and receiver chip set for ultra-high resolution digital display. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, pages 260–261, Feb 2000.
- [22] D. K. Jeong et al. Design of PLL-Based Clock Generation Circuits. *IEEE Journal of Solid-State Circuits*, 22(4):255–261, April 1987.
- [23] M. Horowitz et al. PLL Design for a 500 MB/s Interface. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, pages 160–161, February 1993.
- [24] T. H. Lee et al. A 2.5V CMOS Delay-locked Loop for 18 Mbit, 500 Megabyte/s DRAM. *IEEE Journal of Solid-State Circuits*, 29(12):1491–1496, December 1994.
- [25] C.-K. K. Yang et al. A 0.5 μ m CMOS 4Gb/s serial link transceiver with data recovery using oversampling. *IEEE Journal of Solid-State Circuits*, pages 713–722, May 1998.
- [26] R. Farjad-Rad et al. A 0.3- μ m CMOS 8-Gb/s 4-PAM serial link transceiver. *IEEE Journal of Solid State Circuits*, 35(5):757–764, May 2000.
- [27] L.I. Andersson et al. Silicon bipolar chipset for SONET/SDH 10-Gb/s fiber-optic communication links. *IEEE Journal of Solid-State Circuits*, 30(3):210–218, Mar 1995.
- [28] A. Momtaz et al. Fully-integrated SONET OC48 transceiver in standard CMOS. *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pages 76–77, Feb 2001.
- [29] M. Horowitz et al. High-speed electrical signaling: overview and limitations. *IEEE Micro*, Jan 1998.

- [30] S. Sidiropoulos. *High performance inter-chip signalling*. Ph.d. dissertation, Stanford University, Stanford, CA, June 1998.
- [31] E. Yeung and M. Horowitz. A 2.4Gb/s/pin simultaneous bidirectional parallel link with per pin skew compensation. *IEEE Journal of Solid-State Circuits*, 35(11):1619–1628, Nov 2000.
- [32] J. Kim and M. Horowitz. Adaptive Supply Serial Links With Sub-1-V Operation and Per-Pin Clock Recovery. *IEEE J. Solid-State Circuits*, 37:1403–1413, Nov 2002.
- [33] T. Takahashi et al. 110GB/s simultaneous bi-directional transceiver logic synchronized with a system clock. *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pages 176–177, Feb 1999.
- [34] M. Haycock and R. Mooney. A 2.5Gb/s bi-directional signaling technology. *Hot Interconnects V Symposium Record*, pages 149–156, Aug 1997.
- [35] A. DeHon et al. Automatic impedance control. *International Solid-State Circuits Conference Digest of Technical Papers*, pages 164–165, Feb 1993.
- [36] K. Donnelly and et al. A 660 MB/s interface megacell portable circuit in 0.3um-0.7mm CMOS ASIC. *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pages 290–291, Feb 1996.
- [37] C.K. Yang. *Design of high-speed serial links in CMOS*. Ph.d. dissertation, Stanford University, Stanford, CA, December 1998.
- [38] W. Ellersick et al. GAD: A 12-GS/s CMOS 4-bit A/D converter for an equalized multi-level link. *IEEE Symposium on VLSI Ciruits Digest of Technical Papers*, June 1999.
- [39] S. Sidiropoulos and M. Horowitz. A 700 Mbps/pin CMOS Signalling Interface Using Current Integrating Receivers. *IEEE Journal of Solid-State Circuits*, 32(5):681–690, May 1997.

- [40] A. Hajimiri. Noise in phase-locked loops. *Southwest Symposium on Mixed-Signal Design*, 2001.
- [41] J. G. Maneatis. Low-jitter Process-Independent DLL and PLL based on Self-Biased Techniques. *IEEE Journal of Solid-State Circuits*, 31(11):1723–1732, Nov 1996.
- [42] S. Sidiropoulos et al. Adaptive Bandwidth DLLs and PLLs using Regulated Supply CMOS Buffers. *IEEE Symposium on VLSI Circuits, Digest of Technical Papers*, pages 124–127, June 2000.
- [43] K. H. Mueller and M. Muller. Timing Recovery in Digital Synchronous Data Receivers. *IEEE Trans. on Communications*, COM-24(5):516–531, May 1976.
- [44] V. Stojanovic et al. Adaptive Equalization and Data Recovery in a Dual-Mode (PAM2/4) Serial Link Transceiver. *IEEE Symposium on VLSI Circuits Digest of Technical Papers*, June 2004.
- [45] R. Kollipara et al. Design, Modeling and Characterization of High-Speed Backplane Interconnects. *DesignCon*, 2003.
- [46] W.J. Dally and J. Poulton. Transmitter equalization for 4-Gbps signaling. *IEEE Micro*, 17(1):48–56, Jan 1997.
- [47] A. Fiedler et al. A 1.0625 Gbps transceiver with 2x-oversampling and transmit signal pre-emphasis. *IEEE International Solids-State Circuits Conference, Digest of Technical Papers*, pages 238–239, Feb 1997.
- [48] A. Ho et al. Common-mode Backchannel Signaling System for Differential High-speed Links. *IEEE Symposium on VLSI Circuits Digest of Technical Papers*, June 2004.
- [49] J. Proakis and M. Salehi. *Communication Systems Engineering*. Prentice Hall, 1994.

- [50] J. Zerbe et al. Design, Equalization and Clock Recovery for a 2.5-10Gb/s 2-PAM/4-PAM Backplane Transceiver Cell. *IEEE International Solid-State Circuits Conference*, Feb 2003.
- [51] M.Q. Le et al. An Analog DFE for Disk Drives Using a Mixed-Signal Integrator. *IEEE J. Solid-State Circuits*, pages 592–598, May 1999.
- [52] K.K. Parhi. High-Speed architectures for algorithms with quantizer loops. *IEEE International Symposium on Circuits and Systems*, 3:2357–2360, May 1990.
- [53] S. Kasturia and J.H. Winters. Techniques for high-speed implementation of nonlinear cancellation. *IEEE J. Selected Areas in Communications*, 9(5):711–717, June 1991.
- [54] D. L. Mathine. The Integration of III-V Optoelectronics with Silicon Circuitry. *IEEE JSTQE*, 3(3), June 1997.
- [55] D. R. Rolston et al. A hybrid-SEED smart pixel array for a four-stage intelligent optical backplane demonstrator. *IEEE J. Selected Topics in Quantum Electronics*, 2:97–105, 1996.
- [56] M. B. Venditti et al. Design and test of an optoelectronic-VLSI chip with 540-element receiver-transmitter arrays using differential optical signaling. *IEEE J. Selected Topics in Quantum Electronics*, 9:361–379, Mar 2003.
- [57] A. C. Walker et al. Design and construction of an optoelectronic crossbar switch containing a terabit per second free-space optical interconnect. *IEEE J. Selected Topics in Quantum Electronics*, 5:236–249, Mar/Apr 1999.
- [58] D. A. B. Miller. Dense Two-Dimensional Integration of Optoelectronics and Electronics for Interconnections. *Critical Reviews Conference of SPIE's Symp. on Photonics West, Optoelectronics*, January 1998.
- [59] Y. Liu et al. Smart-pixel array technology for free-space optical interconnects. *IEEE, Proceeding of*, 88:764–768, Jun 2000.

- [60] E.L. Walker. A model for the a-profile multimode optical fiber channel: a linear systems approach. *IEEE Journal of Lightwave Technology*, 12(11):1901–1906, NOV 1994.
- [61] K. W. Goossen et al. GaAs MQW Modulators Integrated with Silicon CMOS. *IEEE Photonics Technology Letters*, 7:360–362, 1995.
- [62] D. A. B. Miller et al. Bandedge Electro-absorption in Quantum Well Structures: The Quantum Confined Stark Effect. *Phys. Rev. Lett.*, 53:2173–2177, 1984.
- [63] A. V. Krishnamoorthy et al. Ring oscillators with optical and electrical readout based on hybrid GaAs MQW modulators bonded to 0.8 um silicon VLSI circuits. *Electronics Letters*, 31:1917–1918, 1995.
- [64] A. V. Krishnamoorthy et al. 3-D integration of MQW modulators over active sub-micron CMOS circuits: 375 Mb/s transimpedance receiver/transmitter circuit. *IEEE Photon. Technol. Lett.*, 7:1288–1290, Nov 1995.
- [65] D. A. Van Blerkom et al. Transimpedance Receiver Design Optimization for Smart Pixel Arrays. *IEEE J. Lightwave Technology*, 16:119–126, Jan 1998.
- [66] D. V. Plant et al. 256 channel bidirectional optical interconnect using VCSELs and photodiodes on CMOS. *IEEE J. Lightwave Technology*, 19:1093–1103, Aug 2001.
- [67] T. K. Woodward et al. 1 Gb/s operation and bit-error rate studies of FET-SEED diode-clamped smart-pixel optical receivers. *IEEE Photonics Technology Letters*, 7:763–765, July 1995.
- [68] C. Debaes et al. Receiver-less optical clock injection for clock distribution networks. *IEEE J. Selected Topics in Quantum Electronics*, 9:400–409, Mar 2003.
- [69] T. K. Woodward et al. Clocked-sense-amplifier-based smart-pixel optical receivers. *IEEE Photon. Technol. Lett.*, 8:1067–1069, Aug 1996.

- [70] D. Agarwal et al. Latency Reduction in Optical Interconnects Using Short Optical Pulses. *IEEE J. Selected Topics in Quantum Electronics*, 9:410–418, Apr 2003.
- [71] A. A. Abidi. On the choice of optimum FET size in wide-band transimpedance amplifiers. *IEEE J. Lightwave Technology*, 6:64–66, Jan 1988.
- [72] M. Ingles and M.S.J. Steyaert. A 1 Gb/s, 0.7 μ m CMOS Optical Receiver with Full Rail-to-Rail Output Swing. *IEEE J. Solid-State Circuits*, 34:971–977, July 1999.
- [73] A. Apsel and A. Andreou. A 10 milliwatt 2 Gbps CMOS optical receiver for optoelectronic interconnect. *Proceedings of the 2003 International Symposium on Circuits and Systems*, 1:77–80, May 2003.
- [74] B. Razavi. *Design of Integrated Circuits for Optical Communication Systems*. McGraw-Hill, 2003.
- [75] S. M. Park and H. Yoo. 1.25-Gb/s Regulated Cascode CMOS Transimpedance Amplifier for Gigabit Ethernet Applications. *IEEE J. of Solid State Circuits*, 39(1):112–121, Jan 2004.
- [76] B. Zand et al. A transimpedance amplifier with DC-coupled differential photodiode current sensing for wireless optical communications. *Proc. IEEE Custom Integrated Circuits Conf.*, pages 455–458, May 2001.
- [77] S. M. Park and C. Toumazou. Low noise current-mode CMOS transimpedance amplifier for Giga-bit optical communication. *Proceedings of 1997 IEEE International Symposium on Circuits and Systems*, 1:209–211, June 1997.
- [78] B. Razavi. A 622Mb/s 4.5PA/ \sqrt{Hz} CMOS Transimpedance Amplifier. *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pages 162–163, Feb 2000.
- [79] S.S. Mohan et al. Bandwidth Extension in CMOS with Optimized On-Chip Inductors. *IEEE J. Solid-State Circuits*, 35:346–355, March 2000.

- [80] F. Chien and Y. Chan. Bandwidth Enhancement of Transimpedance Amplifier by a Capacitive-Peaking Design. *IEEE J. Solid-State Circuits*, 34:11667–1170, Aug 1999.
- [81] P. J. Restle et al. A clock distribution network for microprocessors”, journal = ”Symposium on VLSI Circuits, 2000. Digest of Technical Papers. pages 184–187, 2000.
- [82] J. Montanaro et al. A 160-MHz, 32-b, 0.5-W CMOS RISC microprocessor. *IEEE Journal of Solid State Circuits*, 31:1703–1714, Nov 1996.
- [83] M. J. M. Pelgrom et al. Transistor Matching in Analog CMOS Applications. *IEEE International Electron Devices Meeting (IEDM), Technical Digest*, pages 915–918, Dec 1998.
- [84] M. E. Lee et al. Low-Power Area-Efficient High-Speed I/O Circuit Techniques. *IEEE J. of Solid State Circuits*, 35:1591–1599, Nov 2000.
- [85] R. Ho et al. Applications of on-chip samplers for test and measurement of integrated circuits. *IEEE Symposium on VLSI Circuits*, pages 138–139, June 1998.
- [86] L. Wu and Jr. W. C. Black. A Low-Jitter Skew-Calibrated Multiphase Clock Generator for Time-Interleaved Applications. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, pages 396–397, Feb 2001.
- [87] K. Yamaguchi et al. 2.5GHz 4-phase Clock Generator with Scalable and No Feedback Loop Architecture. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, page 398399, Feb 2001.
- [88] M. Mansuri and C-K.K. Yang. Jitter optimization based on phase-locked loop design parameters. *IEEE J. Solid-State Circuits*, 37(11):1375–1382, Nov 2002.

- [89] A. Demir et al. Phase noise in oscillators: a unifying theory and numerical methods for characterization. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 47(11):655–674, May 2000.
- [90] J. G. Maneatis and M. Horowitz. Precise Delay Generation Using Coupled Oscillators. *IEEE Journal of Solid-State Circuits*, 28(12):1273–1282, Dec 1993.
- [91] H. Partovi et al. Flow-through latch and edge-triggered flip-flop hybrid elements. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, pages 138–139, Feb 1996.
- [92] R. C. Walker et al. A Two-Chip 1.5GBd Serial Link Interface. *IEEE Journal of Solid-State Circuits*, 27(12):1805–1811, Dec 1992.
- [93] S. Sidiropoulos and M. Horowitz. A semidigital dual delay-locked loop. *IEEE Journal of Solid-State Circuits*, pages 1083–1092, Nov 1997.
- [94] M. E. Lee et al. A second-order semi-digital clock recovery circuit based on injection locking. *IEEE International Solid-State Circuits Conference (ISSCC), Digest of Technical Papers*, Feb 2003.