A FULLY-INTEGRATED COMPLEMENTARY METAL-OXIDE-SEMICONDUCTOR RECEIVER WITH AVALANCHE PHOTODETECTOR

by

Spoorthi Gopalakrishna Nayak

B.Tech., National Institute of Technology Karnataka, India, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2017

© Spoorthi Gopalakrishna Nayak, 2017

Abstract

Optoelectronic links play a key role in data-center, high-performance computing, sensor and biological applications. Photodetectors (PDs) are used at the front end of every optoelectronic receiver (RX). Traditionally, PDs are fabricated in expensive technologies to enhance their performance. However, wirebonding an external PD to a complementary metal-oxide-semiconductor (CMOS) RX or flip-chip assembly results in several issues – increase in manufacturing and packaging cost, possible decrease in yield, additional packaging parasitics degrading the sensitivity of the RX, crosstalk between the bondwires degrading RX performance, requirement for electro-static discharge (ESD) devices, etc.

Some of the above problems can be surmounted by fabricating the PD in a CMOS process. However, CMOS PDs have very low responsivity. To improve the responsivity, the PD can be operated in the avalanche region where it has a higher current gain. But there are two major concerns – first, Avalanche PDs (APDs) require high bias voltage for its operation, and second, it is very sensitive to variations in operating conditions. Degradation in the APD performance can reduce RX bandwidth and sensitivity. In this thesis, we present an opto-electrical RX which incorporates on-chip APD, bias generation and stabilization for 850-nm optoelectronic interconnect applications.

The proposed receiver consists of CMOS-APD, transimpedance amplifier (TIA), main amplifiers, offset correction loop and 50Ω buffers in the high speed path. APDs are designed and measured to have a -3 dB bandwidth (BW) of 3.5 GHz in 130nm CMOS process, and 6 GHz BW in 65 nm

CMOS process, respectively. The electrical -3dB BW of RX, designed and measured in 130nm CMOS process, is approximately 4.5 GHz. A fully integrated APD-RX system is implemented in 130 nm CMOS process that also comprises of a control loop consisting of an analog-to-digital converter (ADC), synthesized controller, digital-to-analog converter (DAC) and voltage booster. The voltage booster biases the APD with a voltage higher than nominal supply voltages in CMOS, and the control loop stabilizes this bias voltage from temperature variations. On-chip APD based RX with bias generation and stabilization have tremendous potential in optoelectronic links due to inherent advantages of high gain, low cost, reduced ground-bounce and bond-wire parasitics.

Lay Summary

.

Warehouse-scale datacenters require the interconnection of servers spaced apart by a distance varying from a meter to several kilometers. Beyond a few meters, electrical interconnects become unsuitable for large data rate transmission due to high frequency losses, signal crosstalk and reflections.

Optoelectronic interconnects, on the other hand, have negligible signal distortion, frequencydependent losses and crosstalk. Avalanche photodetector based receivers are popular optoelectronic interconnects to improve the sensitivity of the receivers. However, large voltage supply requirements and stability issues of avalanche photodetector have imposed limitations on their widespread use in receiver design, including the need for multi-chip system design. In this thesis, we propose the first fully integrated optoelectronic receiver with on-chip avalanche photodetector, bias generation and stabilization in a standard CMOS 130nm process.

Preface

The research described herein has been conducted under the supervision of Professor Sudip Shekhar in the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, between September 2014 and August 2017.

This thesis is an original work done by the author S. Nayak. She did the literature survey, designed the APDs, control logic, floor-planned and integrated the entire system, carried out the design tapeouts and did all the measurements. A. Sharkia designed and did the layout of the voltage booster. A. H. Ahmed designed and did the layout of the high-speed RX front-end. ADC design was carried out in collaboration with M. W. AlTaha. H. Jayatilleka assisted in setting up MATLAB codes for computing ADC integral non-linearity (INL) and differential non-linearity (DNL). The APD-RX design was fabricated twice in 130 nm CMOS process. The first chip had electro-static discharge (ESD) and latch-up issues. S. Nayak was then solely responsible for the second tapeout. Further changes were made by S. Nayak in the second tapeout, including resolving the latch-up in the voltage booster, improving the BW of the RX front-end, incorporating the package parasitics and improving the design and layout of RX front-end. The noise-cancelling balun was designed by A. H. Ahmed. S. Nayak did a noise comparison of the traditional differential amplifier to that of the implemented balun architecture to characterize its benefits.

All the work presented henceforth has been carried out in the System-on-Chip (SoC) Laboratory at the University of British Columbia, Point Grey campus.

Table of Contents

Abstractii
Lay Summary iv
Prefacev
Table of Contents vii
List of Tables xi
List of Figures xii
List of Abbreviations xvii
Acknowledgementsxx
Dedication xxii
Chapter 1: Introduction1
1.1 Outline
Chapter 2: Background4
2.1 Opto-electronic transceiver

2.2 Optical transmitter 5
2.3 Optical receiver
2.3.1 Receiver noise and sensitivity
2.3.2 Receiver bandwidth 10
2.3.3 Equalization
Chapter 3: Motivation12
3.1 Working of a photodetector
3.2 Link budget
3.2.1 Overview
3.2.2 Link analysis 17
3.3 On-chip PD vs. off-chip PD 20
Chapter 4: Avalanche Photodetector23
4.1 Working of an avalanche photodetector
4.2 APD gain and link budget
4.3 Noise considerations in an APD-RX

4.4	Design challenges for APD-RXs	
4.4.	.1 Need for large bias voltage	
4.4.	.2 Limited gain-bandwidth of APD	
4.4.	.3 Sensitivity to reverse bias voltage and temperature	
4.5	CMOS APD design	
4.6	CMOS APD measurements	
Chapter	r 5: Monolithic CMOS optical receiver	43
5.1	Transimpedance amplifier	
5.2	Control loop	50
5.3	Analog-to-digital converter	56
5.4	Voltage booster	60
Chapter	r 6: System design considerations	64
6.1	Packaging model	65
6.2	Latchup	67
6.3	Printed circuit board	67

6.4	Optical probe station	
Chapter	r 7: Conclusion and future work	76
Bibliogr	aphy	
Append	ices	84
Apper	ndix A Latchup	
Apper	ndix B Optical Set-up	
B .1	Set-up 1	
B.2	Set-up 2	

List of Tables

Table 4.1 Performance summary of CMOS APDs and comparison to prior art	41
Table 5.5.1 RX power consumption	49

List of Figures

Figure 2.1 Typical block diagram of an optical link
Figure 2.2 (a) A chart showing different optical modulation techniques. (b) Direct modulated
laser and (c) Indirect modulated laser
Figure 2.3 Basic schematic of a VCSEL driver
Figure 2.4 Optical receiver
Figure 2.5 PD signal and noise model9
Figure 2.6 (a) Pulse response at the output of the TIA, and (b) DFE and IIR equalization 11
Figure 3.1 (a) Intrinsic semiconductor and (b) Extrinsic semiconductor
Figure 3.2 A Basic block diagram of optical link
Figure 3.3 Losses in optical link
Figure 3.4 Effect of RX sensitivity on laser power
Figure 3.5 (a) External PD or APD with CMOS RX, (b) fully-integrated RX with APD in a
modified CMOS process and external biasing
Figure 4.1 Effect of APD gain on RX sensitivity and Laser power
xi

Figure 4.2 APD signal and noise currents	28
Figure 4.3 SNR vs. APD gain	29
Figure 4.4 Reverse bias voltage vs. bandwidth of APD.	30
Figure 4.5 (a) Heating/Cooling APD. (b) Temperature sensors to monitor APD temperature and	1
alter bias voltage. (c) Dumpy APD to stabilize APD bias voltage.	33
Figure 4.6 APD structure fabricated in 130 nm CMOS	36
Figure 4.7 APD structure fabricated in 65 nm CMOS	37
Figure 4.8 Frequency response measurement set-up for APD	38
Figure 4.9 IV curve of N+/Pwell/Deep Nwell APD fabricated in 130 nm CMOS	38
Figure 4.10 M vs. reverse bias voltage of the APD fabricated in 130 nm CMOS	39
Figure 4.11 Effect of reverse-bias on APD bandwidth in 130 nm CMOS	39
Figure 4.12 Effect of core area on APD bandwidth in 65 nm CMOS	40
Figure 4.13 Die image of APDs in 130 nm CMOS	40
Figure 4.14 Die image of APDs in 65 nm CMOS	42
Figure 5.1 Fully-integrated APD-RX in standard CMOS process.	43
Figure 5.2 Front end of the APD-RX. The high-speed circuits are designed by A. H. Ahmed	44
	xii

Figure 5.3 A differential amplifier to convert single-ended TIA output to differential signals for
the MAs
Figure 5.4 A noise-cancelling active balun to convert single-ended TIA output to differential
signals for the MAs. Designed by A. H. Ahmed
Figure 5.5 Electrical eye diagram of the RX at 8 Gbps
Figure 5.6 Electrical S21 measurements of the RX
Figure 5.7 Die image of TIA fabricated in 130 nm CMOS 50
Figure 5.8 Block diagram of the proposed APD-RX system
Figure 5.9 (a) APD IV curve shows APD breakdown due to very high biases and (b) Derivative of the average TIA output vs. the reverse bias voltage
Figure 5.10 Control logic FSM for APD bias stabilization
Figure 5.11 Control logic cases: (a) locking at the avalanche region (b) dithering at the top (c)
positive temperature drift, and (d) negative temperature drift
Figure 5.12 Transient response of APD-RX system. (a) APD current, (b) LPF output and (c) APD bias voltage
Figure 5.13 Block diagram of the SAR ADC. Comparator and SAR logic is designed by M. Al-
Taha

Figure 5.14 Common-centroid arrangement of 7-bit capacitor DAC	58
Figure 5.15 Die image of ADC and controller	59
Figure 5.16 (a) Voltage booster circuit (b) Conventional Dickson voltage multiplier [32] (c)	
Back-compensated voltage multiplier [33] and (d) Implemented voltage booster core. The	
voltage booster is designed by A. Sharkia	61
Figure 5.18 Die image of voltage booster fabricated in 130 nm CMOS	63
Figure 5.17 Measured output vs. input voltage characteristic for the voltage booster	63
Figure 6.1 (a) Bondwire model and (b) Bondwire simulation model	64
Figure 6.2 Effect of bondwire	66
Figure 6.3 PCB test fixture	68
Figure 6.4 PAD clearance for wirebonding and probing	69
Figure 6.5 PCB clearance for probing	70
Figure 6.6 LDO Bypass	71
Figure 6. / VCSEL alignment	72
Figure 6.8 S21 measurement set-up for the VCSEL	73
Figure 6.9 Frequency Response of the VCSEL	74

Figure 6.10 Measurement set-up for VCSEL
Figure 6.11 VCSEL diagram at 6.5 Gbps measured using Lab Buddy 75
Figure 6.12 Optical probe station75
Figure A.1 Cross section of a PMOS next to an NMOS, leading to potential latchup
Figure A.2 Latchup equivalent circuit
Figure B.1 TOSA VCSEL
Figure B.2 Lumentum VCSEL eyediagram
Figure B.3 Probe positioner adaptor 90
Figure B.4 Optical probe holder
Figure B.5 Wirebonded VIS VCSEL91
Figure B.6 Initial VCSEL alignment

List of Abbreviations

ADC	Analog-to-Digital Converter
APD	Avalanche Photodetector
BW	Bandwidth
CML	Current Mode Logic
CMOS	Complementary Metal-oxide-semiconductor
CQFP	Ceramic Quad Flat Package
DAC	Digital-to-Analog Converter
DFE	Decision Feedback Equalization
DNL	Differential Non-Linearity
DRC	Design Rule Check
EAM	Electroabsorption Modulator
ENOB	Effective Number of Bits
ESD	Electro-Static Discharge

IIR	Infinite Impulse Response
INL	Integral Non-Linearity
ISI	Inter Symbol Interference
LDO	Low-dropout Regulator
MA	Main Amplifier
MMF	Multimode Fiber
OMA	Optical Modulation Amplitude
PD	Photodetector
RX	Receiver
SAR	Successive Approximation Register
SCR	Silicon-Controlled Rectifier
SMD	Surface Mount Device
SMF	Single Mode Fiber
SNR	Signal to Noise Ratio
SoC	System-on-Chip

STI	Shallow Trench Isolation
TX	Transmitter
VCSEL	Vertical-Cavity Surface Emitting Laser
VNA	Vector Network Analyzer

Acknowledgements

During the time I spent at the University of British Columbia, I was able to expand my knowledge and professional skills. I am thankful to many people for their guidance and support throughout this period.

First and foremost, I am very grateful to my supervisor, Prof. Sudip Shekhar. His teaching, hard work and design expertise is very inspiring. I am indeed fortunate to have the opportunity to learn from him.

I would like acknowledge the support of Prof. Nicolas Jaeger and Prof. Lukas Chrostowski for their valuable feedbacks and support for building my test set-up. I would also thank Prof. Shahriar Mirabbasi and Prof. Nicolas Jaeger for reviewing my thesis and serving in my committee.

Mohammed Altaha has always been there during tape-out deadlines and I thank him for dropping me home late at nights. I would like to thank Hasitha Jayatilleka, Hossam Shoman, Julian Schmidt, and Gene Polovy for their help on the optical side of my research. My conversations and discussions with Ajith Sivadhasan Ramani and Amit Kumar Mishra have always been insightful, supporting me during times of stress and motivating me to tackle problems afresh. Ajith has been a true support during electrical and optical measurements. I truly enjoyed working with all the colleagues at System-on-Chip (SoC) Laboratory and I have benefitted a lot from our interactions. I would like to thank Roozbeh Mehrabadi for Computer aided design (CAD) tools assistance, Dr. Roberto Rosales for all the technical help, David Feixo Irazabal for all the help in building optical set-up, Natural Sciences and Engineering Research Council of Canada (NSERC) for research support, CMC Microsystems for access to CAD tools and facilitating the chip fabrications, and FAST Semiconductor Packaging for wirebonding the vertical-cavity surface emitting laser (VCSEL).

Last but not the least, I would like to express my sincere gratitude to my family. I cannot thank them enough for all their support and guidance throughout my life. To my dear parents and little baby.

Chapter 1: Introduction

With the rapid growth of datacenters, there is a significant demand to leverage the high-volume manufacturing capabilities of the silicon industry and reduce the overall hardware cost of the interconnect links connecting different servers and switches in datacenters. Traditionally, electrical interconnect links are widely used at short distances of up to a few meters. However, with the increase in the data rate per link, electrical interconnects have increased loss with distance due to skin effect, dielectric loss in the copper traces, reflections due to impedance mismatch, etc. These inherent losses in the electrical link cause pulse dispersion resulting in inter symbol interference (ISI), leading to eye closure in an unequalised link. Equalization schemes like Decision Feedback Equalization (DFE), Feed-Forward Equalization (FFE), Continuous-time Linear Equalization (CTLE) at the receiver (RX) end, and FFE at the transmitter (TX) end, are often employed to enhance the data rate through the lossy electrical link. However, these techniques consume substantial power to compensate for channel loss larger than 25-30 dB at the Nyquist rate.

On the other hand, loss in an optical channel (fiber) is almost independent of the data rate and is negligible when the link distance is less than 300 m (short link). Most of the optical interconnects in the existing datacenters today span a distance of less than 300 m, and operate at 850 nm with multi-mode fibers (MMF), vertical-cavity surface emitting laser (VCSEL) and photodetectors (PDs) [1], because of its low cost and superior performance. With the optoelectronic transceiver market poised towards a significant annual growth, it is imperative to reduce the bill of materials and lower the cost of the transceivers. As most of these transceivers utilize an off-chip photodetector (PD), a fully-monolithic implementation of the PD will further reduce the

component cost and the associated packaging cost. From the link perspective, there are also requirements to increase the data rate from 10 Gb/s to 25 Gb/s and higher, and reduce the overall power consumption. A single chip RX with on-Chip PD and CMOS circuits can not only reduce the packaging parasitics and improve bandwidth, but also ease the overall RX design.

Beyond the transceiver market, a fully-integrated RX will also be attractive to several other industrial applications. PDs are widely used in high-performance computing and sensors, as well as in biomedical instrumentation. PD sensor arrays are widely used in position sensors, image sensors, medical scanners, etc. Having a single chip PD-RX solution makes the design compact, low power, and easily scalable. Hence, realizing integrated and dense sensor-RX arrays would be highly desirable due to reduced complexity, manufacturing and packaging costs.

The aim of this research is to promote a fully-integrated, single chip CMOS RX incorporating CMOS avalanche PD (APD).

1.1 Outline

This thesis begins with a brief review of optical link. Chapter 2 presents the basic principles of high-speed optical links. It also gives an overview of optical transceivers. Chapter 3 presents the motivation behind the work. It analyzes the link budget and discusses the impact of PD gain on overall power consumption of the link. It also describes the pros and cons of using off-chip and on-chip PD. Chapter 4 describes APDs in detail. It discusses the advantages of having an on-chip APD and explore the benefits it has on the RX design. A brief review of APD characteristics is provided, followed by design and measurement results of CMOS APDs in 130nm and 65nm

CMOS processes. Chapter 5 presents a fully-integrated APD-RX optoelectronic system. Various blocks of the proposed APD-RX system are described and their simulation and measurement results are presented. Chapter 6 describes the measurement setups used in characterizing the designed chips, and also presents various learnings related to the design and prototyping of the system. Conclusions and future work are described in Chapter 7.

Chapter 2: Background

In this chapter, we give an overview of an opto-electronic link. Different architectures for TX are briefly described. Design considerations and key specifications for an RX are also presented.

2.1 Opto-electronic transceiver

A general opto-electronic link is shown in Figure 2.1. It consists of an electro-optical TX, which converts the electrical signal into optical signal, optical channel (i.e., optical fiber) and an opto-electronic RX that converts the optical signals back to electrical signal for processing.



Figure 2.1 Typical block diagram of an optical link

At the TX end, to minimize the I/O (input/output) pins, data is serialized and forwarded to the high speed TX driver. Electrical drivers modulate the laser signal, and the modulated signal is coupled to the channel. At the front of the RX, a PD receives the incoming optical signals and produces an equivalent current signal. The electrical current signal is then converted to voltage and amplified 4

by the TIA. The voltage output of the TIA is further boosted by a set of amplifiers, often called Main Amplifiers (MAs), before being resolved to digital levels by a comparator and eventually deserialized.

2.2 Optical transmitter

In an optical TX, optical devices are modulated by high-speed TX drivers. The TX drivers are generally implemented using current mode logic (CML) to meet the BW requirement of 10 Gb/s and higher data rates. Optical modulators can be broadly divided into direct modulators and indirect modulators, as shown in Figure 2.2 (a). In direct modulation, Figure 2.2 (b), the electrical driver modulates the laser source directly by changing the laser bias. In indirect modulators, the driver modulates the light coming out of continuous-wave laser, Figure 2.2 (c). Indirect modulation



Figure 2.2 (a) A chart showing different optical modulation techniques. (b) Direct modulated laser and (c)

Indirect modulated laser

can be further classified into electro-optical modulation where the electric field is applied to change

the optical path length, and electro-absorption modulation where the applied electric field changes the amount of light absorbed.

Vertical Cavity Surface-Emitting Laser (VCSEL) is a popular direct modulated laser source. The output power of the VCSEL can be modulated by changing its bias current. A basic VCSEL driver is shown in Figure 2.3. Here, a pseudo-differential CML is used with a VCSEL on one of its arms and a dummy load on the other. The current is steered between the two driver arms based on the input data stream. The VCSEL needs to be biased at its threshold voltage (knee voltage of the diode) which is often higher than the nominal CMOS supply voltage, V_{DD}. Therefore, the driver usually has a higher supply voltage, V_{DDH}, for its VCSEL arm.



Figure 2.3 Basic schematic of a VCSEL driver

Micro-ring modulators (MRMs) and Mach-Zender modulators (MZMs) are some of the examples of electro-optical modulators. By changing the electric field across the modulator, refractive index of the modulator is changed, and hence the path length is also changed.

Similar to electro-optical modulators, electroabsorption modulators (EAM) are used to modulate the output of continuous-wave laser. Electrical field applied across EAM changes the bandgap energy, and hence there is a change in the absorption spectrum. This phenomenon is called Fran-Keldysh effect. Generally, EAM are implemented as waveguides with electrodes which are driven by high-speed electrical drivers.

2.3 Optical receiver

Optoelectronic RXs convert optical signal to electrical signals which could be used for further processing or storage. Figure 2.4 shows the general architecture of an optoelectronic RX, which can be broadly described into three parts:

• *Optical Detector*: The front end of the optical RX is a PD which receives an optical signal and generates the corresponding electrical current. The light-to-current relationship of the PD can be defined as its responsivity, *R* [2]:

$$I_{sig,PD} = \mathbb{R}P_{in} \tag{2.1}$$

• *Electrical Amplifiers*: The current generated by the photodetector is fed into a TIA which converts the current signals to amplified voltage signals. MA further amplifies the TIA output. The TIA and/or the MAs can also include equalization circuits to improve the signal quality. Some RXs have limiting amplifiers to get rail-to-rail voltage swing.

• *Decision circuits and De-serializer*: The last block of the receiver is a comparator which digitizes the analog signal and then the digital data stream is de-serialized. The comparator typically gets the clock from a clock and data recovery (CDR) circuit. It could also include equalization circuits to improve the signal quality.





2.3.1 Receiver noise and sensitivity

The equivalent circuit model of the PD is shown in Figure 2.5. The PD can be modeled as a signal current source, $I_{sig,PD}$, noise current source, $I_{noise,PD}$, and a capacitor, C_{PD} [2]. The characteristic behavior of both the current sources are described later in Chapter 4. The signal current of the PD is proportional to the input optical power and has a linear relationship. The noise current of the PD is dependent on signal and can be considered to be mostly white noise in spectrum. The input-referred noise of the electrical amplifier is dominated by the TIA noise and is represented as $I_{n,rx}$.



Figure 2.5 PD signal and noise model

The input-referred noise of the TIA often sets the sensitivity limits for the electrical front end. The optical sensitivity, $\overline{P_{op-sen}}$, is defined as minimum received optical power, averaged over time, necessary to achieve a specified bit-error-rate (BER) [2].

$$\overline{P_{op-sen}} = \frac{Q(i_{n,0}^{rms} + i_{n,1}^{rms})}{2R}$$
(2.2)

where $i_{n,0}^{rms}$ and $i_{n,1}^{rms}$ are rms noise of level zero and level one. Electrical sensitivity, i_{sen}^{pp} , is defined as the minimum peak to peak signal current at the input of the RX necessary to achieve a specified BER [2]:

$$i_{sen}^{pp} = Q(i_{n,0}^{rms} + i_{n,1}^{rms})$$
(2.3)

$$\overline{P_{op-sen}} = \frac{i_{sen}^{pp}}{2R}$$
(2.4)

2.3.2 Receiver bandwidth

Receiver BW is governed by the bit rate, *B*. Ideally an infinite BW would preserve the signal integrity without causing any distortion. But there is a tradeoff between noise and distortion. The white noise power is proportional to the BW. If the BW is too large, the noise integrated over this BW lowers the signal to noise (SNR) ratio, where as if the BW is too small, more ISI and distortions are generated. The optimum 3-dB bandwidth of the receiver, BW_{3dB} , for a NRZ data stream is often given by [2]:

$$BW_{3dB} \approx \frac{2}{3} B \tag{2.6}$$

Assuming it is a single pole system. When there are multiple blocks in the RX chain, the overall bandwidth is as shown below [2]:

$$\frac{1}{(BW_{3dB})^2} = \frac{1}{(BW_1)^2} + \frac{1}{(BW_2)^2} + \dots$$
(2.7)

where BW_1 , BW_2 are the bandwidth of individual circuit blocks. All individual blocks should have bandwidth greater than BW_{3dB} [2].

2.3.3 Equalization

The signal may undergo substantial distortions due to ISI by the time it reaches the decision circuits. Some of the causes for the ISI include optical fiber dispersions, limited electrical input BW of the APD-RX, and distortions caused by TIA and MAs. These distortions can result in increasing the Bit-Error Rate (BER). Figure 2.6 (a) shows an example of a distorted pulse response.



Figure 2.6 (a) Pulse response at the output of the TIA, and (b) DFE and IIR equalization

 P_0 is the main cursor, P_{-1} is the pre-cursor and P_1 and P_2 are the post cursors. Equalizations techniques such as DFE, as shown in Figure 2.6 (b), can be used to partially mitigate the effect of ISI by cancelling the post cursors. The decision circuit uses the knowledge of the previous quantized symbol to remove its ISI effect on the present symbol. The resolved data is subtracted/added to the incoming data stream based on the polarity and amount of equalization required to remove the ISI effect. If the effect of the ISI can be modeled as a pole (as shown in Figure 2.6 (a)) and implemented as the exponential tail of an RC filter, *infinite impulse response* (IIR), similar to a DFE tap can be used to further reduce the ISI, as shown in Figure 2.6 (b). DFE improves the SNR as it helps in improving the signal quality without boosting noise. However, DFE suffers from error propagation. If there is any quantization error, it is fed back into the system.

Chapter 3: Motivation

3.1 Working of a photodetector

Photodetectors work on the principle of photoconduction. In a semiconductor material, when energy of the incident photon is more than the energy band gap of the material (E(eV)), electronhole pairs are created generating additional free charge carriers. The phenomena of increase in free carriers and hence the electrical conductivity due to incident photons is called intrinsic photoconductivity. To increase the conductivity for a low energy-level incident photon, the band gap of the material must be lowered. This is done by adding impurity to the semiconductor material. Absorption from (or to) impurity sites in the gap creates free carriers in conduction or valence band, as shown in Figure 3.1.. This phenomenon is called extrinsic photoconduction.



Figure 3.1 (a) Intrinsic semiconductor and (b) Extrinsic semiconductor

The equation below shows the relation between the bandgap of the material and the wavelength of the incident photon (λ):

$$E = hv = \frac{hc}{\lambda} \tag{3.1}$$

$$\lambda = \frac{1240}{E} nm \tag{3.2}$$

The energy of the incident light is given by E = hv, where *h* is the plank's constant (6.63×10⁻³⁴ m²kg/s) and *v* is the frequency of the incident light. Bandgap energy is inversely proportional to the wavelength of the light. Silicon with bandgap energy of E = 1.12 eV can operate up to 1100 nm wavelength. Thus, silicon PDs are suitable for communication links which use 850 nm VCSELs. For operating at higher wavelengths, 1310 nm and 1550 nm, silicon PDs are modified by doping them with more expensive Ge or InP-InAsP to reduce the bandgap.

Not all of the incident photons lead to generation of electron-hole pair. A PD has a limited capability of collecting incident photons and converting them to electric current. Some incident photons are reflected by the surface and some are absorbed by the materials and lost in the form of heat. *Quantum efficiency*, η , is the measure of electrons produced per incident photon, which is usually expressed as a percentage [2].

$$\eta = \frac{no.of \ electrons \ generated}{no.of \ incident \ photon} \times 100$$
(3.3)

$$\eta = \frac{I_{PD}/q}{P_{in}/h\nu} \tag{3.4}$$

where, I_{PD} is the photo current generated by the PD, q is the electron charge (1.6x10⁻¹⁹ C) and P_{in} is the incident optical power.

Responsivity signifies the current gain of the photodetector, defined as the ratio of the photocurrent generated to the incident optical power [2].

$$R = \frac{I_{PD}}{P_{in}} (A/W)$$
(3.5)

From equation (3.4) and (3.5) we can show that the responsivity is a function of wavelength of incident light.

$$R = \frac{\eta \lambda}{hc} \tag{3.6}$$

For a wavelength of 850nm and a quantum efficiency of 0.64 [2], responsivity is 0.43 A/W.

3.2 Link budget

This section focuses on the link budget for a VCSEL-PD based optical link. An overview of an optical transceiver is first provided, and various factors which affect the link budget are described. Advantages and challenges of using on-chip PDs in comparison to off-chip PDs are also discussed.

3.2.1 Overview

Figure 3.2 shows a basic optical link. The link can be seen as an electro-optical TX connected to an opto-electronic RX. The optical modulator shown in Figure 3.2 is assumed to be a directly-modulated 850nm VCSEL in this Chapter.



Figure 3.2 A Basic block diagram of optical link

Some of the major considerations that affect the link performance concerning different link components are described below:

• *Modulator Driver*: In a VCSEL-based link, the VCSEL driver must be capable of providing high output current swing in order to achieve a higher difference between high and low optical power levels, defined as optical modulation amplitude (OMA). For complex
modulation schemes, such as 4-level Pulse Amplitude Modulation (PAM4), non-linearity of the VCSEL imposes difficulties on the driver design.

- *VCSEL*: Laser efficiency is defined as the ratio of output laser power to that of the input electrical power consumed by the laser. Higher efficiency of the laser significantly improves the total energy efficiency of the link.
- Optical Fiber: As the radius of the VCSEL's emission area is typically in the range of 20 µm, VCSELs can be very efficiently coupled to multimode fiber (MMF) that has a core diameter of 50 µm instead of single mode fiber (SMF) that has a core diameter of 9 µm. However, MMF suffers from high losses due to attenuation and dispersion. Though VCSELs are very popular because of their high performance and low cost, MMF limits the length the VCSEL-based links. Hence VCSEL-based links are used where reach is generally less than 300 m [3]. Furthermore, various coupling losses like VCSEL to fiber, connectors, and fiber to photodetector should be minimized to improve the link budget.
- *Photodetector*: The characteristics of the PD set the limit on both RX and TX design. The TX must be designed to meet the sensitivity of the RX. At the RX end, PD with higher responsivity relaxes the sensitivity requirements of the electrical front end of the RX. Recall from (3.5) that PD responsivity is a measure of its gain.

$$I_{PD} = RP_{in} \tag{3.7}$$

The relation between the optical sensitivity, $\overline{P_{op-sen}}$, and the electrical sensitivity, i_{sen}^{pp} , of RX is shown in equation (2.4).

Higher responsivity of the PD therefore significantly improves the RX optical sensitivity. Traditionally, external PDs implemented in III-V technologies have R in the range of 0.6 to 0.9 A/W [2]. One approach to boost the responsivity is to leverage the avalanche effect [4] and design an APD. APDs are described in the next Chapter. Besides responsivity, PD capacitance plays a role in setting the RX performance. The parasitics capacitance, C_{PD}, along with the packaging and ESD capacitances of PD limits the RX bandwidth. The bias of the PD also affects its bandwidth and gain, and hence must be tuned.

• *TIA*: Being the first electrical stage in the front end of the optical RX, TIA plays a crucial role in the link performance through its gain and bandwidth. A TIA is required to provide sufficient gain for the signal to meet the sensitivity requirements of the sense amplifiers. The noise of the TIA sets the sensitivity limit on the RX front end along with the PD noise.

3.2.2 Link analysis

Consider a typical VCSEL-based MMF link in a datacenter. The optical power output, P_{TX} (in dB), of the VCSEL with an efficiency, η_{VCSEL} , when directly modulated by the CMOS driver which consumes P_{TX_ELEC} (mW), can be shown as:

$$P_{TX} = 10 \log(\eta_{VCSEL} P_{TX_ELEC}) \tag{3.8}$$

VCSEL output signal is then coupled to an MMF, with a coupling loss, P_{MM-CPL} . MMF also introduces attenuation losses, $P_{MMF-att}$, and dispersion losses, $P_{MMF-dis}$ during transmission. At the Rx end, the signal see another coupling loss when the PD is connected to the fiber.



Figure 3.3 Losses in optical link

The link budget of such a link can be described by the following equation [3]:

$$P_{RX} = P_{TX} - 2P_{MMF-CPL} - P_{MMF-att} - P_{MMF-dis} - P_{pen} - P_{margin} = \frac{i_{sen}^{pp}}{2MR}$$
(3.9)

where, P_{pen} , includes the link penalty due to crosstalk, ISI and relative intensity noise, P_{margin} , is the link margin for the RX and P_{RX} represents the received optical power at the PD.

Consider a link with 300 m length of OM4 MMF with 3.5 dB/Km of $P_{MMF-att}$, P_{MM-CPL} of 1.1 dB each, $P_{MMF-disp}$ of 5 dB, P_{pen} of 4.8 dB and a 3 dB margin, 17% efficiency for the VCSEL and a

baseline R of 0.43 A/W for the PD. Figure 3.4 shows the VCSEL wall-plug power required for different values of RX electrical sensitivities for different PD gain.



Figure 3.4 Effect of RX sensitivity on laser power

Clearly, improving the responsivity of the PDs can significantly reduce the overall power consumption of the link. Conversely, for the same laser power, the design of the TIA can be considerably relaxed.

3.3 On-chip PD vs. off-chip PD



Figure 3.5 (a) External PD or APD with CMOS RX, (b) fully-integrated RX with APD in a modified CMOS process and external biasing.

Traditionally, PDs are implemented in a separate process compared to the CMOS RX to increase the PD performance, namely, -3dB BW of its frequency response, and its responsivity. PDs are generally made in expensive technologies such as Ge [5][6], GaAs [7][8] or InP-InGaAs [9] to enhance their performance. However, connecting an external PD (or APD) to a CMOS RX (Figure 3.5 (a)) using wirebonding or flip-chip assembly results in several issues – increase in manufacturing and packaging cost, possible decrease in yield, crosstalk between the bondwires especially when implemented as arrays of PDs connecting multiple RXs, requirement for ESD devices, additional packaging parasitics that further degrade the sensitivity of the RX, etc. Consider the system shown in Figure 3.5 (a) where an external PD (or APD) is followed by a transimpe dance amplifier (TIA) used as the gain stage to convert the PD current, I_{PD} , to voltages that can be further amplified by the main amplifiers. Considering an inverter based TIA, the input bandwidth of the receiver, BW_{in} , is given by [3]:

$$BW_{in} = \frac{1}{2\pi C_T \left(\frac{R_f}{1 + (g_{mn} + g_{mp})(r_{dsn} \parallel r_{dsp})}\right)}$$
(3.10)

Here R_f is negative feedback resistor, g_{mn} (g_{mp}) is the transconductance and r_{dsn} (r_{dsp}) is the output impedance of the NMOS (PMOS) transistor of the inverter, and C_T is the total input capacitance at the TIA input, $C_T = C_{PD} + C_{ESD} + 2C_{PAD} + C_{TIA,in}$, where C_{PD} , C_{ESD} , C_{PAD} and $C_{TIA,in}$ are parasitic capacitances of photodetector, ESD devices, pads and TIA input, respectively. The input BW of the RX is therefore inversely proportional to C_T , and additional capacitance from the pads and ESD significantly limit the maximum achievable data rate (MADR). For example, in 0.13-µm CMOS process, typical values of $C_{TIA,in} = 100$ fF, $2C_{PAD} = 160$ fF, $C_{ESD} = 200$ fF, and $C_{PD} = 100$ fF imply that the pads and ESD constitute about 72% of total input capacitance.

To overcome the aforementioned problems with a discrete PD, a fully monolithic CMOS PD can be implemented by using additional modifications to the CMOS process [10] (Figure 3.5 (b)). However, adding Ge to CMOS process to improve the PD performance increases manufacturing cost and complexity of fabrication, and is detrimental to the performance of CMOS transistors [11]. A Ge-based APD has high optical absorption only in 1.3-1.55 µm wavelength range, and is there not suitable for 850nm applications. Fully-integrated bulk CMOS PDs [12][13] have low responsivity, and are not very attractive for high-speed links.

On-chip PDs have many advantages over off-chip PDs. However, they suffer from low responsivity. To overcome this problem, avalanche gain in APDs can be exploited to improve the RX sensitivity or save TX power. In Chapter 4, we describe the operation of APDs and their impact

on optical links, then present our design and measurement results of APDs implemented in two different CMOS processes.

Chapter 4: Avalanche Photodetector

4.1 Working of an avalanche photodetector

To achieve a higher responsivity, avalanche effect in a PD can be used. An APD is operated at a high reverse bias voltage close to its breakdown, which causes generated electron-hole pair charge carriers to accelerate in the depletion region and produce additional carriers due to impact ionization (avalanching). As a result of the avalanche effect, the effective responsivity of the APD, R_{APD} , is increased [2]:

$$I_{APD} = M R P_{in} \tag{4.1}$$

$$R_{APD} = MR = \frac{i_{sen}^{pp}}{2P_{op-sen}}$$
(4.2)

where the avalanche multiplication gain, *M*, is given by [14]:

$$M = \frac{1 - k}{e^{-((1 - k)\alpha_e w} - k)}$$
(4.3)

where *w* is the width of the depletion region, α_e is the ionization co-efficient of the electrons and *k* is the ratio of the ionization co-efficient of holes (α_h) to the ionization co-efficient of electrons (α_e). The value of impact-ionization co-efficient, (α_h) and (α_e), depends on semiconductor material and on the electric field that accelerates electron and holes [14]. Assuming α_e to be 10^4 cm⁻¹ [14], *w* to be around 2 µm and *k* for silicon based APD to be 0.03 [2], we get *M* as 10. Assuming *R* is 0.43 A/W and *M* is 10, the APD has an effective responsivity of 4.3 A/W.

4.2 APD gain and link budget

The multiplication gain of the APD has a huge impact on the overall link budget. Assuming similar link parameters as in Section 3.2.2, and RX electrical sensitivities of -15 dBm and -10 dBm at 10 Gbps and 25 Gbps, respectively, Figure 4.1 shows the VCSEL wall-plug power required for different values of *M*. A baseline *R* of 0.43 A/W is assumed for the APD.



Figure 4.1 Effect of APD gain on RX sensitivity and Laser power

Clearly, use of APDs, and improving M for the APDs can significantly reduce the overall power consumption of the link. Conversely, for the same laser power, the design of the RX can be considerably relaxed.

Unfortunately, just like the signal, noise in the APD also experiences similar gain. Next, we describe different sources of noise in an APD.

4.3 Noise considerations in an APD-RX

Photodiode noise current: Noise current plays a major role in setting the sensitivity of the RX. APD noise current, $I_{noise,APD}$, consists of two main components – shot noise and thermal noise.

• *Shot noise* is white and is a function of the PD current, and the noise bandwidth, BW_{noise} . The shot noise current ($\overline{I_S^2}$ A rms) comprises of many short pulses distributed randomly with time, and given by [2].

$$\overline{I_S^2} = 2qM^2F(\mathbb{R}P_{in} + I_D)BW_{noise}$$
(4.4)

where *F* is the excess noise factor. The relationship between *F* and *M* can be shown to be [2]:

$$F = kM + (1-k)\left(2 - \frac{1}{M}\right)$$
(4.5)

where k is the ratio of ionisation co-efficient of electron and hole. For silicon APDs, k is approximately 0.02 to 0.05 [2].

Assuming *R* as 0.43, for P_{in} of -15 dBm, the desired PD signal current is $RP_{in}M = 12.9 \times M \mu A$, whereas the dark current component is only in the range of $5 \times M$ nA. So, in equation (4.4), the dark current component can be ignored.

• *Thermal noise*, $\overline{I_T^2}$, is a function of temperature, *T*, and is inversely proportional to the TIA input resistance (*R_{SH}*) which acts as the shunt load resistance for the PD. It is given by [14]:

$$\overline{I}_{T}^{2} = \left(\frac{4kTBW_{noise}}{R_{SH}}\right) \tag{4.6}$$

where k is Boltzmann's constant (1.38×10^{-23} J/K).

Assuming TIA input resistance to be 200 Ω and the bandwidth to be 5 GHz for a 10 Gbps RX, the RMS thermal noise is 0.64 μ A and can be reduced by increasing the input resistance of TIA at the cost of RX input BW. On the other hand, the input-referred RMS noise of a CMOS TIA, $i_{n,TIA}^{rms}$, for 10 Gbps is usually in the range of 1 to 2 μ A [2]. Hence the overall input-referred noise of an optical RX is often dictated by the input-referred noise of the TIA in comparison to the thermal noise of the APD.

Dark Current: PDs produce a small amount of current even when they are not exposed to light, called dark current, I_{Dark} . I_{Dark} depends on junction area, temperature and bias voltage. For traditional PDs, I_{Dark} is usually small, of the order of 5 nA [14]. In an APD, the internally generated dark current also undergoes amplification. However, dark current due to surface leakage (I_{ds}) does not see amplification, as it does not flow through the avalanche region [15]. The total dark current in an APD is given by [15]:

$$I_{Dark} = I_{ds} + MI_D \tag{4.7}$$

Assuming M > 10, I_{ds} can be ignored.

$$I_{Dark} \approx M I_D \tag{4.8}$$

The dark current in an APD-RX results in an current offset, and must be accounted for in the RX design for optimal performance.

Signal to Noise Ratio: Signal to noise ratio (SNR) is defined as the ratio of mean-free average signal power to the average noise power [2]. For a DC balanced signal, mean-free power is $(i_s^{pp}/2)^2$ [2]. Assuming equal number of ones and zeros, the noise power is calculated as sum of noises at zero and one, $(\overline{\iota_{n,0}^2} + \overline{\iota_{n,1}^2})/2$ [2].

$$SNR = \frac{\left(i_s^{pp}/2\right)^2}{\left(\overline{\iota_{n,1}^2} + \overline{\iota_{n,0}^2}\right)/2}$$
(4.9)

$$\overline{u_{n,1}^{2}} = \overline{I_{S}^{2}} + \left(i_{n,TIA}^{rms}\right)^{2} = \left(2qM^{2}F(\mathbb{R}P_{in})BW_{noise} + \left(i_{n,TIA}^{rms}\right)^{2}\right)$$
(4.10)

$$\overline{\iota_{n,0}^2} = \left(i_{n,TIA}^{rms}\right)^2 \tag{4.11}$$

$$SNR = \frac{(MRP_{in})^2}{2\left(\left(2qM^2F(RP_{in})BW_{noise} + (i_{n,TIA}^{rms})^2\right) + (i_{n,TIA}^{rms})^2\right)}$$
(4.12)

At very low bias voltage, the avalanche gain is negligible and hence the signal and the shot noise is minimal. The total noise at this lower bias voltage is dictated by the TIA noise. As the reverse bias increases, APD gain increases, so does the signal strength and shot noise. After avalanche breakdown, shot noise increases significantly as the gain is very high. There exists an optimal region for the bias voltage near the avalanche region where the SNR reaches its maximum, as shown in Figure 4.2.



Figure 4.2 APD signal and noise currents

The optimum gain, M_{OPT} , defined as the M which gives the maximum SNR, can be computed as:

$$\frac{d(SNR)}{d(M)} = 0 \tag{4.13}$$

$$M^{3} + \frac{(1-k)}{k}M = \frac{2(i_{n,TIA}^{rms})^{2}}{qkRP_{in}BW_{noise}}$$
(4.14)

An approximate solution can be shown to be:

$$M_{OPT} \approx \sqrt[3]{\frac{2(i_{n,TIA}^{rms})^2}{qkRP_{in}BW_{noise}}}$$
(4.15)

Figure 4.3 shows the SNR for the RX based on equation (4.16), assuming k = 0.03 for silicon APD [2] with a baseline responsivity of 0.43A/W and an incident optical power of -30 dBm. The RMS receiver noise is assumed to be 1.4 μ A at 10 Gbps [2].



Figure 4.3 SNR vs. APD gain

4.4 Design challenges for APD-RXs

Despite the advantages of CMOS APDs, there are certain drawbacks that have limited their practical use. Next, we describe these drawbacks, and propose solutions to overcome them.

4.4.1 Need for large bias voltage

APDs in bulk CMOS process require bias voltage of up to 10 V [4]. Even the Ge-based APDs require high bias voltages. Figure 4.4 shows the bias voltage requirement for different APDs. As the nominal voltage supplies on bulk CMOS processes are typically limited to 3.5 V (for I/Os), this has resulted in APDs being used only as external components. In Section 5.4, we present an on-chip voltage booster to generate the required APD voltage in a standard CMOS process.



Figure 4.4 Reverse bias voltage vs. bandwidth of APD.

4.4.2 Limited gain-bandwidth of APD

Bandwidth of the APD, BW_{APD} , is dependent on three main factors: (i) the parasitics of the APD (R_{APD} , C_{APD}), (ii) the transit time (τ_{TR}) of the photocarriers in the depletion region, and (iii) the slow diffusion component in the photocurrent [14].

The parasitics of APD is mainly because of reverse biased APD junction capacitance C_{PD} and the electrical contact resistance R_{PD} . The packaged APD also has additional parasitics due to pads, ESD, and bondwires. Ignoring the packaging parasitics, the time constant of the APD, τ_{RC} , can be shown as [2]:

$$\tau_{RC} = R_{APD} C_{APD} = R_{APD} \left(\frac{\varepsilon A}{W} \right)$$
(4.11)

where ε is the permittivity, *A* is the area of active region and *w* is the width of the depletion region. The transit time can be defined as the time taken by the carriers to drift through the depletion region. Assuming the drift velocity is v_d and the depletion width is *w*, the transit time constant can be shown as [2]:

$$\tau_{TR} = \frac{w}{v_d} \tag{4.12}$$

The photon absorption outside the depletion region results in slow diffusion current. The free carriers generated outside must diffuse to the depletion region boundary, which is a slow process. This leads to pulse spreading [14]. There are many techniques to prevent the slow diffusion current from reaching the depletion region. One of the techniques is to increase the depletion region width so that most of the incident light is absorbed within the depletion region [14], hence minimizing the effect of slow diffusion current on the BW of the APD. The APD BW, BW_{APD} , is, therefore, mainly a function of its parasitics and transit time constants and can be shown as [2][14]:

$$BW_{APD} = \frac{1}{2\pi(\tau_{TR} + \tau_{RC})}$$
(4.18)

$$BW_{APD} = \frac{w}{2\pi \left(R_{APD} \varepsilon A + \frac{w^2}{v_d} \right)}$$
(4.19)

With increase in the reverse bias voltage, the width of the depletion region, *w*, increases. Thus, more electron-hole pairs are generated, resulting in higher multiplication gain of the APD. The

relationship between *M* and *w* is given by equation (4.3). However, generation and collection of the secondary electron-hole pair (carrier multiplication) takes additional time [14]. This additional time requirement for generation of electron-hole pair cannot be achieved at high frequency which results in gain reduction [14]. This sets the trade-off between gain-BW product [14]. Thus, to a first order, changes in reverse bias voltage do not impact the gain-BW product for an APD [2]. Typical gain-BW product for a silicon APD ranges from 300 to 800 GHz [4][10][16].

4.4.3 Sensitivity to reverse bias voltage and temperature

The gain of the APD fluctuates with change in the chip temperature and the applied reverse bias[15]-[18]. Performance of the APD is also subjected to random process variations during its manufacturing.

Several ideas have been proposed in prior-art to maintain steady gain and BW of APD using temperature monitoring and compensation techniques. These systems sense the temperature and compensate for the variations either by changing the bias voltage of the APD or by heating or cooling the APD. [20] uses the cooling characteristic of a Peltier cell in conjunction with a thermistor to maintain a constant low temperature for APD, as shown Figure 4.5(a). This implementation is, however, difficult to realize in standard CMOS technology. [21][22] uses a temperature sensor in the vicinity of APD to maintain a stable bias voltage with a pre-tabulated temperature vs. APD bias voltage data, as shown in Figure 4.5(b). The pre-tabulated data is obtained with a standalone APD. However, generating an accurate pre-tabulated data for a monolithic implementation where the temperature of the APD and the RX changes together over



Figure 4.5 (a) Heating/Cooling APD. (b) Temperature sensors to monitor APD temperature and alter bias voltage. (c) Dumpy APD to stabilize APD bias voltage.

time is difficult. [20] and [23] uses thermistor based logic to compensate for the change in temperature by changing its bias voltage.

[24] uses clock and data recovery (CDR) logic to make decisions based on eye quality. This technique increases complexity and is only applicable in implementations where the CDR is implemented on the same chip. [25] and [26] uses matched APDs, one biased in unity-gain region and another biased in high-gain mode through a feedback loop to attain constant multiplication gain, as shown in Figure 4.5(c). The main difficulties in these implementations are the required matching of APDs, defining and maintaining unity-gain bias.

In Chapter 5, we propose the first fully-integrated optical RX with on-chip biasing, tuning and stabilization of APD in a standard CMOS process using nominal voltage supplies. The details of the implemented APDs are provided next.

4.5 CMOS APD design

APD is basically a PN diode which is reverse biased. Figure 4.6 and Figure 4.7 show the N+/P well and P+/ Nwell based APDs implemented in 130nm CMOS process, respectively. At the junction edges, high concentration of electric field can result in premature breakdown. This can result in lower breakdown voltage and hence lower multiplication gain. In order to facilitate high electric field concentration at the junction, guard rings are used. In [27], various guard ring structures are discussed and it is shown that APD with shallow trench isolation (STI) guard rings have best performance. Therefore, in our designs, we have used STI guard rings to enhance breakdown voltage and to provide higher avalanche gain.

N+/Pwell junction based APD suffers from slow diffusion current at Pwell – Psubstrate junction. When light penetrates into the depletion region, it creates electron-hole pairs leading to diffusion current. The penetration depth of light in silicon is around 20 μ m which is more than the depth of the depletion region (around 2 μ m). Thus, there is a very high chance for the light to enter the P-substrate and produce electron-hole pair which results in a slow diffusion current. When this slow diffusion current reaches the depletion region, it will increase the total signal current and hence improve responsivity. On the other hand, due to slow diffusion, the bandwidth of the APD is reduced. We use Deep N-well to shield Pwell from Psubstrate to prevent the slow diffusion current from reaching the depletion region in order to boost the bandwidth of the APD.

4.6 CMOS APD measurements

APDs are fabricated and measured in 130 nm and 65 nm CMOS processes. N+/P well/Deep Nwell and P+/ Nwell/Deep Nwell based APDs, shown in Figure 4.6, are fabricated in 130 nm. However, because of pad placements (discussed in Chapter 6), P+/ Nwell APDs could not be tested. N+/P well and P+/ Nwell based APDs fabricated in 65 nm CMOS are shown in Figure 4.7. In the 65 nm CMOS technology, available to us, a deep well is not present. Thus, the N+/P well showed a lower bandwidth due to the slow diffusion current reaching the diffusion region.

Different structures of APD are fabricated with varying optical opening window $(l \times l)$. The effect of *l* on the capacitance of the APD is given by:

$$C_{APD} = \left(\frac{\varepsilon l^2}{W}\right) \tag{4.21}$$

Thus, a smaller APD opening, *l*, helps in reducing the capacitance, but also result in alignment difficulties.

The measurement set-up is shown in Figure 4.8. An 850 nm VCSEL die is used as an optical source and its output is coupled to the fiber using a collimated lens set-up with a coupling efficiency of 50%, where the optical power of the VCSEL output is measured using an optical spectrum analyser (Agilent 86146B). The coupled light of -12 dBm is flashed on the APD using a 50/125µm multimode lensed fiber, where the coupling loss between the fiber and the APD

interface is approximately 2 dBm. Next, the VCSEL is modulated with an alternating data pattern (1010) from a pattern generator (Anrtisu MP1800A) and using the setup shown in Figure 4.8, the output of the APD is directly connected to an electrical spectrum analyser (Rohde & Schwarz 26.5 GHz FSW Spectrum Analyser).

APD IV curve and gain, *M* vs. reverse bias voltage curve is shown in Figure 4.9 and Figure 4.10 respectively for 130nm N+/P well/Deep Nwell. At low bias (0.2 V), the gain is almost unity. The gain of the APD is calculated as the ratio of photocurrent at a given bias to the photocurrent at 0.2 V. The gain of the APD near avalanche breakdown is measured to be 778. Figure 4.11 shows the effect of bias change on APD's normalized large-signal frequency response. Because of the low incident power on the APD, and lack of an external pre-calibrated TIA in our measurement setup, small-signal frequency response could not be measured due to low output current levels. It also





(b) APD Top veiw

(c) APD P+/Nwell cross section

Figure 4.6 APD structure fabricated in 130 nm CMOS

posed a restriction for pulse response, which is often an alternative method to characterize the small-signal frequency response of a PD. Instead, the VCSEL is modulated with alternating data pattern from 0 to 0.8 V, and the output electrical current from the APD is observed on a 50 Ω spectrum analyzer. As the frequency of the alternating data pattern is swept from few hundreds of MHz to GHz, the output power is recorded, and thus, a large-signal frequency response is obtained. Normalized frequency responses of 65 nm APDs with core area of 40 μ m × 40 μ m, 30 μ m × 30 μ m and 20 μ m × 20 μ m are shown in Figure 4.12. The die images are shown in Figure 4.13 and Figure 4.14. Table 4.1 compares the performance of the APDs to prior-art.



Figure 4.7 APD structure fabricated in 65 nm CMOS



Figure 4.8 Frequency response measurement set-up for APD



Figure 4.9 IV curve of N+/Pwell/Deep Nwell APD fabricated in 130 nm CMOS



Figure 4.10 M vs. reverse bias voltage of the APD fabricated in 130 nm CMOS



Figure 4.11 Effect of reverse-bias on APD bandwidth in 130 nm CMOS



Figure 4.12 Effect of core area on APD bandwidth in 65 nm CMOS



Figure 4.13 Die image of APDs in 130 nm CMOS

	[MJLeeOE2010]	[jsyung2012]	[JSYoun15]	This work	
Process	65 nm	130 nm	65 nm	130 nm	65 nm
APD	N + /P-well Deep	P+/Nwell	P+/Nwell	N + /P-well	P+/Nwell
Structure	Nwell			Deep Nwell	
Optical Window	30 x 30 μm	10 x 10 μm	10 x 10 μm	30 x 30 µm	20 x 20 μm
Bandwidth (GHz)	3.2	6.3	5	3.5*	6.2*
Gain (M)	569	-	14.4	987	

Table 4.1 Performance summary of CMOS APDs and comparison to prior art

*large signal bandwidth



Figure 4.14 Die image of APDs in 65 nm CMOS

Chapter 5: Monolithic CMOS optical receiver

In this chapter, we propose a fully-integrated optical RX with on-chip biasing, tuning and stabilization of APD in a standard CMOS process with nominal voltage supplies. Figure 5.1 presents a simple block diagram of the proposed system, where the APD is designed in the same



Figure 5.1 Fully-integrated APD-RX in standard CMOS process.

bulk CMOS 130nm process as the RX circuits in order to reduce cost, manufacturing complexity, bondwire crosstalk, and parasitics. Due to monolithic integration, no external signal pads or additional ESD structures are needed. The APD-RX system can be broadly divided into two segments – high-speed data path and APD bias-generation and control path, as shown in Figure 5.2. The slow drift of the TIA output provides the temperature-dependent variation of the APD characteristics. A control block uses this information to stabilize the bias of the APD. Next, we describe the different blocks in more detail.

5.1 Transimpedance amplifier

The high-speed RX path is shown in Figure 5.2. A TIA converts the current signals from the APD to voltage signals, followed by main amplifiers (MAs) for further voltage amplification. As TIA is the first block in the electrical RX link, gain, noise and bandwidth of the TIA are imperative to the overall receiver's performance.



Figure 5.2 Front end of the APD-RX. The high-speed circuits are designed by A. H. Ahmed.

The Friis equation shown below gives the noise figure of receiver:

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \dots + \frac{F_N - 1}{G_1 G_2 \cdots G_{N-1}}$$
(5.1)

where F_i and G_i are noise figure and power gain of ith block in the link. For the optical receiver, we have:

$$F = F_{TIA} + \frac{F_{BALUN} - 1}{G_{TIA}} + \frac{F_{MA} - 1}{G_{TIA}G_{BALUN}} + \frac{F_{BUF} - 1}{G_{TIA}G_{BALUN}G_{MA}}$$
(5.2)

44

where F_{TIA} (F_{BALUN} and F_{MA}) and G_{TIA} (G_{BALUN} and G_{MA}) are the noise figure and gain of TIA (Balun and MA), respectively. From equation (5.2), for receivers where the gain of the TIA is large, the noise contribution of the TIA is dominant, and noise of the successive stages are attenuated by G_{TIA} . Minimizing the Input referred noise of the TIA would improve the overall noise figure of the receiver. If the gain of the TIA is not sufficiently large, as is usually the case for CMOS TIAs with limited supply voltages, the noise performance of the second stage is also important.

The -3dB transimpedance BW of the TIA, BW_{TIA} , plays a crucial role in setting the overall BW of the RX. The input bandwidth of the receiver, $BW_{TIA,i}$, for an inverter based TIA, can be shown to be inversely proportional to C_T , the total input capacitance at the input of the TIA [3]. This input capacitance sets a limit on the TIA bandwidth. A good discussion for an inverter-based TIA design is given in [28]. The output of the PD is single-ended and is sensitive to ground bounce. As differential designs are preferred in an RX, a differential TIA driven by a PD and a dummy PD can be implemented [29], where the dummy PD does not have any light incident on it. A more common architecture is shown in Figure 5.3. The single-ended output of the TIA is connected to a differential amplifier, with the other input of the differential amplifier connected to a replica TIA. However, these methods cause mismatch in gain and phase of the differential output, resulting in asymmetric signals. Moreover, dummy TIA also increases the power consumption of the RX. Let us consider the noise performance of the differential amplifier in Figure 5.3 in more detail.



Figure 5.3 A differential amplifier to convert single-ended TIA output to differential signals for the MAs

Transistor M1 is connected to the output of the TIA and M2 is connected to a dummy TIA for better matching. Transconductance (g_m) and the load (R) of both the transistors M1 and M2 are assumed to be matched. The output referred noise, $\overline{V_{o,n}^2}$, and the input referred noise, $\overline{V_{l,n}^2}$, of the differential amplifier can be calculated as follows:

$$\overline{V_{o,n}^2} = 2[4kTR + 4kT\gamma g_m R^2] = 8kTR[1 + \gamma g_m R]$$
(5.3)

$$\overline{V_{l,n}^2} = \frac{8kTR[1 + \gamma g_m R]}{(g_m R)^2} = \frac{8kT[1 + \gamma g_m R]}{g_m^2 R}$$
(5.4)

In the proposed system, a single-ended inverter-based push-pull TIA is followed by a self-noise cancelling active balun to convert the single ended TIA output to differential signals. Figure 5.4 shows the active balun implementation, inspired by noise-cancelling low noise amplifiers [30].



Figure 5.4 A noise-cancelling active balun to convert single-ended TIA output to differential signals for the MAs. Designed by A. H. Ahmed.

 V_{in} represents the single-ended signal from the TIA, R_s is the output resistance of the TIA, V_B is the gate bias for the common-gate transistor M2, and V_P and V_N are differential signal outputs. Here, g_{m1} and g_{m2} are transconductance of transistors M1 and M2, respectively. R_4 and R_3 are the loads seen by M1 and M2 transistors and can be approximated to $1/g_{m4}$ and $1/g_{m3}$, respectively. It can be shown that for matched swings at V_P and V_N , $g_{m1}R_4 = g_{m2}R_3$. It can also be shown that, if $R_s = 1/g_{m2}$, the effective noise of M2 is cancelled. The gain of the balun, A_{Balun} , can then be shown to be:

$$A_{Balun} = g_{m2}R_3 \tag{5.5}$$

Considering noise from other sources, M1, R_3 and R_4 , we can calculate the output referred noise, $\overline{V_{0,n}^2}$, and the input referred noise, $\overline{V_{1,n}^2}$, as follows:

$$\overline{V_{o,n}^2} = 4kTR[2 + \gamma g_m R] \tag{5.6}$$

$$\overline{V_{l,n}^2} = \frac{4kT[2+\gamma g_m R]}{g_m^2 R}$$
(5.7)

Comparing this with equation (5.8), the input referred noise of the active balun is less than that of a differential amplifier with dummy TIA. The active balun is followed by 4 stages of MAs, with each stage implemented as a differential amplifier employing shunt-peaking active inductors for bandwidth extension [31]. The amplified signal is finally buffered to the output using 50 Ω drivers for measurement purposes. The output signal has a peak-to-peak signal swing of 300mV.

The output of the MA is low pass filtered and sent to an error amplifier for offset cancellation. The low pass filter (LPF) is implemented as a simple RC filter with a fixed capacitance of 2.3 pF and a 4-bit thermometric tunable resistor bank with four 12 k Ω MOS-resistors.

The power consumption of the RX is summarised in Table 5.1. The measured eye diagram for the RX, for 8Gbps PRBS7 electrical input, is shown in Figure 5.5. The electrical -3dB BW of the RX is measured to be 4.6 GHz, as shown in Figure 5.6. The optical source (VCSEL) was damaged during the measurements, so an optical eye-diagram could not be measured. The die image is shown in Figure 5.7.

Circuits	Power Consumption (mW)		
TIA	1.68		
Balun and MA	3.84		
Total Power	5.52		
Output Buffers	25.32		

Table 5.5.1 RX power consumption



Figure 5.5 Electrical eye diagram of the RX at 8 Gbps



Figure 5.6 Electrical S21 measurements of the RX

5.2 Control loop

Figure 5.8 shows a detailed block diagram of the entire APD-RX system. A closed control feedback loop tunes the voltage bias of the APD and stabilizes it for temperature variation. A voltage booster inside the loop generates the large voltage bias needed for the avalanche operation of the APD.



Figure 5.7 Die image of TIA fabricated in 130 nm CMOS

As described in Chapter 3, the performance of APD is sensitive to the reverse bias voltage and temperature, and all of the prior-art in bias stabilization circuits have been limited to off-chip 50



Figure 5.8 Block diagram of the proposed APD-RX system.

implementations. Typically, a shift in temperature by 1°C changes the APD bias by 0.05% [10] for Silicon APD and 0.2% [2] for Ge APD. To maintain a stable performance, either a constant operating temperature can be maintained for the APD, or the bias voltage of the APD can be tuned. Heating/cooling APD is power inefficient, bulky and not easily compatible for CMOS monolithic applications.

Based on the fact that any change in temperature leads to change in APD I-V characteristics [4][15][18], no temperature sensors are needed in the proposed system. A change in the biasing of the APD due to temperature variations leads to a change in its responsivity, and therefore, the overall gain of the CMOS receiver is also affected. The average of the signal extracted at the output of the MA using a low pass filter has the information of the varying responsivity of the APD. Because the high-speed signal path inherently has a low pass filter in the offset correction loop so as to effectively ac-couple the incoming high speed current from the APD, the output of this LPF is
tapped by both the error amplifier of TIA, as well as a 7-bit successive approximation registerbased (SAR) ADC. The SAR ADC is briefly described in Section 5.3. The digitalized output from the ADC is processed by the control logic to generate a control voltage, V_{REF} , from a 7-bit Digital to Analog Converter (DAC). V_{REF} is then boosted by the voltage booster as described in Section 5.4, and applied to the APD as a reverse-bias.

The output signal current of the APD shows a steady rise with increasing reverse-bias voltage till the breakdown voltage. Beyond the breakdown voltage, there is an exponential increase in the noise current leading to saturation of the RX. If we consider the low pass envelope output as a function of the reverse-bias voltage, we see that there is a steady increase in derivative of the slope till the photodetector reaches the avalanche region. Near the avalanche region there is a sudden increase in the derivative of the slope as a result of avalanche effect, and beyond the avalanche



Figure 5.9 (a) APD IV curve shows APD breakdown due to very high biases and (b) Derivative of the average TIA output vs. the reverse bias voltage



Figure 5.10 Control logic FSM for APD bias stabilization

region the derivative of the slope reduces as the current becomes nearly linear. Thus, the plot for the second derivative of the low pass filter output ($\Delta^2 V_{LPF}$) vs. the reverse-bias voltage peaks near the avalanche region as shown in the Figure 5.9 (a).

A modified hill-climbing algorithm is implemented in the digital control to track the peak in the second derivative of LPF output which is a function of the bias voltage of the APD. In Figure 5.9 (a), the yellow circle indicates the avalanche region. Note that if the reverse-bias voltage of the APD is increased much beyond its avalanche point, the diode can get permanently damaged due to very high currents flowing through the APD as shown in Figure 5.9 (a). Therefore, a major difference between conventional hill-climbing algorithms and the proposed algorithm is that the algorithm takes the slope at avalanche, Δx_{ref} , as input and doesn't allow the system to overshoot this value. The reference slope, Δx_{ref} , is fed to the control logic based on the characteristic



Figure 5.11 Control logic cases: (a) locking at the avalanche region (b) dithering at the top (c) positive temperature drift, and (d) negative temperature drift

calibration of the APD. The control algorithm also takes into account shifts in I-V curve because of temperature effects while setting the bias of the APD.

The logic implemented for the control algorithm has four states – START, UP, DOWN and WAIT. Figure 5.10 shows the FSM of the control logic. The system is reset at the START state. In the UP or DOWN state, the control voltage is incremented or decremented respectively. In the WAIT state, the control voltage is kept steady. The step size of the control voltage may be incremented or decremented in steps of one or two to achieve faster settling.

At every clock cycle, the present ADC value, x(n) is compared with the previous value, x(n-1) and the difference between them is computed, Δx . This difference is then compared with the reference slope, shown in Figure 5.11 (a), $[\Delta ref = \Delta x_{ref} - |\Delta x|]$. The control remains in UP state until it reaches 54



Figure 5.12 Transient response of APD-RX system. (a) APD current, (b) LPF output and (c) APD bias voltage.

the avalanche region. Once it reaches avalanche it either moves to wait state or dithers at the top based on the value of Δ ref as shown in Figure 5.11 (a) and Figure 5.11 (b). The cases for positive and negative temperature drift are highlighted in Figure 5.11 (c) and Figure 5.11 (d), respectively. As shown in Figure 5.11 (c), if there is a rise in temperature, the bias voltage is increased to lock to the avalanche region of the shifted curve. Similarly, the loop adjusts the bias voltage in case of a temperature fall.

The control logic is written in Verilog and synthesized. Total area occupied by the synthesized control logic is 100 μ m×100 μ m, clocked at 8 MHz. Simulation results for APD-RX system are shown in Figure 5.12. Figure 5.12 (a) shows the high-speed output current of the APD as a function of time. Figure 5.12 (b) plots the LPF output, which is equivalent to the average value of the APD

current. As seen in Figure 5.12 (c), the control loop ensures that the bias voltage of APD is slowly increased till it reaches the avalanche region and then is kept steady. The ripple after settling is 5mV.

5.3 Analog-to-digital converter

The output of the LPF used in the offset-cancellation feedback path also extracts the slow DC drift associated with the data signal due to APD bias variation. The filter cutoff is decided based on the required range of offset cancellation of TIA and rate of APD bias variation. The extracted DC drift is then fed into a SAR ADC, as shown in Figure 5.13. The SAR logic is coded in Verilog and synthesized. The total area of the SAR logic is 150 μ m × 150 μ m.A 7-bit binary-weighted capacitor bank, C0 to C7, is used in the DAC. Each capacitor is implemented as a sum of parallel unit capacitors with a unit capacitance, C₀, of 56 fF. This the minimum mimcap available in the 130 nm CMOS technology and dictates the overall size of the DAC. Each unit capacitor is accompanied by its switch with either connects it to supply or ground based on the control bit from the SAR logic. To optimize the area of the DAC, switches are placed under the unit capacitors. Under ideal conditions the resolution of an *N*-bit ADC is 2^{*N*}. However, due to mismatch, noise and distortions the resolution of ADC is reduced, represented by an effective number of bits (ENOB).

Design and layout of the DAC capacitor bank determines the linearity of the ADC. Any mismatch in the capacitor bank increases non-linearity leading to spurs and degrades the ADC ENOB. Mismatches can be broadly classified as systemic, gradient or random.



Figure 5.13 Block diagram of the SAR ADC. Comparator and SAR logic is designed by M. Al-Taha.

Systemic mismatch arises due to asymmetric circuit or layout design. Mismatch in wire length/loading and use of different capacitor/resistors leads to mismatch. This could be minimized by careful design techniques. While designing a capacitor bank to minimize systemic mismatch, multiple replicas of smaller unit capacitors are used instead of a large single capacitor. This ensures better matching of the capacitor bank.

Gradient mismatch is the mismatch over longer lengths across the chip arising during the chip fabrication. This can be minimized by following simple layout techniques. Devices are placed close to each other to ensure similar working temperature and supply. The size and the orientation of the unit devices are kept uniform. Common-centroid arrangement is followed, which reduces gradient mismatch and parasitic mismatch in layout. As shown in the Figure 5.14, capacitors are arranged in a square matrix with smaller capacitors near the center and larger capacitors at the

outer edge. Black squares represent dummy capacitors placed to ensure symmetric fringe capacitances between the capacitor plates.



Figure 5.14 Common-centroid arrangement of 7-bit capacitor DAC.

Random mismatch mainly arises because of variation in process parameters during lithographic process. Device length, doping concentrations, sheet resistance/capacitance, and etching length vary during fabrication. One way to minimize the effect of random mismatch is to increase the area/perimeter ratio. Square shapes assure minimum etching mismatch error for a given area. Thus, capacitors are kept square to minimize over-etching mismatch. Dummy capacitors prevent uneven lithography from encroaching on the array. A dummy capacitor perimeter provides a better matching, however, it is not used due to area limitations.

An identical 7-bit capacitor DAC, as used in the ADC, operates on the output of the digital control logic to drive the voltage booster as shown in Figure 5.8.

Parasitic capacitance from the top plate of the capacitor array, C_x , at node X in Figure 5.13, affects the overall performance of the ADC and degrades the ENOB. After the first-pass layout, C_x was extracted as 1.2 pF, which degraded the ENOB to < 2. To minimize the parasitic capacitance of the routing, all wire of top-plates of the capacitor array are made to be of minimal width and the unit capacitors are kept at the minimum distance. Top metal layers are used for routing, but the top most layer is avoided as the minimum width design rule check (DRC) requirements for the top metal layer is very large in the 130 nm CMOS technology available to us. After several layout iterations, C_x is reduced to 488 fF, with an ENOB of 4.4. Both integral non-linearity (INL) and differential non-linearity (DNL) are simulated to be less than 1.



Figure 5.15 Die image of ADC and controller

5.4 Voltage booster

In order to provide the high reverse-bias voltage needed for the APD from an external supply of 2.5V, a fully integrated voltage booster is implemented to generate up to 12 V (VDD_{HI}). Figure 5.16 (a) shows the simplified schematic of the proposed voltage booster. The power for the voltage booster is provided through, V_{REF} , which can be swept from 0.5V to 2.5 V. V_{REF} can also be controlled from the output of a DAC, as described later in Section 5.3. A bias voltage, V_{bias} , controls the dropout across the transistor M1 to provide a variable supply, V_{DDX} , to a pair of inverters driven by differential clock phases, CLK and CLKB. Although provided externally in our prototype, these differential signals can be easily generated by an on-chip ring oscillator. The output of the inverters swing from 0 to V_{DDX} , and can therefore be varied as needed. V_{DDx} also provides the input voltage to the core of the voltage booster.

The core of the voltage booster is based on a Dickson voltage multiplier [32]. Figure 5.16 (b) shows a conventional NMOS based Dickson voltage booster in which diode-connected NMOS transistors are connected in series, and the intermediate nodes share capacitors. The bottom plate of these capacitors are connected to differential phases of a clock in an alternating fashion. A disadvantage of the traditional Dickson architecture is that the diode connected NMOS transistors turn-off when the gate-source bias falls below the threshold voltage V_{TN} . The threshold drop can be compensated by using a back-compensated voltage booster implemented using PMOS transistors, as proposed in [33]. A PMOS transistor requires a negative gate-source bias voltage is provided by a diode connecting the PMOS gate to the source of the previous stage instead of the



Figure 5.16 (a) Voltage booster circuit (b) Conventional Dickson voltage multiplier [32] (c) Backcompensated voltage multiplier [33] and (d) Implemented voltage booster core. The voltage booster is designed by A. Sharkia.

traditional diode connection. The voltage of the intermediate nodes increases across the stack where the top plates of the capacitors are connected. As the bottom plates of the capacitors are fed with differential signals of the same swing, these charge pump architectures impose a high voltage stress on the capacitors of the last few stages. A modified version of the threshold-compensated PMOS rectifier based voltage booster core is shown in Figure 5.16 (d), where voltage drops are limited to no more than 2.5 V (maximum value of V_{REF}) between any two nodes of any of the capacitors or transistors. Furthermore, thick-oxide I/O transistors and MIM capacitors are used. This allows the proposed circuit to achieve high voltages in standard CMOS process without reliability issues.

The proposed voltage booster has 11 cascaded stages. Ideally, each stage should increase the voltage by $V_{REF} - V_{TP}$, however, due to leakage and parasitics, the amount by which the voltage is increased in each stage diminishes as more stages are added. The voltage booster works with a wide range of clock frequencies, from 120 MHz to 2.4 GHz. Higher frequencies are preferred to minimize the output ripple and the overall circuit area. However, synthesis of a high frequency clock consumes power, leading to a tradeoff between power consumption and output ripple and area of the voltage booster. To further reduce the ripple two boosters cores can be used in parallel with opposite clock phase. The output of the voltage booster can be varied from 0 V to 12 V through V_{bias} (V_{DDX}). The voltage booster occupies an area of 330 µm × 140 µm, mostly limited by the area of the capacitors. The measurement result of the voltage booster is shown in Figure 5.17. With a nominal I/O supply of 2.5 V in this process, the voltage booster output varies from 0 to 7.42 V. The power consumption of the voltage booster is 2.4 mA. The die image is shown in Figure 5.18.



Figure 5.18 Measured output vs. input voltage characteristic for the voltage booster.



Figure 5.17 Die image of voltage booster fabricated in 130 nm CMOS

Chapter 6: System design considerations

In this section, we discuss various challenges faced during the design, layout and measurement of the APD-RX system, and the steps taken to solve them. An optical probe-station was custom-designed for carrying out the 850 nm optical measurements, and key considerations are described here.

1nHXl R .000,

 Length of bondwire in mm
 R – Resistance depends on frequency (1Ω for 10Gbps)



Figure 6.1 (a) Bondwire model and (b) Bondwire simulation model

6.1 Packaging model

CMOS dies are generally packaged for commercial applications. The pads of the CMOS die are electrically connected to the package through wirebonds or flip-chip bumps. Wire bonds are most commonly used for packaging as it is cheaper. Gold or Aluminum bondwires of $25-250 \ \mu m$ in diameter is usually used depending upon the application. In this design, we used CQFP80 pin package with gold bonds. These wirebonds have low resistance but are dominated by inductance. The estimated parasitic inductance is around 1nH per millimeter of its length.

In a differential circuit, a fixed DC current drawn from the supply is steered between two branches. On the other hand, in single ended circuits, when an AC current is drawn from the power supply through the bondwires, data dependent fluctuation is experienced at the supply node of the chip, as shown in Figure 6.1(b). These voltage fluctuations severely affect the performance of high speed circuits.

The package model was not included during the chip design, which severely affected the measured performance. The bondwire inductance for the CQFP80 package used in our design is estimated to be 5 nH. The performance of RX with and without bondwire model in simulation is shown in Figure 6.2.

65





Figure 6.2 Effect of bondwire

6.2 Latchup

During the measurement of the first prototype chip, we found excess current being drawn between supply and ground due to a low resistance path causing permanent damage. Upon investigation, we realized that we had ignored many latch-up DRC errors. Furthermore, the ESD structures were not sufficient to provide the required protection.

Latchup and various techniques to mitigate them are discussed in detail in Appendix A. For the second protoype, all of the recommended guidelines as prescribed by the DRC file, and discussed in the Appendix A, were employed to minimize the occurrence of latchup.

6.3 Printed circuit board

As discussed in Section 6.1, wire-bond affects the performance of high speed systems. Probing the die eliminates the need for wire-bonds and is generally preferred for rapid testing of designs which are sensitive to bond wire or in systems where the length of wire-bond cannot be customized. Systems with few I/O pads can be tested using a semiconductor probe station, where RF probes are used for high speed I/O pads and DC wedges ranging from 12 to 20 pins are used for power supply, biases and control signals (up to few tens of MHz). However, with the increase in number of I/O pads, it becomes cumbersome to perform probe-test. As our prototype had large number of I/O pads, it was packaged and tested on a printed circuit board (PCB) fixture. However, our first prototype chip lacked accessibility to many of the intermediate nodes, which made the debug difficult during the test process. In the second design, standalone test structures and the APD-RX

system with visibility to internal nodes were included. TIA output and optical input to the PD were probed, while rest of the pads were packaged. A PCB was designed for testing the chip, as shown in Figure 6.3. A 2-layer board with 1 oz copper trace thickness was designed, where the top metal layer was used for routing and rest of the area was covered with ground plane. The bottom layer was mainly used as ground shield. Top and the bottom ground plan were shorted using via arrays.



Figure 6.3 PCB test fixture

Following are some of the considerations behind layout and PCB design and testing:

If some of the pads are probed and some are wirebonded for testing, care must be taken to during CMOS layout and PCB layout to facilitate probing. If the probing pads are too close to the wirebonded pads, probe heads could push the wirebonds leading to shorts or even damage the wirebond. Based on probes used (Cascade GSSG, Cascade GSG, GGB GSSG), 250 µm is a safe buffer distance to maintain between probing pads and the nearest adjacent wirebond pads in all directions during layout, as shown in Figure 6.4.



Figure 6.4 PAD clearance for wirebonding and probing

- VCSELs and APDs require connection to optical fibers. Optical fibers should also be considered as probe and given a buffer of 250 μm.
- Components with a height more than a couple of millimeters (headers, SMA, MMCX, potentiometers etc.) soldered on the PCB in the direction of probing can block the probe and prevent it from landing. At least 3 cm of clearance should be provided on the PCB for the probe, as shown in Figure 6.5. Shorter surface mount devices (SMDs) like capacitors or resistor can be easily placed as needed.



Figure 6.5 PCB clearance for probing

- Low-dropout regulators (LDO) are used on the PCB to suppress the supply noise. As the LDOs draw power, in order to get a true measure of the actual circuit power, a bypass path for LDO is required. A provision to bypass the LDO is shown in Figure 6.6. Here the H1 header provides VDD IN, the input supply for the LDO, and VDD, the supply used during bypass. The supply to the chip, VDD CHIP, can be shorted either to bypass supply, VDD, or the regulated output from the LDO, VDD LDO.
- In order to suppress supply noises, capacitors are placed very close to the package. These capacitors act as storage banks and help in preventing short surges. However, SMD capacitors come with inherent lead inductance. With the increase in capacitance, lead inductance also increases. Beyond the self-resonance frequency of the capacitor, it behaves as inductor rather than a capacitor. To provide a wide range of capacitor decoupling, a set

of four capacitors are used in parallel, 0.1 μ F, 1 μ F, 10 μ F and 100 μ F. This ensures better noise suppression from the supply for a larger frequency range.



Figure 6.6 LDO Bypass

6.4 Optical probe station

All measurements for the CMOS APD are done using an 850 nm VCSEL as a light source. Since VCSELs have a light-emitting cavity of around 20 µm and are multi-mode, MMFs are used for coupling. Initial set-ups used for optical light coupling are discussed in detail in Appendix B.

For the final setup, we used a lensed fiber to pick up light from the VIS wirebonded VCSEL. As the VCSEL had a large divergence angle of 15°, it was not possible to efficiently couple the light directly to the lensed optical fiber. Instead, we used a two lens system to pick up the optical light. The first lens collimates the VCSEL light, while the second one focus the light onto to optical fiber. The new set-up has several controls for alignment. The tilt of the VCSEL and all of the



Figure 6.7 VCSEL alignment

lenses can be controlled, along with X, Y and Z movements, as shown in Figure 6.7. With this setup, we could achieve around 50% of coupling efficiency. The final probe station is shown in Figure 6.12.

The VCSEL gain, S_{21} , was measured using a Lab Buddy and an Agilent E836A VNA. The measurement set-up is shown in Figure 6.8. The measured S21 plot is shown in Figure 6.9.

The eye diagram of the VCSEL was measured using an Anritsu MP1800A BERT, a Lab Buddy



Figure 6.8 S21 measurement set-up for the VCSEL

and an Agilent DSAX93204A oscilloscope as shown in Figure 6.10. The eye diagram of the VCSEL at 6.5 Gbps is shown in Figure 6.11. Although the VCSEL is rated to operate at 25 Gbps, our measurements could not reach such high data rates. We suspect that this is due to the Lab Buddy, which has poor responsivity at 850 nm.



Figure 6.9 Frequency Response of the VCSEL



Figure 6.10 Measurement set-up for VCSEL



Figure 6.11 VCSEL diagram at 6.5 Gbps measured using Lab Buddy



Figure 6.12 Optical probe station

Chapter 7: Conclusion and future work

APD-based RXs greatly relax the sensitivity requirements of the optical RX because of the inherent avalanche gain of the APD. However, due to high reverse bias requirement and temperature sensitivity of the APD, APD-based RXs have been traditionally implemented as multidie solutions. This thesis proposes the first fully on-chip solution of a CMOS based optical RX with APD. Fully-monolithic implementation further improves the bandwidth at the input of the RX by eliminating package parasitics. A simple solution for on-chip biasing and bias stabilization of APD is presented.

APDs fabricated in 130 nm and 65 nm technology compare favorably to the prior-art. The electrical performance of the designed RX is demonstrated up to 8 Gbps. The voltage booster is measured to provide up to 7.42 V of output voltage with a nominal supply of 2.5 V in a standard CMOS technology.

None of the Verilog synthesized blocks worked well in the prototype. This is attributed to errors in the mixed-signal flow, but lack of access to the ARM IP blocks and digital standard cell library prevented us from a proper debug. The blocks that failed to function properly include the ADC and the control loop as it could not be tested without a working ADC. It is recommended that a custom-drawn standard cell library be used for any future mixed-signal design.

Several methods discussed in this thesis were tried out before arriving at the final efficient lens based coupling for VCSELs. However, lack of a fully-characterized optical source/modulator at 76

850 nm wavelength hindered the full measurements of the proposed optical receiver. Due to limitations of the Lab Buddy at 850 nm, the measurements of APDs are pessimistic. We were not able to accurately characterize the losses in the optical path or measure the BW of the VCSEL. These limitations also prevented us in measuring the small-signal BW of the APDs, and only the large signal BW could be measured.

As a next step, we should fully characterize the VCSEL and the path loss to decouple losses and get a more accurate measurement of the on-chip APDs gain and bandwidth.

As an interesting future work, a receiver with a one-tap decision feedback equalization (DFE) and infinite impulse response (IIR) filter feedback can be implemented to ameliorate the maximum achievable data rate with the limited gain-bandwidth of the CMOS APD.

Bibliography

- [1] D. Mahgerefteh *et al.*, "Techno-economic comparison of silicon photonics and multimode VCSELs," *IEEE Journal of Lightwave Technology*, vol. 34, no. 2, pp. 233-242, 2016.
- [2] E. Sackinger, "Broadband circuits for optical fiber communication," Hoboken, NJ, USA: Wiley, 2005.
- [3] A. H. Ahmed *et al.*, "Silicon-photonics microring links for datacenters challenges and opportunities," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 22, no. 6, pp. 194-203, 2016.
- [4] M. J. Lee *et al.*, "A silicon avalanche photodetector fabricated with standard CMOS technology with over 1 THz gain-bandwidth product," *Optics Express*, 18(23):24189-94, 2010.
- [5] Y. Kang *et al.*, "Monolithic germanium/silicon avalanche photodiodes with 340 GHz gain–bandwidth product," *Nature Photonics*, 3, 59 63, 2009.
- [6] N. Duan *et al.*, "310 GHz gain-bandwidth product Ge/Si avalanche photodetector for
 1550 nm light detection," *Optics Express*, 20(10):11031-6, 2012.
- [7] J. Choi *et al.*, "A monolithic GaAs receiver for optical interconnect systems," *IEEE Journal of Solid-State Circuits*, vol. 29, no. 3, pp. 328-331, March, 1994.

- [8] C. Takano *et al.*, "Monolithic integration of 5-Gb/s optical receiver block for short distance communication," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 10, pp. 1431-1433, Oct, 1992.
- [9] J. H. Jang *et al.*, "Long-wavelength In_{0.53}Ga_{0.47}As metamorphic p-i-n photodiodes on GaAs substrates," *IEEE Photonics Technology Letters*, vol. 37, no. 11, pp. 707-708, 2001.
- [10] J. E. Bowers *et al.*, "High-gain high-sensitivity resonant Ge/Si APD photodetectors," *Proceedings of SPIE., Infrared Technology and Applications XXXVI*, 2010.
- [11] J. Wang *et al.*, "Ge-photodetectors for Si-based optoelectronic integration," Sensors, 696-718, 2011.
- [12] F. Tavernier *et al.*, "High-speed optical receivers with integrated photodiode in 130 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 10, pp. 2856-2867, October, 2009.
- [13] T. S. C. Kao *et al.*, "A 5-Gbit/s CMOS optical receiver with integrated spatially modulated light detector and equalization," *IEEE Transaction on Circuits and Systems*, vol. 57, no. 11, pp. 2844-2857, Nov., 2010.
- [14] G. P. Agarwal, "Fiber-optic communications systems," NY, John Wiley & Sons, 2002.
- [15] Hamamatsu Photonics, "Characteristics and use of Si APD," 2004.

- [16] W. S. Zaoui *et al.*, "Frequency response and bandwidth enhancement in Ge/Si avalanche photodiodes with over 840GHz gain-bandwidth-product," *Optics Express*, 2009.
- [17] T. Ikagawa *et al.*, "Performance of large-area avalanche photodiode for low-energy Xrays and gamma-rays scintillation detection," *Nuclear Instruments and Methods in Physics Research*, 10.2172/826645, 2003.
- [18] H. T. Chen *et al.*, "High sensitivity 10Gb/s Si photonic receiver based on a low-voltage waveguide-coupled Ge avalanche photodetector," *Optics Express*, 2015.
- [19] G. Kurczveil et al., "A compact, high-speed, highly efficient hybrid silicon photodetector," *IEEE Optical Interconnects Conference*, pp. 98-99, 2016.
- [20] M. A. Perez Garcia et al., "Low-cost temperature stabilization in APD photo sensors by means a high frequency switching DC/T converter," *IEEE Instrumentation and Measurement Technology Conference*, vol. 2, pp. 1733-1737, 2002.
- [21] J. Kataoka *et al.*, "An active gain-control system for avalanche photo-diodes under moderate temperature variations," *Nuclear Instruments and Methods in Physics Research*, vol. 564, pp. 300-307, 2006.
- [22] N. Zhang *et al.*, "Temperature compensation schemes for APD detectors in PET," *Nuclear Science Symposium Conference Record*, pp. 2995-2996, 2011.

- [23] Lin Tian *et al.*, "Bias voltage compensating circuit," Google Patents CN103940507 A, 2014.
- [24] W. Wang, "Dynamic control of photodiode bias voltage," US Patent 7103288 B2, 2006.
- [25] S. Deng, "Control circuits for avalanche photodiodes," PhD Thesis, University College Cork, 2013.
- [26] D. O'Connell, "Miniature gain and bias control circuit for avalanche photodiodes," *Electronic Letters*, vol. 43, no. 5, pp. 67-68, March, 2007.
- [27] M. J. Lee *et al.*, "Effects of guard-ring structures on the performance of silicon avalanche photodetectors fabricated with standard CMOS technology," *IEEE Electron Device Letters*, vol. 33, no. 1, pp. 80-82, Jan., 2012.
- [28] F. Y. Liu *et al.*, "10-Gbps, 5.3-mW optical transmitter and receiver circuits in 40-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 9, pp. 2049-2067, Sept., 2012.
- [29] J. S. Youn *et al.*, "High-speed CMOS integrated optical receiver with an avalanche photodetector," *IEEE Photonics Technology Letter*, vol. 21, no. 20, pp. 1553-1555, Oct., 2009.
- [30] F. Bruccoleri *et al.*, "Wide-band CMOS low-noise amplifier exploiting thermal noise canceling," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 2, pp. 275-282, Feb., 2004.

- [31] E. Sackinger *et al.*, "A 3-GHz 32-dB CMOS limiting amplifier for SONET OC-48 receivers," *IEEE Journal of Solid-State Circuits*, pp. 158-159, 2000.
- [32] J. F. Dickson, "On-chip high-voltage generation in MNOS integrated circuits using improved voltage multiplier technique," *IEEE Journal of Solid-State Circuits*, vol. 11, no. 3, pp. 374-378, 1976.
- [33] Z. Hameed *et al.*, "Fully-integrated passive threshold-compensated PMOS rectifier for RF energy harvesting," *IEEE International Midwest Symposium on Circuits and Systems*, pp. 129-132, 2013.
- [34] T. Uemura *et al.*, "Preventing single event latchup with deep P-well on P-substrate," *IEEE International Reliability Physics Symposium*, pp. SE.3.1-SE.3.4, 2014.
- [35] H. W. Tsai *et al.*, "Compensation circuit with additional junction sensor to enhance latchup immunity for CMOS integrated circuits," *European Conference on Circuit Theory and Design*, pp. 1-4, 2015.
- [36] Z. Huang *et al.*, "Low-voltage Si-Ge Avalanche photodiode," *IEEE 12th International Conference on Group IV Photonics*, pp. 41-42, 2015.
- [37] Y. Kang *et al.*, "Avalanche photodiode with low breakdown voltage," US Patent, WO2013101110 A1, 2011.

- [38] H. Y. Jung *et al.*, "A high-speed CMOS integrated optical receiver with an under-damped TIA," *IEEE Photonics Conference*, pp. 347-350, 2015.
- [39] K. Iiyama *et al.*, "Hole-Injection-Type and Electron-Injection-Type Silicon Avalanche Photodiodes Fabricated by Standard 0.18-µm CMOS Process," *IEEE Photonics Technology Letters*, vol. 22, no. 12, pp. 932-934, 2010.
- [40] N. J. D. Martinez *et al.*, "Characterization of high performance waveguide-coupled linear mode avalanche photodiodes," *IEEE Optical Interconnects Conference*, pp. 100-101, 2016.
- [41] J. J. Ackert *et al.*, "10 Gb/s bit error free performance of a monolithic silicon avalanche waveguide-integrated photodetector," *Optical Fiber Communications Conference and Exhibition*, pp. 1-3, 2014.

Appendices

Appendix A Latchup

Latchup is a phenomenon where a low-impedance path is created between the power and ground rails due to interaction of pnp and npn parasitic transistors. This low-impedance path between the power and the ground leads to virtual shorting of VDD and ground resulting in a high current that could permanently impair the circuit. Latch-up in CMOS technology is the result of closely placed PMOS and NMOS transistors forming a pnpn junction. The pnpn junction is generally referred to as a thyristor or a silicon-controlled rectifier (SCR). Figure A.1 shows a cross section of a PMOS-NMOS pair placed besides each other. Such an arrangement leads to a formation of coupled bipolar transistors Q1 (pnp) and Q2 (npn). R_{SUBV} and R_{NWV} are the vertical resistances of substrate and n-well, and are proportional to the p+ and n+ contact area, respectively. R_{SUBL} and R_{NWL} are the



Figure A.1 Cross section of a PMOS next to an NMOS, leading to potential latchup

lateral resistance of substrate and n-well, respectively, and depend mainly on the distance of well contact from the source/drain diffusion.

A simplified equivalent circuit for latchup is shown in Figure A.2, where R_{SUB} and R_{NW} are total substrate and n-well resistance. If there is a supply bounce, transistor Q1 turns on. If the R_{SUB} resistor is large, small amount of pnp collector current turns on Q2. Similarly, if there is a ground bounce, npn transistor turns on which in turn triggers the Q1 transistor. The positive feedback increases the current, resulting in circuit failure. If the gain of the Q1 and Q2, β 1 and β 2,



Figure A.2 Latchup equivalent circuit

respectively, is higher than unity, then the transistors will continue to conduct even if the perturbation subsides.

Some of the causes for latchup include disturbance in supply and ground during start-up, voltage spikes in supply and ground rails, input or output signals swings above supply or below ground, and ESD injection of minority carriers from the power clamp into either the substrate or the n-well.

There are many design techniques that can be employed to provide certain amount of protection against latch-up. The current gain of the bipolar transistors can be reduced. With reduced gain (< 1), the feedback action cannot be sustained for a long time and thus the short circuit current can be minimized, preventing permanent circuit damage. The gain can be reduced by lowering the lifetime of the minority charge carriers by gold doping of the substrate. However, it is not under the control of designers and depends on the fabrication company. Another way to reduce the current gain is to increase the spacing between NMOS and PMOS transistors. Having an isolation trench between them further reduces the gain.

 R_{SUB} and R_{NW} are inversely proportional to the amount of current required to trigger Q2 and Q1 transistors, respectively. Minimizing R_{SUB} and R_{NW} helps in controlling latchup. The sheet resistance of the layers, and hence the vertical resistance of substrate and n-well (R_{SUBV} and R_{NWV}), are fixed by the process technology. The lateral resistance can be minimized by placing substrate contact (p+ contact) and n-well contact (n+ contact) close to n+ and p+ junction.

[34] proposes using deep P-well on P-substrate to prevent latch-up. Deep P-well decreases R_{SUB} suppressing variations in the well. Furthermore, the gain of Q2 is reduced as a result of increased base doping concentration. In [35], during latchup event, the ESD protection devices are modified to generate the compensation current, which can reduce the latchup trigger current that flows into the internal circuits.

Certain design rules and good practices must be employed to reduce the substrate and n-well resistances. Minimum n-well and p-well (substrate) contact area can be imposed to minimize vertical resistance. The lateral resistances are determined by the distance between the diffusion junction and their well contacts. Maximum n-well and p-well tap spacing can be imposed to minimize lateral resistance. To further prevent external latch-up, n+ and p+ junctions connected to I/O pads should have extra protection. N-well and p-well guard rings must be used to collect minority charges P-well (substrate) guard ring is used for p+ devices connected to I/O pads. N-well and p-well are used for n+ devices connected to I/O pads. The guard rings must have a sufficient number of vias and connect to metal wiring to maintain very low wiring resistance.
Appendix B Optical Set-up

B.1 Set-up 1

For the first set up measurements, 10 Gbps packaged VCSELs from Lumentum, shown in Figure B.1, are used. These VCSELs come with TOSA packaging of 5 leads for electrical biasing/modulations, output of monitoring PD and a pig tail LC connector for coupling the light into the fiber. The other end of the optical fiber is cleaved and stuck to an old RF probe, to focus the light onto the on-chip APD. The VCSEL is tested using the Discovery Semiconductor Lab Buddy to characterize it. Lab Buddy is an optical front-end receiver with a wide range of wavelength for its PD and TIA. However, our Lab Buddy is mainly characterized at 1550 nm and its responsivity is low at 850 nm of wavelength, requiring high power out of the laser.

Our measurements showed that the VCSEL had a limited BW and could only be reliably operated up to 4 Gbps, as shown in Figure B.2. Also, the probing technique proved to be prone to errors. The glue used to stick the fiber onto the probe made it brittle and damaged the fiber repeatedly.



Figure B.1 TOSA VCSEL



Figure B.2 Measured eye diagram for the Lumentum VCSEL



We renovated an old probe station to build an optical set-up. The old Signaton probe station was dismantled and rebuilt on an optical table. Some of the components were mounted using adaptors and in some places plastic modules were fabricated to minimize the cost of procuring very flat metal adaptors. The Signaton probe station had a metallic platen for magnetic positioners. Holes were drilled to the metallic platen to accommodate both magnetic positioners and Cascade positioners. Two rows of holes were added to provide the flexibility during alignment. As the distance between the Cascade positioners and the chuck was large and could not be reached by the probe head, adaptors were fabricated to facilitate probe landing, as shown in Figure B.3. For optical probing, a custom-made plastic optical fiber holder was fabricated and fixed on the positioners, as shown in Figure B.4.



Figure B.3 Probe positioner adaptor



Figure B.4 Optical probe holder

25 Gbps bare die VCSEL from VIS were wire bonded on a PCB test fixture for measurement purposes, as shown in Figure B.5. For initial aligning, we used plug-in collimators from WT&T. The VCSEL was mounted on the optical table with X control, while the fiber and the plug-in collimator were fixed on a mount which had X, Y and Z control, as shown in Figure B.6. There was no control over the tilt of fiber/VCSEL or the alignment of the collimator with respect to the optical fiber. This severely limited the coupling efficiency (< 0.1%).



Figure B.5 Wirebonded VIS VCSEL



Figure B.6 Initial VCSEL alignment