Design, Analysis, and Performance of a Silicon Photonic Traveling Wave Mach-Zehnder Modulator

David Patel

Department of Electrical & Computer Engineering
McGill University
Montréal, Canada

December 2014

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

Abstract

Silicon photonics is a technology platform for integrating passive and active optical devices. The ability to use existing CMOS-VLSI foundry processes to fabricate integrated optics enables large volume production at low costs while maintaining compatibility with electronic integrated circuits. With the recent explosion in network traffic, large datacenters are demanding higher and higher intra-datacenter transmission speeds over distances of up to 20 km. For this application, optical interconnects are more energy efficient and are able to satisfy the bandwidth demands of the foreseeable future. Optical modulators are an essential component of these communication links and immense research is dedicated to developing efficient high bitrate devices. It is desirable to have high bitrate transmission over a single channel because it simplifies multiplexing (e.g., fewer wavelengths and parallel fibers) which can increase costs. The work presented in this thesis uses the silicon photonic platform to realize an integrated Mach-Zehnder modulator (MZM) suitable for datacenter communication links.

A series push-pull traveling wave MZM is designed using T-shaped extensions to increase the microwave index. The modulator is designed to have a characteristic impedance close to 50 Ω for matching with common microwave drivers and terminations. In addition, the group velocities of the microwave and optical wave are closely matched to optimize the electro-optic bandwidth. A 37 GHz -3-dB electro-optic bandwidth was measured at 1 V reverse bias. The DC V_π was measured to be 7 V in the most efficient arm, corresponding to a $V_\pi L_\pi$ of 2.8 V-cm. For on-off keying modulation, open eye diagrams are visible up to 60 Gbps. Error-free operation with a bit error rate (BER) $<1\times10^{-12}$ was measured up to 45 Gbps, the limit of our photo-receiver. Eye diagrams with a clear opening were observed up to 70 Gbps using analog signal processing to compensate for the limitation of our bit pattern generator. The device was further characterized using digital-signal-processing to dispense the analog bandwidth limitations of the entire communication link to the digital domain. Using PAM-4 and PAM-8 modulation formats, 112 Gbps transmission on 2 km (5 km with PAM-8) of SMF fiber is achieved below the pre-forward-error-correction (FEC) BER threshold of 3.8×10^{-3}, corresponding to a post-FEC BER of $<1\times10^{-15}$. The best case energy consumed by the modulator is estimated to be 0.18 pJ/bit. The ability to generate 100 Gbps modulation in a single wavelength, as achieved in this work, is one of the main waypoints in realizing 400 Gbps and 1 Tbps short reach interconnects.

Sommaire

La photonique sur silicium est une plate-forme technologique pour les dispositifs optiques intégrés passifs et actifs. La possibilité d'utiliser les procédés de fonderie CMOS-VLSI existants pour la fabrication de composantes d'optique intégrée permet une production à grande échelle et à moindre coût tout en maintenant la compatibilité des composantes avec les circuits électroniques intégrés. Avec les récentes explosions du trafic dans les réseaux, les grands centres de données exigent une vitesse de transmission interne de plus en plus grande sur des distances allant jusqu'à 20 km. Pour cette application, les interconnexions optiques sont plus efficace énergiquement et sont capables de satisfaire les demandes de bande passante à venir. Les modulateurs optiques sont des composantes essentielles de ces liens de communication et une grande partie de l'effort de recherche est dédiée à développer des dispositifs efficaces à haut débit. Il est souhaitable d'avoir des débits de transmission élevés sur un seul canal car cela simplifie le multiplexage (i.e. moins de canaux et de fibres en parallèle) qui pourrait augmenter les coûts. Le travail de recherche présenté dans cette thèse démontre l'utilisation de la photonique sur silicium pour réaliser un modulateur Mach-Zehnder (MZM) adéquat pour les liens de communication des centres de données.

Un modulateur Mach-Zehnder à ondes progressives, en série et en montage symétrique est conçu avec une extension en forme de T pour augmenter l'indice des micro-ondes. Le modulateur est conçu pour obtenir une impédance caractéristique de près de 50 Ω pour s'adapter aux impédances typiques de terminaisons et de pilotes à micro-ondes. De plus, les vitesses de groupe des ondes optiques et micro-ondes sont adaptés autant que possible afin d'optimiser la bande passante électro-optique. Une bande passante électro-optique de 37 GHz à -3 dB a été mesurée avec une tension à polarisation inverse de 1 V. La tension de polarisation V_π en courant continu mesurée était de 7 V, correspondant à un facteur de mérite $V_\pi L_\pi$ de 2.8 V-cm. Pour la modulation par tout ou rien (OOK), des diagrammes de l'oeil ouverts sont visibles jusqu'à 60 Gbps. Un fonctionnement exempt d'erreur avec un taux d'erreur sur les bits (BER) de moins de 1×10^{-12} a été mesuré jusqu'à 45 Gbps, la limite de notre photorécepteur. Des diagrammes de l'oeil avec une ouverture distincte ont été observés jusqu'à 70 Gbps en utilisant le traitement de signal analogue. Le dispositif a été analysé plus amplement en utilisant le traitement de signal digital afin de ne pas être limité par la bande passante analogue du lien de communication. En utilisant les formats de modulation PAM-4 et PAM-8, une transmission de 112 Gbps sur 2 km de fibre optique unimodale (SMF) a été réalisée sous le seuil de BER avec pré-correction d'erreur sans voie de retour (FEC) de $3,8\times10^{-3}$, correspondant à un post-FEC BER de moins de 1×10^{-15}. La consommation d'énergie du modulateur dans le meilleur des cas est estimée à 0,18 pJ/bit. La possibilité de produire une modulation à 100 Gbps avec une seule longueur d'onde, tel qu'accompli dans le travail présenté, est une des étapes importantes dans la réalisation d'interconnexion à courte portée à 400 Gbps et à 1 Tbps.

Acknowledgments

I sincerely thank my supervisor, Prof. David V. Plant, for his constant encouragement in pursuing graduate studies. I am very grateful for the long discussions I had with him on various topics over the past few years, which he managed to organize despite his busy schedule. Moreover, I highly appreciate Prof. Plant's dedication to his work which has resulted in a laboratory with high-end equipment that is necessary to perform not only my research but also that of other students in the Photonic Systems Group at McGill.

I would also like to thank the current and past students in our group for their tremendous help and guidance: Venkatakrishnan Veerasubramanian, Samir Ghosh, Alireza Samani, Mathieu Chagnon, Mohamed Osman, Wei Shi, Imran Cheema, Zhaobing Tian, and Qiuhang Zhong. The experimental setup used for all silicon photonic related activity would be nonexistent without Samir and the contributions from Venkat, Ali, and Imran. I am particularly indebted to Samir, who had spent nearly as much time as I have in the laboratory to help me test my devices. Mathieu and Mohamed have contributed immensely to the DSP portion of this work. I acknowledge Wei Shi for initiating us to the topic of traveling wave Mach-Zehnder modulators. I thank Zhaobing Tian for introducing me to the workings of the PSG. I would like to reiterate my appreciation of Venkat and Samir for their advice on all issues in and out of the laboratory. I wish you all the best.

Furthermore, I would like to acknowledge Prof. Lukas Chrostowski from the University of British Columbia and NSERC for the CREATE Si-EPIC program. This program has taught me and numerous students nationwide the skills necessary to perform the design and layout of silicon photonic devices. In addition, I would like to thank the Canadian Microelectronics Corporation (CMC), in particular Dan Deptuck, who has made obtaining fabricated silicon photonic devices a reality. CMC also provided access to a wide suite of simulation and CAD tools that were indispensable for this work.

Moreover, I would like to thank our industry collaborators: Michel Poulin from TeraXion for his invaluable technical insight on modulators and Stephane Lessard from Ericsson Montreal for helping us obtain the DAC used in the experiment.

I dedicate this thesis to my parents and my brother. I would not be where I am if it was not for their hard work and perseverance. I am privileged of their unconditional support in my endeavors.

David Patel

Contents

List of Figures

List of Tables

List of Acronyms

ASP	Analog-signal-processing
BER	Bit Error Rate
BERT	Bit Error Rate (Ratio) Tester
BW	Bandwidth
C-band	Conventional Band (1530 - 1565 nm)
CDR	Clock and Data Recovery
CPS	Coplanar Strips
DCA	Digital Communications Analyzer
DSP	Digital-signal-processing
EDFA	Erbium-doped fiber amplifier
E-O	Electro-optic
FEC	Forward Error Correction
ISI	Intersymbol Interference
LCA	Lightwave Component Analyzer
MIS	Metal Insulator Semiconductor
MZI	Mach-Zehnder Interferometer
MZM	Mach-Zehnder Modulator
NRZ	Non-return-to-zero
O-band	Original Band (1260 - 1360 nm)
OCT	On-chip Termination
OOK	On-off Keying
PAM	Pulse-amplitude Modulation
PD	Photo-diode
PR	Photo-receiver

PRBS	Pseudorandom Binary Sequence
Q	Quality Factor
RF	Radio Frequency
RRC	Root-raised-cosine
SNR	Signal-to-noise Ratio
SOI	Silicon-on-Insulator
SPP	Series Push-Pull
SW-TWMZM	Slow-wave Traveling-wave Mach-Zehnder Modulator
TIA	Transimpedance Amplifier
TMM	Transfer Matrix Model
TPS	Thermal Phase Shifter
TWMZM	Traveling-wave Mach-Zehnder Modulator
VNA	Vector Network Analyzer
WDM	Wavelength Division Multiplexing

> “Pick a flower on Earth and you move the farthest star.”
>
> *P. A. M. Dirac*

Chapter 1

Introduction

The past decade has been marked by a significant increase in internet access media and an abundance of cloud-based storage, high-definition multimedia streaming, and services providing software, platforms, and infrastructures (SaaS, PaaS, and IaaS). This has resulted in an unprecedented demand in the speed and volume of data transfers. As forecasted by Cisco (Fig. 1.1), this demand will definitely continue to grow. By the year 2018, 8.6 zettabytes of data will be transferred per year [1]. What is astounding is that 75% of this network traffic is within the datacenter itself (east-west traffic pattern). On top of that, the internet of things or everything (IoT/IoE) is still nascent. It can be speculated that as IoT develops, it will further exacerbate network congestion within datacenters. Moreover, the massive volume of data has resulted in the creation of large ware-house scale datacenters. For such long and fast links, optical interconnects are believed to be superior to other forms of interconnects and are now becoming inevitable.

At the same time, the forefront and the future exascale supercomputers also require high bandwidth interconnects for transferring data from one central-processing-unit (CPU) to another, from the CPUs to the memory, and between different nodes in a cluster. Supercomputers consume large amounts of power, typically in the range of several megawatts, resulting in yearly energy costs of millions of dollars. Most of this power is consumed by interconnects. In fact, it is estimated that transferring data to the nearest chip can cost 200 times the energy needed to perform operations on that same data [2]. Thus, these interconnects must not only support high data rates but must also do this with very low energy consumption. In addition to supercomputers, better interconnects have also become

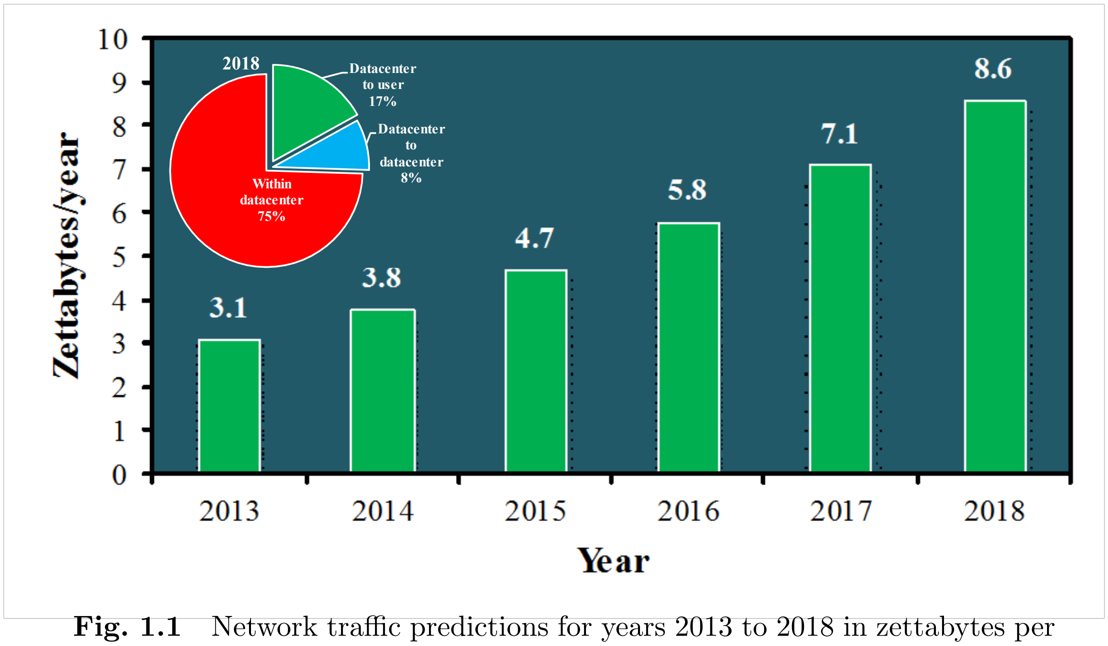


Fig. 1.1 Network traffic predictions for years 2013 to 2018 in zettabytes per year. The pie chart (inset) shows the predicted distribution of traffic in 2018. Figure produced from data in [1].

necessary at a much smaller scale. The transit frequency f_T of transistors has reached hundreds of GHz with the recent CMOS technology nodes. Unfortunately, processor operating speed has remained significantly lower due to bandwidth limitation of the interconnect and the power consumption associated with them. High-speed on-chip clock distribution is also less reliable for higher frequencies with present electrical interconnects and is an additional obstacle to faster processor speeds.

Increased bandwidth demands at various scales has motivated tremendous research in the field of interconnects. Interconnects can be divided into two main categories: wireless and wireline. Wireless communication for chip-to-chip (inter-chip) and within the chip (intra-chip) is believed to be more convenient as complicated on-chip and on-board routing is avoided. It is also believed to be power efficient. However, the distance and the bandwidth is relatively limited. Wireless communication is also highly susceptible to interference, and protocols allowing only few devices to communicate at a time, or schemes using division multiple access need to be implemented [3]. The most common wireline interconnects are electrical and optical. Electrical interconnects have a long delay and are also very lossy. Therefore, they are only advantageous for low bitrates and short distances. Optical channels, on the other hand, have an extremely large bandwidth. For instance, the conventional wavelength band (C-band: 1530-1565 nm) corresponds to a bandwidth of 4.3 THz. The optical channel can also be considered to be relatively lossless for short links. For example, a typical single-mode fiber (SMF) has a loss of only 0.2 dB/km in the C-band. Optical channels, compared to their electrical counterparts, have also very negligible cross-talk and thus high density interconnects are possible [4]. In fact, integrated optical paths can even cross each-other with negligible interactions among them. Although the optical channel may appear to be ideal, it does exhibit non-linearity and it is also quite dispersive in the C-band. These effects, however, become only an issue for longer distances and can often be compensated with signal processing.

While the optical channel is nearly ideal for short-reach communication, data needs to be generated and processed in the electrical domain. Compared to electrical interconnects, the additional steps of converting electrical signals to optical signals and vice-versa are required to make use of optical interconnects. This is a major hurdle in the adoption of optical interconnects as it increases cost and the total power consumption. Hence, for low bitrates and short distances, i.e., at the scale of intra-chip and inter-chip communications, copper interconnects cost less and are more energy efficient. The entire optical interconnect system would need to consume an energy of about 100 fJ/bit to be competitive [4]. The cost benefit begins to favor optical channels for longer distances and/or higher bitrates. For example, interconnects longer than 2-3 m at a bitrate of 25 Gbps favor optical channels over electrical interconnects in terms of cost. At 40 Gbps, optical interconnects are more cost effective than electrical interconnects for links longer than about 1 cm [2]. In a 16 petaflops supercomputer, optical interconnect would consume about 8.5% of the power consumed by a system with 10GBase-T cabling. This translates to savings of about $150 million dollars in operating costs over the life of the machine (~10 years) [2]. Thus, for the applications of datacenter communication links and supercomputer nodes, optical interconnects are the most suitable choice.

Optical transmit-receive systems are typically implemented in a wavelength division multiplexing scheme (WDM), as illustrated in Fig. 1.2. The terahertz bandwidth of the optical channel is too large to be used adequately, therefore, it is usually divided into channels of smaller bandwidths with different center wavelengths. Typical wavelength spacing are 0.4 nm in dense WDM systems or 20 nm in a coarse WDM systems. Coarse systems can tolerate less accurate lasers as variations in the wavelengths can have less impact on neighboring channels, whereas dense systems allow for more channels within an

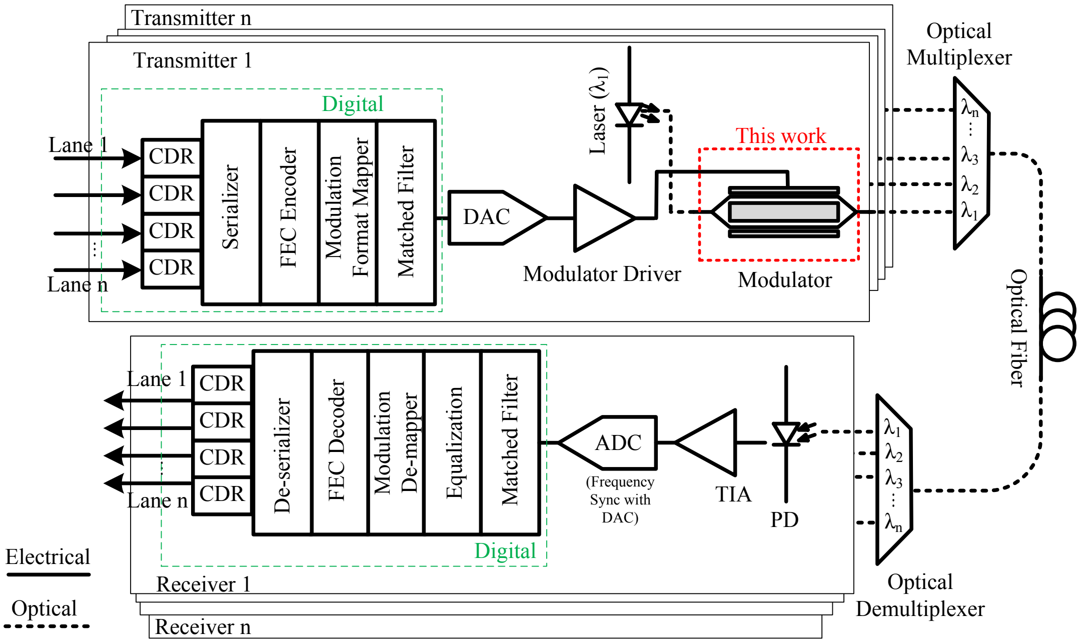


Fig. 1.2 An example of a short-reach optical transmit/receive chain with digital-signal-processing and error correction coding. CDR: Clock and data recovery, FEC: Forward error correction, DAC: Digital-to-analog converter, PD: Photo-diode, TIA: Trans-impedance amplifier, ADC: Analog-to-digital converter.

aggregate bandwidth. The division of the large bandwidth into smaller bands can also reduce noise with bandpass filtering. Moreover, multiplexing allows the relatively slower electronics to make efficient use of the large optical bandwidth. The system in Fig. 1.2 also includes coding and digital-signal-processing (DSP) and would increase the total power consumed by the system. For short-reach communications, coding and DSP could only be justifiable when targeting extremely high datarates with very low bandwidth electronics and electro-optics. The data to be transmitted optically often comes into n electrical lanes operating at a slower bitrate B. A serializer will convert the parallel streams into a single stream at a bitrate of $B \times n$. Forward error correction coding makes it possible to recover data received with a high error count to one that is nearly error-free. This encoded data is then mapped to a desired modulation format, for example, multiple amplitude levels (such as PAM) in intensity modulated systems or quadrature amplitude levels in coherent systems. These modulation formats transmit symbols (multiple bits), instead of single bits within the same time slot and with about the same bandwidth. Thus, higher spectral efficiency, which is defined as the ratio of the occupied bandwidth and the channel's available bandwidth, is achieved. The bit-to-symbol mapping is usually done with a labeling scheme that reduces the Hamming distance and optimizes the signal-to-noise ratio (SNR). After mapping to a modulation format, the data is pulse-shaped to reduce its bandwidth with a filter that is matched at the transmitter and the receiver. Using the same pulse shaping at the transmitter and receiver optimizes the SNR. The output of the digital-to-analog converter is then amplified to voltages required by the optical modulator. A laser provides continuous-wave light at the carrier frequency to be modulated. After propagating through the fiber, at the receiver side, the light is captured with a photo-receiver and digitized with an analog-to-digital converter (ADC). Operations in reverse of the ones at the transmitter are applied in order to recover the transmitted data. Clock and data recovery circuits are used to ensure that the continuous waveform is sampled at the appropriate time. In addition, the ADC and DAC should also sample with minimal frequency offsets between each other, though this can be corrected with additional digital-signal-processing (DSP).

This thesis presents the analysis, design, and testing of one of the sub-components of the optical transmitter: the Mach-Zehnder modulator (MZM) (outlined in red in Fig. 1.2). The MZM is essentially a transducer that converts data from the electrical domain to the optical domain and thus allows the use of the optical channel for data transmission. As such, it is a key component of transceivers. In the experiment discussed in this thesis, the MZM is operated in on-off keying (OOK) modulation and then with PAM modulation using DSP.

1.1 Silicon Photonics

During the same time that bandwidth demands were increasing, significant development has occurred in the field of silicon photonics. Silicon photonic transmitters and receivers have even been commercialized. In this relatively new technology, optical integrated devices and circuits are being built on silicon, the material system employed by electronics for several decades. The silicon photonic fabrication process has matured over the past years, and many companies and institutes now have access to this technology. For instance, IBM has a 90 nm CMOS photonic process with monolithically integrated transistors, i.e., mixed signal (digital and analog) electronic circuits can be created on the same die as the optical circuits.

Luxtera also has a silicon photonic technology based on Freescale's 130 nm with monolithic integration. Other companies have acquired or also developed their own technologies, e.g., Intel, Cisco (acquired Lightwire), ST Microelectronics, Huawei (acquired Caliopa), and Mellanox. Some of these technologies have placed more emphasis on packaging and other forms of integration as oppose to monolithic integration of electronics, for example, hybrid integration techniques for lasers, and 3-D assembly technologies. Moreover, silicon photonic devices have also been built using unmodified IBM 45 nm SOI process with additional post-processing steps [5].

Most of these foundries have remained private, but various organizations worldwide have made multi-project wafer (MPW) runs available for fabrication of silicon photonic devices at foundries such as imec, CEA-LETI, VTT, IHP, and IME A*STAR. This has made possible for academia to obtain silicon photonic integrated circuits similar to microelectronics in the past. In Canada, this has been made possible by CMC Microsystems [6] and the NSERC CREATE Si-EPIC program [7]. The OpSIS-IME was a similar foundry service based in the United States.

The wide interest in silicon photonics arises from its numerous advantages. Since the technology uses existing CMOS-VLSI process steps, it is more reliable and has better yield compared to other integrated photonic technologies. This amounts to silicon photonics having an economy of scale resulting in cheaper integrated photonics for large volume production, such as the volumes expected for datacenter transceivers. Furthermore, it has the potential of monolithic, hybrid, and multi-chip module (system in a package) integration. For datacenter links, monolithic integration with electronics may not be optimal because optical devices are typically on the micrometer scale while electronics are on the nanometer scale. In advanced CMOS technology nodes, die area is expensive and the footprint of optical devices can significantly increase costs [8]. Furthermore, modifications required for monolithic integration with photonics can degrade transistor performance [8].

Another advantage of silicon photonics is its high index contrast ratio between the core and the cladding. This allows for denser optical devices and circuits relative to other technologies.

One of the biggest disadvantage of silicon is that the material has an indirect bandgap, thus it is very inefficient for light emission. Lasers will have to be integrated hybridly meaning additional fabrication steps and optical coupling losses. Furthermore, silicon is almost transparent at wavelengths of interest to telecommunications and photo-detectors cannot be built without the use of an additional alloy such as germanium. The transparency, however, allows for lower propagation loss. In the context of modulators, silicon is relatively inefficient and requires larger driving voltages. Thus far, it has higher overall optical insertion losses, and only recently high bandwidth devices are getting published. A short list of commercially available modulators in competing technologies is presented in Table 1.1.

Table 1.1 Commercially available Mach-Zehnder modulators in other technologies compared with silicon Mach-Zehnder modulator from this work.

Technology	Vendor	Bandwidth [GHz]	Drive Voltage [V]	Insertion Loss [dB]
LiNbO3				
	EO Space	30+	3	<3
	Photline	30+	5	3.5
	Oclaro	30+	5	3.5
GaAs				
	Optelian	50	3.3	4
	Finisar	27	3	10
InP				
	TeraXion	35	1.5	9
Si				
	(This work)	38+[a]	7	15[b]

[a] Bandwidth at 4 V bias is 41 GHz and response at 50 GHz is only -3.2 dB.

[b] Includes fiber coupling loss.

1.2 Silicon Optical Modulators

The technology's potential in building transmitters and receivers is the major motivation of research in this field. In fact, most silicon photonic processes are specifically optimized for modulators and photo-detectors.

In optical transmitters, modulation is generally performed either by directly modulating the laser, i.e., by turning on and off the laser, or by modulating the light out of the laser using an external modulator as shown in Fig. 1.2. Direct modulation can be more compact, simpler, and could also be cost-effective. External modulation, however, has reached higher speeds, has larger modulation depth, and introduces less distortions. Distortions in direct modulation is mostly in the form of chirp, which occurs because of parasitic phase modulation. In external modulation, the chirp can be removed or even reversed for dispersion tolerance. In addition, external modulation can allow other modulation formats such as quadrature phase shift keying.

External modulation in silicon is typically accomplished by a change in the effective index or a change in absorption (electro-absorption modulators). Changes in effective index lead to a change in phase of the optical wave, which can be transformed into changes in intensity using interference, resonance, or slow-light behavior. Common examples of resonance based modulators are micro-ring modulators [9] and Bragg grating modulators [10]. For these devices, the pass band of the filter can be moved with a change in effective index. When the wavelength is fixed, the shift in the passband appears as changes in transmission creating intensity modulation. Other investigated structures aid interferometers achieve better efficiency via slow-light effects using band-stop structures such as photonic crystals [11], gratings [12], and coupling modulated ring assisted Mach-Zehnder interferometers [13]. Slow-light devices require high-resolution lithography and sometimes post-processing steps such as polymer deposition.

In devices fabricated with photo-lithography, micro-ring modulators and Mach-Zehnder modulators have been rigorously researched. Compared to MZMs, micro-ring modulators are compact, have low insertion loss, and are believed to be more energy efficient than traveling wave MZMs. Unfortunately, they are quite sensitive to fabrication and temperature variations. Micro-ring modulators cannot tolerate temperature changes greater than about 0.2°C, whereas asymmetric MZMs can tolerate changes in temperature up to 8.3°C [14]. Thus, micro-ring modulators additionally require heaters and controlling circuits in feedback loops. Nevertheless, even with the extra control circuitry, they are believed to have a relatively low power consumption and are a contender to MZMs in this aspect. In addition, because of their compact size relative to MZMs, they are suitable for high-density intra-chip and inter-chip communication links. Micro-ring modulators were previously limited in bandwidth due to its cavity linewidth, yet modulation at up to 30 Gbps was generally achieved. A recent publication has shown higher modulation speeds of up to 40 Gbps, with a 6.2 dB extinction ratio, a drive voltage of 4.8 Vpp, 1 dB propagation loss, and 0.12 pJ/bit of energy consumption [15]. Other micro-ring modulators have been developed on the coupling modulation effect instead of the cavity limited intra-cavity modulation effect. Such modulators have been shown to achieve data rates twice of that obtainable via intra-cavity modulation [16]. Due to their inherent ability to modulate specific wavelengths on a common bus, micro-ring modulators are convenient for WDM systems. Modulation with micro-rings is typically performed with wavelength biased near the resonance dip, this can lead to a very high modulation loss and poor detection of the transmitted signal.

Mach-Zehnder modulators are typically more energy consuming and have a much larger footprint. They are generally several millimeters long and can be hundreds of micrometers wide. Fortunately, these devices are easier to test, and because they are more resistant to temperature changes, temperature stabilization is not as critical. As such, tremendous research has been performed on silicon MZMs with respect to electrode design for improved bandwidth, and process design for better efficiency and lower optical losses. The work in this thesis focuses on MZMs, which are of immense interest due to their simplicity, robustness, and high bandwidth operation. MZMs are viable for datacenter transmitters. The interest in this field is attested by the vast amount of prior work on this device. A list of notable publications on carrier depletion based MZMs is shown in Table 1.2. A more succinct summary of the state of the art is also presented in [17].

Table 1.2 Recent developments of carrier-depletion based Mach-Zehnder Modulators

Type[1]	Band[2]	Loss[3] [dB]	$V_\pi L_\pi$ [V-cm]	Length[4] [mm]	E-O 3-dB [GHz][5]	Speed[6] [Gbps]	Drive [V_{pp}]	Bias [-V]	ER [dB]	Junction	Transmission Line	Z_T[7]	Energy[8] [pJ/bit]	Ref.	Year
SPP	C	4.2	2.4	2.4		30					CPS			[18]	2014
DD	O	1.93	4	3.5		40	2[f]	2	2.74		CPW	20		[19]	2014
SD/DD	O		1.1	3						Interleaved	CPW (double-pass)			[20]	2014
SD/DD	C	4	0.9	1.6	12 (1V)	25	1.75	0.8	3.1	Lateral	CPW (double-pass)	50		[21]	2014
DD	C	6.2	2.7	3	30[a]	40	1.6[f]	1	3.1	Lateral[d]	CPS	25 (33)	0.64	[22]	2014
SPP	O	5.2	3.9	6	21			4		Lateral	CPS	50		[23, 24]	2014
DD	C	3.34	2.2	0.75	35 (0V)	70	5.32[f]	1.5	6.1	Lateral	CPW	18		[25]	2014
DD	C	7.7	2.73	3.5	27	40	1.6[f]	1	10	Lateral[d]	Shield CPS w/ T-rails	50		[26]	2014
SD	O	1.65	7.65	5.5	30	28	1.3	5	5.66	Lateral[c]	Shield CPW	50	0.8	[27]	2014
SD	C	2.2	5.17	5.5		28	1.3	5	5.97	Lateral[c]	Shield CPW	50	1.1	[27]	2014
SD/DD	C	4.1	2.2	0.95	26	40	7	4	8	Lateral[e]	CPW	50		[28]	2014
SD/DD	C	9	1.85	3	30	50	3.5	5	7	Lateral[c]	CPW (Copper)	50		[29]	2014
SD	C			4	25.6	50	7	5	5.56	Lateral[c]	CPW (SGS)	50		[30]	2013
DD	C	7	2.1	3	30[a]	40	2.5[f]	0.25	5.1	Lateral[d]	CPS	25		[31]	2013
DD	C		0.2	0.625			1[f]			SISCAP	Lumped			[32]	2013
SD/DD	C	8.8	2.4	4	16.2			2						[33]	2013
SD/DD	O	5.5	2.64	3	>30	50	3[f]	0	3.4	Lateral[d]	CPS	25 (33)	0.45	[34]	2013
SD	C	3.5	2.4	0.95	20 (3V)	40	6		7.9	Interleaved	CPW	50		[35]	2013
SD	C	1.9	2.05	0.75	27.8	60	3.6	5	4.4	Lateral[e]	CPW	50 (30)		[36]	2013
SD/DD	C	2.8	1.12	0.5	4.3 (0V)	10	3	2.5	7.5	Interleaved	Lumped			[37]	2012
SD/DD	C	4.6	1.19	3		35		6	11.3	Lateral	CPW	33		[37]	2012

Continued on next page

Type[1]	Band[2]	Loss[3] [dB]	$V_\pi L_\pi$ [V-cm]	Length[4] [mm]	E-O 3-dB [GHz][5]	Speed[6] [Gbps]	Drive [V_{pp}]	Bias [-V]	ER [dB]	Junction	Transmission Line	Z_T[7]	Energy[8] [pJ/bit]	Ref.	Year
SD/DD	C	6.2			15.5 (0V)	25			4.4		Lateral	33		[38]	2012
SD/DD	C	13	1.65	2		12.5	2	0.8	12.79	Lateral	CPW	33	7.8	[39]	2012
SPP	C	4.1	2.4	2	>20	50	4.5	6	4.7	Lateral	CPS	50		[40]	2012
SPP	C	6.6	2.08	4	>20	40	4.5	5	6	Lateral	CPS	50		[40]	2012
SPP	C	9	1.86	6	>20	30	3.5	3	8.5	Lateral	CPS	50		[40]	2012
SD/DD	C	2	1.6	0.75	>20[b]	40	7	3.5	7.01	Interleaved	CPW	50		[41]	2012
SD	C	7.4	2.8	2.04	>40[b]	50	6.5	4	3.1	Lateral[e]	CPW	50	4.2	[42]	2012
DD	C	8.7	1.8	5	10	20	0.63[f]	0	3.7	Lateral	CPS	50 (37)	0.2	[14]	2012
SPP	C		0.7	2	14					Lateral		50		[43]	2012
SD	C	6	3.5	4.7	20	40	7		6.6	Lateral (pipin)	CPW	50		[44]	2012
SD/DD	C	7.7	11	1.35	27[b]	40	6	3	6.5	Wrap-Around	CPW	50		[45]	2011
SD/DD (TM)	C	5.4	14	1.35	27[b]	40	6	3	6.5	Wrap-Around	CPW	50	4.5	[45]	2011
DD	C	15	2.7	3.5		40	4[f]	2	7	Lateral[e]		50	2	[46]	2011
SPP	C	5	2.4	2.4	>30[b]	30	6	3	7.2	Lateral	CPS w/ T-rails	50		[47]	2011
SD	C	6.4	4	1	20	30	6.5	3.5		Vertical	CPW	19		[48]	2008

[1] SD: Single Drive, DD: Dual Drive, SPP: Series Push-pull, SD/DD: Not explicitly indicated if modulator is driven single-ended.

[2] O: O-band, C: C-band.

[3] Loss includes MZM coupler loss.

[4] Not necessarily the phase shifter (or active) length.

[5] At reverse bias voltage unless indicated in brackets.

[6] Maximum speed tested in the reference.

[7] Termination impedance. Transmission line characteristic impedance shown in brackets.

[8] As reported in reference.

[a] Peak of 2.5 dB at 6 GHz, dip near -3 dB at 16 GHz, plateau at -3 dB from 25 to 30 GHz.

[b] Reported value is the RF -6 dB point.

[c] Junction uses doping compensation.

[d] Junction uses 6-levels of doping.

[e] Junction uses an offset doping for p-type dopant.

[f] V_{pp} differential.

1.3 Thesis Objective and Structure

The design of an optical modulator involves an intricate tradespace between optical loss, energy consumption, fabrication capabilities, and modulation speed. The objective of the thesis is to discuss this tradespace, present simulation results, and report experimentally measured performances. The device developed in this thesis was optimized for the 3-dB electro-optic bandwidth within the constraints imposed by the foundry processes. This was achieved by using an electrode geometry, p-n junction structure, and MZM configuration that was only briefly explored in this technology. With this structure, we measured an electro-optic bandwidth that is at least 8 GHz higher than reported in the literature for similar devices. The design of the electrodes used planar microwave engineering techniques as an attempt to match the impedance, match the velocity of the microwave to the velocity of the optical wave, and reduce the microwave propagation loss. As part of a communication link, the modulator is tested by performing bit-error-rate measurements under various conditions such as different bias and driving voltages. These experiments illustrate the trade-offs in the transmitter's performance with respect to energy consumption, extinction ratio, and small-signal bandwidth. Furthermore, by using DSP and PAM modulation, it is shown that the modulator can transmit beyond 100 Gbps on a single channel with error correction. This methodology was previously applied for O-band silicon modulators [24, 23]. In this work, this is accomplished using a C-band modulator and a slow-wave traveling-wave MZM (TWMZM) structure with a bandwidth that is 18 GHz higher. The achievement of 100 Gbps per channel is fundamental for developing the future 400 Gbps Ethernet using four wavelength channels.

Due to the immense research in the field of TWMZMs, this topic has been covered by many dissertations in other technology platforms and recent publications for silicon. Therefore, this thesis shares a structure similar to the theses for GaAs [49], InP [50, 51], and the basics of silicon photonics from the books [52] and [53].

This introductory chapter presented a motivation for performing research in short-reach silicon optical interconnects. A brief comparison was made with the silicon photonic technology and modulators in the $LiNbO_3$, GaAs, and InP technology. The advantages and disadvantages of the silicon technology along with the state of the art was also discussed.

Chapter 2 presents a general description of silicon MZMs and a simple model for simulation of the optical device is shown. Several key modulator metrics and simulation results are described. This is followed by a discussion of modulation mechanisms in the silicon platform. With respect to the carrier depletion effect, several simulated results for various p-n junction geometry are also presented.

Chapter 3 focuses on the RF and MW aspects of a TWMZM modulator design. The chapter begins by demonstrating the necessity of traveling wave designs as opposed to simple lumped element modeling. Then, simulations are presented for co-planar strips transmission lines without p-n junctions. Next, the effect of the p-n junction on the transmission lines are analytically determined and simulations with an electromagnetic finite element model solver are also presented. The electrical and optical solutions are subsequently combined to determine the overall electro-optic performance of the device.

Chapter 4 summarizes elements from Chapter 2 and Chapter 3 with respect to the design that was fabricated and characterized. Additional layout considerations are also discussed in this chapter.

The measured performance and discussion is reported in Chapter 5.

Finally, the thesis concludes in Chapter 6 with a summary of the key points and achievements.

Chapter 2

Silicon Mach-Zehnder Modulators

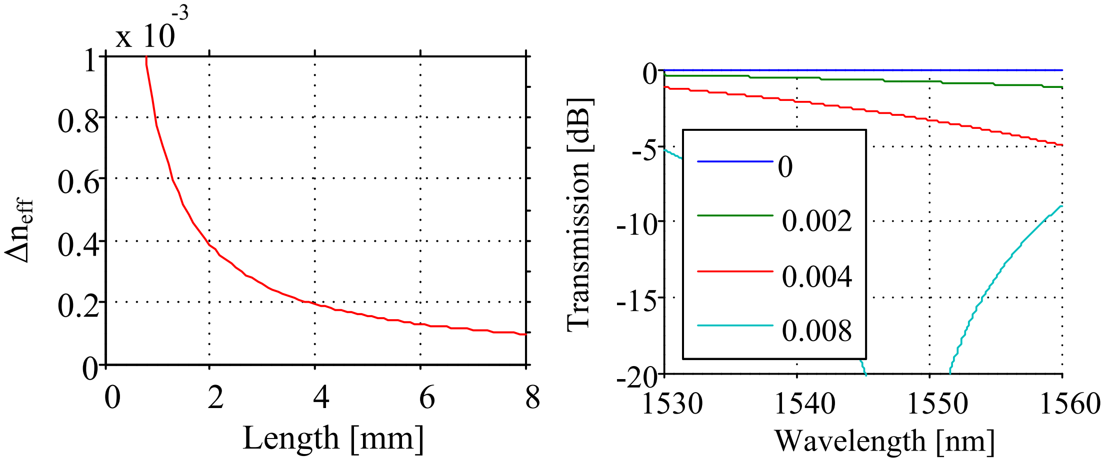
2.1 Introduction

This chapter discusses the fundamentals of Mach-Zehnder modulators (MZM) with emphasis on their implementation in a silicon platform. Section 2.2 presents a brief overview of typical Mach-Zehnder interferometers (MZI). This is followed by a description of the transfer matrix model (TMM) in Section 2.3 for primitive, but fast and informative, simulations of MZIs. Critical design issues from these basic simulations are presented. Afterwards, in Section 2.5, modulation in silicon is examined. Section 2.6 considers different p-n junction design parameters and summarizes trends that were observed via simulations, and in the literature. At last, different modulator driving configurations are compared in Section 2.4.

2.2 Mach-Zehnder Modulator Basics

Generally speaking, a MZI is used to detect relative phase shifts between two optical paths. A MZI is formed by splitting light from a common source, guiding light in two paths that are subjected to phase shifts, and then recombining the light. For electro-optic modulators, the phase shift in one or both optical paths is intentionally introduced by modulating the refractive index with an electric stimulus. For other applications such as sensing, this change in refractive index can be brought upon by a different stimulus, for instance, the analyte itself. As summarized by Eq. 2.1, this change in refractive index, or more specifically the effective refractive index Δn_{eff}, modifies the phase constant β of an optical wave with a free-space wavelength λ_0. This results in a different phase shift $\Delta\phi$ accumulated by light as it travels a distance L than it would if no stimulus was present. When different phase shifts are introduced in the two arms of the interferometer, changes in intensity can be observed at the output of the interferometer because of interference. For example, in an ideal case, a π-phase shift relative to the other arm would lead to destructive interference and all light will be extinct. However, if identical phase shift is experienced by both arms, light will interfere constructively resulting in the maximum possible light out of the interferometer.

$$\Delta\phi = [\beta]L = [\frac{2\pi}{\lambda_0}\Delta n_{eff}]L \tag{2.1}$$



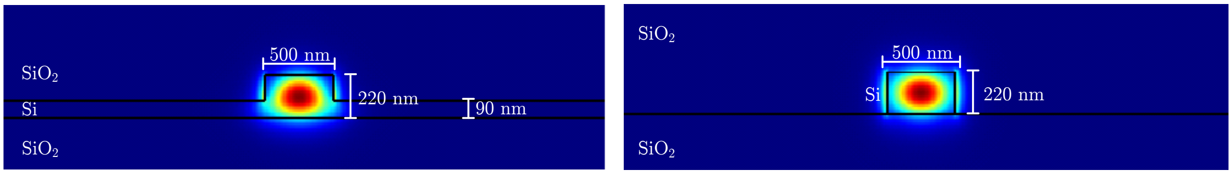
(a) (b)

Fig. 2.1 (a) Required change in effective index for a π phase shift as a function of the phase-shifting length for 1550 nm. (b) Effect on transmission spectrum due to a Δn_{eff} between two arms of a 3 mm long MZI.

In Fig. 2.1(a), the required change in effective index (relative to the other arm of the MZI) from Eq. 2.1 for a π phase shift is plotted for different phase shifting lengths at 1550 nm. This plot demonstrates the trade-off between length and the amount of modulation that is required for complete destructive interference. For a phase shifter of 3 mm, the required change in Δn_{eff} between the two arms of an MZI is about 2.6×10^{-4}. This number is technology independent. In Fig. 2.1(b), the change in transmission calculated with the TMM, to be described in Section 2.3, is shown for changes in effective index at the third decimal place for a 3 mm long MZI. This figure illustrates that for a specific wavelength, even a seemingly minute change in effective index between the arms of a MZI can have a strong impact on the transmission.

2.2.1 Silicon Waveguide Design

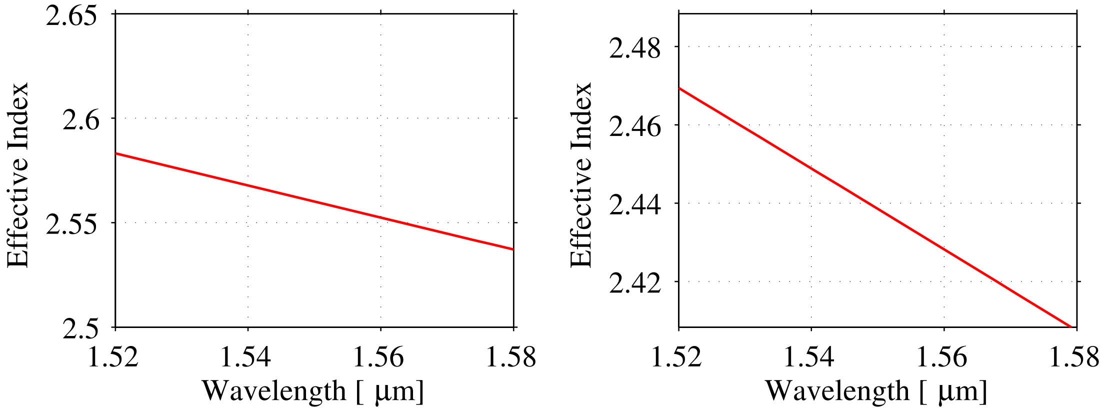
The arms of the interferometer are waveguides which are designed to confine and propagate light. In a SOI technology with a 220-nm thick top silicon, light is predominantly guided by buried channel (strip) or rib waveguides. Buried channel waveguides are created by surrounding a core material that has a higher refractive index with a cladding material which has a lower refractive index for total internal reflection. In SOI, silicon is the core material which has a refractive index of approximately 3.48 near 1550 nm at room temperature, and silicon dioxide is the cladding material with a refractive index of 1.44 near 1550 nm [54].



(a) (b)

Fig. 2.2 Optical TE mode at 1550 nm in silicon (a) rib, and (b) channel waveguide.

Since the strip waveguide is surrounded by the cladding on all sides, it has a high optical confinement and lower cross-talk with other waveguides in the vicinity. However, strip waveguides can have higher scattering loss due to side-wall roughness from manufacturing [55]. Rib waveguides, on the other hand, have a slab on which the strip is placed and they have a lower confinement. The lower confinement, compared to strip waveguides, requires larger separation between waveguides, about 2 µm, for low cross-talk and larger bending radii for lower radiation losses in 90° bends [55]. Rib waveguides, however, have lower optical propagation loss because the optical mode is less intense at the side-walls, therefore it is less affected by the roughness of the side-walls introduced during manufacturing. Typical SOI buried channel waveguides for the TE mode at 1550-nm wavelength are 500-nm wide and 220-nm high. Rib waveguides in this process are formed with a 90 nm slab. These dimensions were determined from a trade-off of single mode operation and propagation loss. The confined TE optical mode, and the outline of each type of waveguide, are shown in Fig. 2.2. The average propagation loss in the technology used to fabricate the modulators in this work (OpSIS-IME) were 2.4 dB/cm for channel waveguides [31] and 1.5 dB/cm for rib waveguides [56].



(a)

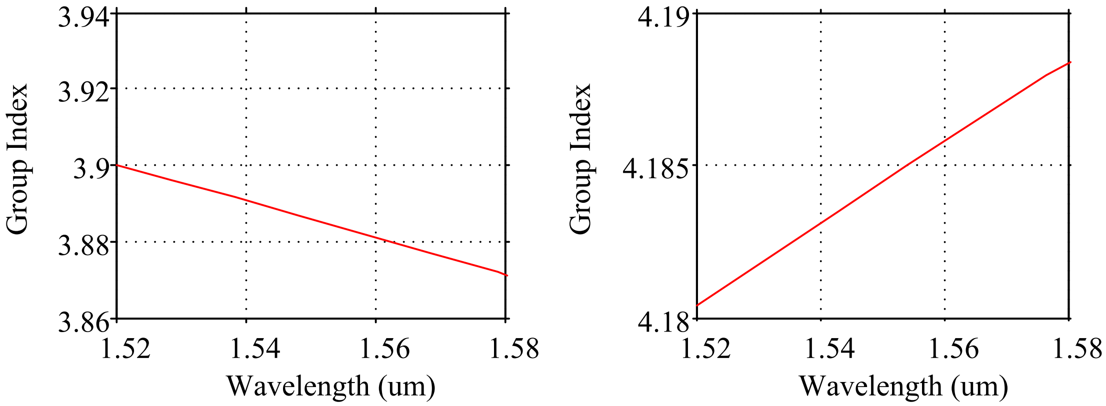
(b)

Fig. 2.3 Simulated effective index in silicon (a) rib, and (b) channel waveguide.

Fig. 2.3 shows the simulated effective index for the fundamental TE mode in the rib and channel waveguides as a function of wavelength using the Lumerical MODE solver and procedure described in [52]. It is observed from Fig. 2.3 that SOI waveguides are dispersive since the effective index varies with the wavelength. Dispersion results in a group index that can be approximated by:

$$n_{og} = n_{eff} - \lambda_0 \frac{dn_{eff}}{d\lambda_0} \tag{2.2}$$

Using the simulated effective index of Fig. 2.3, and Eq. 2.2, the optical group index of the channel and rib waveguides were obtained and are plotted in Fig. 2.4. This group index will be used as the target for the microwave group velocity for the electrode design.



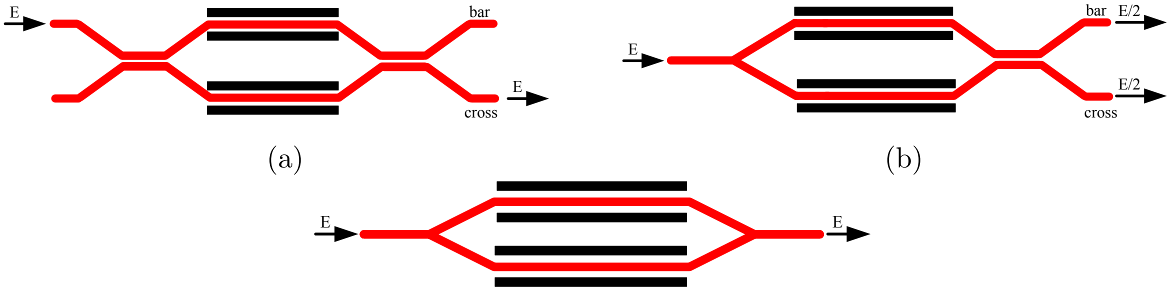
(a) (b)

Fig. 2.4 Simulated group index in silicon (a) rib, and (b) channel waveguide.

2.2.2 MZI Input/Output Coupling

The splitting of the light at the input and the merging of the arms at the output of the interferometer can be performed using y-branches, 3-dB directional couplers, or with multimode interference (MMI) couplers. In SOI, very compact y-branches with low insertion-loss, and large bandwidth have been demonstrated [57]. MMI couplers have lower polarization dependence, are also low-loss, and are more tolerant to fabrication variations [58]. Their footprint, however, is significantly larger than y-branches. Directional couplers formed with channel waveguides are also compact, but their performance degrades substantially with minute variations in fabrication. Better tolerance to fabrication variations has been recently shown with rib directional couplers [59], and long adiabatic directional couplers [60].

A MZI can be further classified as an $n \times m$ MZI by the number of inputs n, and outputs m [53]. This is determined by the choice of the coupling device, as illustrated in Fig. 2.5, at the input and output of the interferometer. In a 2×2 MZM with directional couplers and no additional phase shift in the arms, the output appears at the cross port with a phase shift of π relative to the input.When a π phase shift is applied to the MZI arms, the light is output to the bar port. Such 2×2 MZI devices are the building blocks of interferometric optical switches. The 1×2 MZI, without any phase shift, behaves as a 3-db directional coupler. When a $\pi/2$ phase shift is introduced, the light is sent to one output port only. With a $-\pi/2$ phase shift, the light is output from the other port. Without any phase shift,



(a) (b)

(c)

Fig. 2.5 Different configurations of MZI: (a) 2×2 MZI, (b) 1×2 MZI, and (c) 1×1 MZI.

light is output from both ports. For modulators, all three configurations are applicable, however, the 1×1 MZI is the simplest.

2.2.3 Unbalanced and Asymmetrical MZIs

A MZI is symmetrical if both arms have the same optical path length, i.e., physical path length × refractive index, and asymmetrical if they have different optical path lengths [53]. Moreover, a MZI is balanced if the physical path length is identical in both arms, otherwise it is termed unbalanced. An imbalance introduces additional phase shift in one arm and results in a periodic transfer function with respect to wavelength. The period is defined as the free spectral range (FSR) and it decreases with a larger imbalance. The FSR, in frequency and wavelength, is expressed as:

$$FSR[Hz] = \frac{c_0}{n_{og}\Delta L} \tag{2.3}$$

$$FSR[m] = \frac{\lambda^2}{n_{og}\Delta L} \tag{2.4}$$

Where, c_0 is the speed of light in vacuum, n_{og} is the optical group velocity, and λ is the wavelength at which the group velocity is obtained. The simulated group index in silicon rib and channel waveguide are shown in Fig. 2.4.

Ideally, an imbalance in a MZM is unnecessary because without an imbalance and without additional phase shift, light will interfere constructively for all wavelengths allowing for broadband operation. In reality, it is difficult to fabricate two arms of the interferometer that are exactly the same in terms of length and effective indices. In addition, it is also difficult to obtain identical splitting and recombination at the I/O couplers of the interferometer. This makes additional phase shifters and/or imbalances necessary in practical MZM designs for two main purposes: (1) modulator biasing, and (2) phase-shift measurements. This is further explained with TMM simulations in Section 2.3.3.

2.2.4 Silicon Thermo-optic Effect

The refractive index of bulk silicon increases with temperature as dictated by the thermo-optic coefficient:

$$\frac{dn}{dT} = 1.86 \times 10^{-4} K^{-1} \tag{2.5}$$

This thermo-optic effect is another phase-changing mechanism available in silicon. The thermo-optic coefficient is approximately constant for typical MZM operating temperatures. The relatively strong variation in index with temperature is advantageous for thermo-optic modulation and biasing of silicon modulators; however, it is also an obstacle with respect to thermal stability and strongly affects the behavior of certain silicon photonic devices such as ring resonators. From simulation, plotted in Fig. 2.6, the effective index of a silicon rib waveguide was determined to vary linearly with temperature at about the same rate as the thermo-optic coefficient.

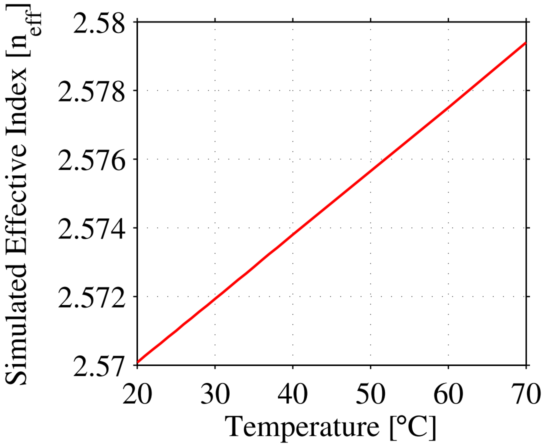


Fig. 2.6 Change in effective index in a rib waveguide at 1550 nm for increasing temperature.

2.3 Transfer Matrix Model

The different configurations of the MZI presented in Section 2.2.2 can be studied using the transfer matrix model (TMM). Each sub-component of the MZI is represented by a vector or a matrix and the final outputs are obtained by multiplying the individual matrices.

The matrix formulation of the electric field outputs of the y-branch is:

$$\begin{bmatrix} E_{Y,O1} \\ E_{Y,O2} \end{bmatrix} = \begin{bmatrix} \sqrt{\epsilon_1} \\ \sqrt{\epsilon_2} \end{bmatrix} [E_I] \tag{2.6}$$

Where, E_I, $E_{Y,O1}$, and $E_{Y,O2}$ are the electric field phasors of the input, output of arm 1, and output of arm 2, respectively. The ϵ_1 and ϵ_2 are the output power ratio of arm 1 and arm 2 of the y-branch relative to the input. For a lossless y-branch with identical splitting, ϵ_1 and ϵ_2 are 0.5.

The outputs of the 3-dB directional coupler, ignoring loss, in matrix form is:

$$\begin{bmatrix} E_{DC,O1} \\ E_{DC,O2} \end{bmatrix} = \begin{bmatrix} \sqrt{1-\epsilon} & j\sqrt{\epsilon} \\ j\sqrt{\epsilon} & \sqrt{1-\epsilon} \end{bmatrix} \begin{bmatrix} E_{I1} \\ E_{I2} \end{bmatrix} \tag{2.7}$$

Where, ϵ is the power splitting ratio between the two arms. The imaginary unit j represents the $\pi/2$ phase shift between the direct and cross coupled inputs.

The propagation in the arms of the interferometer is modeled by modifying the amplitude and the phase of the phasors.

$$\begin{bmatrix} E_{Oarm1} \\ E_{Oarm2} \end{bmatrix} = \begin{bmatrix} exp(-j\phi_1 - \frac{\alpha_1}{2}L_1) & 0 \\ 0 & exp(-j\phi_2 - \frac{\alpha_2}{2}L_2) \end{bmatrix} \begin{bmatrix} E_{Iarm1} \\ E_{Iarm2} \end{bmatrix} \tag{2.8}$$

Where, ϕ_1 and ϕ_2 is the total phase shift in arm 1 and arm 2, α_1 and α_2 is the optical propagation loss (in Nepers/unit length) in the respective arms, and L_1 and L_2 are the total lengths of each arm. The terms E_{Iarm1} and E_{Iarm2} are the input fields at the two arms of the interferometer, respectively, and E_{Oarm1} and E_{Oarm2} are the fields after the two phase shifting regions, respectively.

Phase shift differences between the two paths of the interferometer, in silicon MZMs, can be introduced by an imbalance in length (Section 2.2.3), thermal tuning (Section 2.2.4), and modulation by the plasma dispersion effect(Section 2.5). The imbalance can be unintentional, for instance due to variations in fabrication or intentional to create a measurable FSR and facilitate phase shift measurements. The total phase shift and loss can be divided into these three individual components, thus, the total phase shift for arm 1 and arm 2 can be expressed as:

$$\phi_1 = \frac{2\pi}{\lambda_0}[n_{eff} L_{nm1} + n_{eff}(V) L_{active1} + n_{eff}(T) L_{thermal1}] \tag{2.9}$$

$$\phi_2 = \frac{2\pi}{\lambda_0}[n_{eff} L_{nm2} + n_{eff}(V) L_{active2} + n_{eff}(T) L_{thermal2}] \tag{2.10}$$

Where, L_{nm1} and L_{nm2} are the non-modulated lengths of each arm, $L_{active1}$ and $L_{active2}$ are the electrically modulated lengths of each arm, and $L_{thermal1}$ and $L_{thermal2}$ are the thermally modulated lengths of each arm. The effective index of refraction is a function of voltage for electrical modulation and a function of temperature for thermal modulation. With this partition, the total length of an arm L is the sum of the lengths L_{nm}, L_{active}, and $L_{thermal}$ of that arm.

Finally, the total transmission can be obtained by multiplying the matrices together. For a 1×1 MZI, the output electric field is obtained as:

$$[E_O] = [Y]_{out} \, [MZM] \, [Y]_{in} \, E_{in} \tag{2.11}$$

Where, $[Y]$ represents the vector of the Y-branches (Eq. 2.6) at the output and input, and $[MZM]$ is the 2×2 matrix for the phase shift in the arms of the interferometer (Eq. 2.8).

The output intensity I_{out} is then obtained by the square of the absolute value of the output electric field. The optical intensity in decibels is:

$$I_{out,dB} = 10 \, log_{10}(|E_O|^2) \tag{2.12}$$

A simplified expression for the electric field at the output of the 1×1 MZI can be obtained by considering only one type of phase changing mechanism and assuming equal physical path lengths. Eq. 2.13 shows the electric field output for this special case and that it has both an amplitude modulation and a phase modulation component. For intensity modulation, the amplitude modulation is desired but the phase modulation is unwanted. On the other hand, for phase shift keying modulation formats, the amplitude modulation that occurs is undesired.

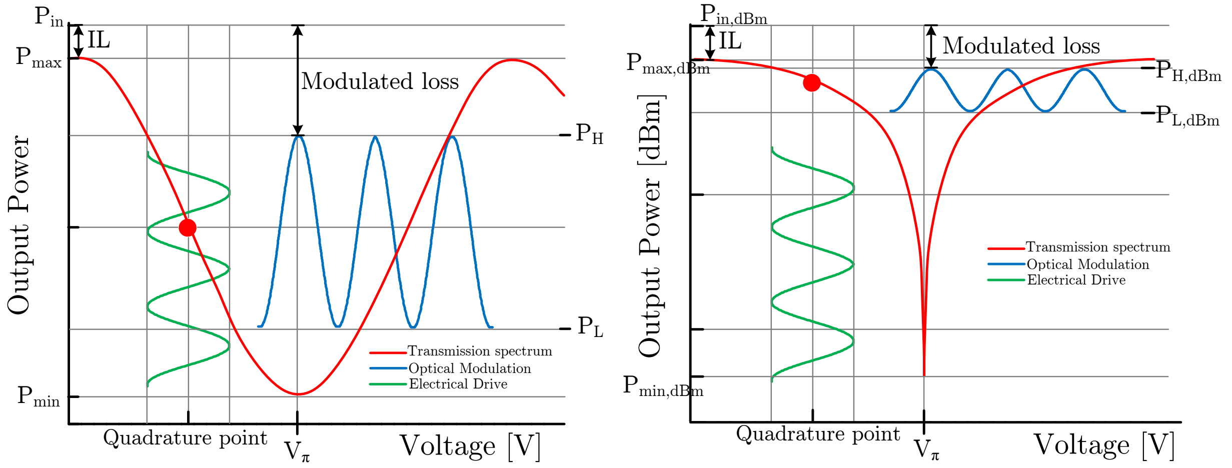
$$E_O = E_{in}\, \overbrace{cos\left[\left(\frac{\Delta\beta_2 - \Delta\beta_1}{2}\right) L\right]}^{\text{amplitude modulation}}\, \underbrace{exp\left[j\left(\frac{\Delta\beta_2 - \Delta\beta_1}{2}\right) L\right]}_{\text{phase modulation}} \tag{2.13}$$

Similarly, a simplified expression for the intensity at the output of the 1×1 MZI can be obtained by employing the exponential-hyperbolic and Euler's relation. The transmission of the MZI is:

$$T = \frac{1}{2}[1 + sech(\frac{\Delta\alpha}{2}L)cos(\frac{2\pi}{\lambda_0}(n_{eff1} - n_{eff2})L)] \tag{2.14}$$

Where, $\Delta\alpha$ $[Np/m]$ is the difference in loss between the two arms. The argument in the cosine term is the relative phase shift leading to the cosine term evaluating to 1 for constructive interference, and -1 for destructive interference.

2.3.1 Modulator Metrics



(a)

(b)

Fig. 2.7 Transmission of a MZM with applied voltage in (a) linear units, and (b) in decibel-milliwatts.

A diagram of the output power of an MZM with a linear change in index due to applied voltage is shown in Fig 2.7(a) with linear units and in Fig. 2.7(b) with units of decibel-milliwatts. Certain key metrics can be understood from these figures. For on-off keying (OOK) modulation, the modulator is biased at quadrature point, which corresponds to the mid-point of the linear transmission curve, or -3 dB of the maximum on the decibel transmission curve. At the quadrature point, the modulator is in the most linear regime. Incidentally, the largest change in transmission for a change in voltage ($\frac{dT}{dV}$) also occurs around this point.

The **insertion loss** (IL) of the device is determined by subtracting the maximum of the transmission from the input power, in decibels:

$$IL\,[dB] = P_{in,dBm} - P_{max,dBm} \tag{2.15}$$

The **extinction ratio** is defined as the ratio of the maximum power to the minimum power. It can be calculated for the modulated optical signal and for the optical device.

The **static** or **optical ER** (ER) is a metric of the passive MZM and can be calculated as:

$$ER\,[dB] = 10log_{10}\left[\frac{P_{max}}{P_{min}}\right] \tag{2.16}$$

It is related to the loss imbalance in the arms of an MZI as explained further in Section 2.3.2.

The **modulated ER** (ER_{mod}) is determined from the modulated optical signal. In decibels, it is expressed as:

$$ER_{mod}\,[dB] = 10log_{10}\left[\frac{P_H}{P_L}\right] \tag{2.17}$$

It should be noted that the modulated ER depends on the swing of the driving voltage and also the modulator biasing point. For instance, in OOK modulation, if the biasing point is moved from the quadrature point to a point corresponding to lower output power, i.e., down the transmission curve, then the zero-level optical power of the output (P_L) will also decrease. As the power of the modulated zero-level decreases, the modulated ER will increase.

However, the **optical modulation amplitude** (OMA), which measures the difference between the two optical power levels, i.e., the peak-peak height of the eye, will remain almost the same. It will be perturbed by the non-linearity of the transfer curve, but in the linear portion of the curve, OMA would mainly change with the driving voltage.

$$OMA = P_H - P_L \tag{2.18}$$

Thus, by changing the bias point, the modulated ER can be increased or decreased, however, the OMA remains about the same.

Changing the bias point for a fixed drive voltage also changes the **modulated loss** (ML) (also referred to as bias loss). ML is the difference between the input power and the power of the highest modulated bit.

$$ML\ [dB] = P_{in,dBm} - P_{H,dBm} \tag{2.19}$$

For instance, by moving the bias point down the transmission curve, the highest level of the modulated signal (P_H) will also lower. Therefore, in this situation, a higher modulated ER would be achieved with increased modulated loss and the same OMA.

Modulated loss can also differ from the insertion loss when the modulator is adequately biased at the quadrature point. This happens when the drive voltage is chosen to be lower than the voltage required for π phase shift (V_π). Then, the highest level (P_H) does not reach the maximum transmission of the modulator (P_{max}).

For a communication link, better metrics of performance than OMA and ER is the bit-error-rate (BER) measurement. BER measurements are also dependent on the performance of the detector and receiver.

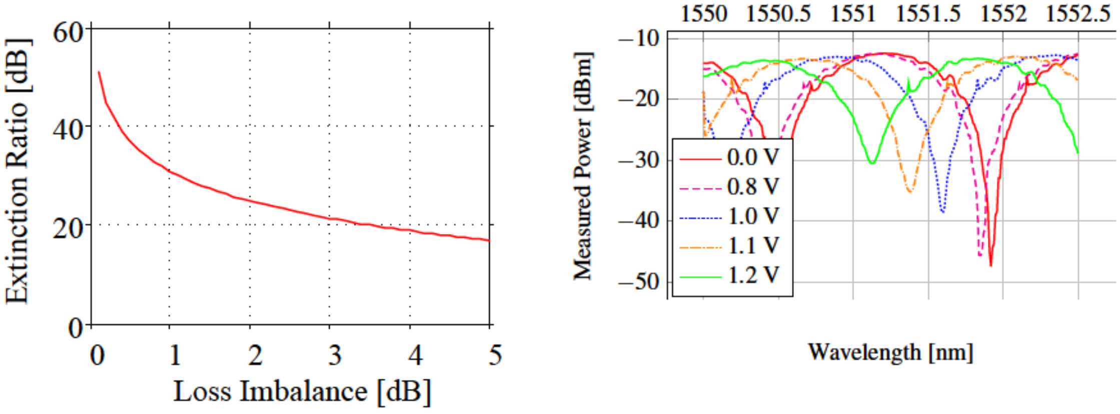
For electro-optic modulators, the voltage at which a phase shift of π is obtained is called the half-wave voltage (V_π). If it is assumed that Δn_{eff} is directly proportional to the applied voltage, as is the case with electro-optic modulation based on the Pockels effect, then from Eq. 2.1 it is known that a smaller voltage would be required for longer modulators to achieve a π phase shift. This leads to the definition of the **$\mathbf{V_\pi L_\pi}$ figure-of-merit** (FOM), which can be used to compare the efficiency of modulators. A low value of this FOM indicates a more efficient modulator. In silicon, it is important to note that phase shift is non-linear with applied voltage, as will be shown with experimental measurements in Section 5.2. Therefore, knowing the $V_\pi L_\pi$ FOM does not allow the accurate calculation of V_π for a given modulator length or the calculation of L_π for a given voltage. Furthermore, the phase shift in silicon saturates at higher voltages. This signifies that for short modulators, it may be impossible to obtain a π phase shift regardless of the magnitude of the voltage applied. It also signifies that the efficiency in silicon is dependent on the bias voltage and higher reverse bias voltages reduce the modulation efficiency.

2.3.2 Extinction Ratio Dependence on Loss Imbalance

The theoretical optical extinction ratio can be found by taking the ratio of the transmission (Eq. 2.14) at constructive interference and at destructive interference.

$$ER = \frac{[1 + sech(\frac{\Delta\alpha}{2}L)]}{[1 - sech(\frac{\Delta\alpha}{2}L)]} \tag{2.20}$$

Eq. 2.20 shows that the extinction ratio depends on the optical loss imbalance, and is depicted in Fig. 2.8(a). When a silicon modulator is being driven, the arm with the higher carrier concentration will have a higher loss than the other arm resulting in a degradation in the optical extinction ratio. This is most apparent with silicon MZMs operating in forward bias. For such modulators, the carrier injection in one arms leads to a significant increase in optical loss compared to the other arm of the modulator. Fig. 2.8(b) is the measured transmission spectrum of a silicon modulator in forward bias from [61]. It is observed that the extinction ratio of the green curve (1.2 V bias) is 18 dB lower than the extinction ratio of the red curve (0 V). The degradation in extinction ratio also occurs for reverse bias operation but the effect is typically negligible.



(a)

(b)

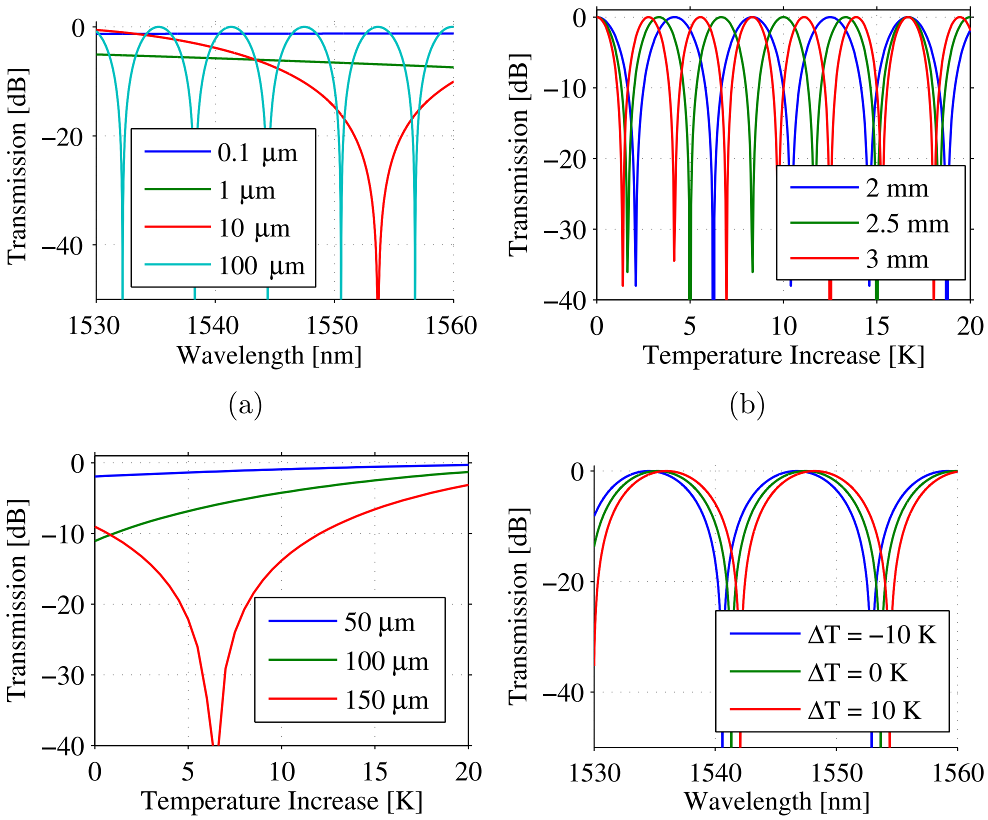
Fig. 2.8 (a) Reduction in maximum achievable extinction ratio as a function of loss imbalance in the arms of a Mach-Zehnder interferometer. (b) Transmission spectrum of a single-drive forward bias modulator depicting the reduction in extinction ratio (figure reproduced from [61]).

2.3.3 Temperature Stability, Imbalance, and Fabrication Variations

Several simulations were performed using the TMM (Eq. 2.11) for an ideal lossless MZI to study the effect of arm length imbalance, temperature stability, and variations in effective index.

In Fig. 2.9(a), the transmission spectrum for different lengths of imbalance is plotted. As expected from Eq. 2.3 and Eq. 2.4, the FSR decreases with increasing imbalance. With minute imbalances, e.g., 100 nm and 1 μm, the FSR is large enough that the spectrum is visibly flat over a large wavelength range. However, even with such minute imbalances, a significant decrease in the transmission is observed because the maximum of the transmission falls outside the simulated wavelength range. In reality, the maximum transmission may occur at wavelengths that is outside the range of a tunable laser or far away from the desired wavelength of operation. When such imbalances occur due to layout mistakes or fabrication errors, the reduction in transmission may be falsely interpreted as increased insertion loss. This subfigure also illustrates that optically broadband MZMs should ideally have zero imbalance in arms.

MZMs with imbalances should incorporate additional phase shifters in order to set the quadrature point of the modulator at a specific desired wavelength. These phase shifters need to be able to introduce a significant change in phase to decrease or increase the FSR,



(a) (b)

(c) (d)

Fig. 2.9 Simulated transmission characteristics of an ideal, lossless, 1×1 MZI: (a) change in the FSR for different imbalance lengths, (b) change in transmission resulting from a temperature increase in one arm for a balanced MZI with different lengths at 1550 nm, (c) change in transmission due to a temperature increase in both arms of the MZI for different imbalance lengths at 1550 nm (d) shift in transmission spectrum due to a temperature increase in both arms of the MZI for and imbalance of 50 µm.

therefore, thermo-optic or forward bias phase shifters should be used for this function.

Because of fabrication variations, it is very difficult to obtain symmetrical MZMs in a silicon photonic process. These variations include SOI wafer thickness, and deviations in waveguide widths due to limitations associated with resists and the photolithography processes. This asymmetry has been shown to be worse in a 248 nm lithography process compared to a 193 nm process [62]. For this reason, additional phase shifters are necessary when operation of a MZM is required at a specific wavelength. For silicon MZMs based on the carrier depletion effect, it is also important to have imbalances or additional active phase shifters because the modulation effect is weak and becomes even weaker with increasing reverse-bias voltage. Thus, it is not always possible to bias short modulators at the quadrature point with a DC bias applied to the RF signal. Even for long modulators, if a high reverse bias voltage needs to be applied to be at the quadrature point, this would lead to a poorer modulation efficiency.

The change in transmission as temperature increases in one arm of the MZI for different MZI lengths is shown in Fig. 2.9(b). A 2 mm arm with a temperature change of 4.1°C results in a 2π phase shift. In Fig. 2.9(c), the change in transmission as a function of temperature increase in both arms of the MZI for MZIs with different imbalance lengths is shown. This subfigure demonstrates that a MZI is inert to temperature variations only if it is perfectly symmetrical. With an imbalance, either intentional or due to fabrication variations, this stability degrades. For a MZI with a 50 μm imbalance, a change in temperature of +/- 10°C can lead to a significant change in the transmission spectrum as depicted in Fig. 2.9(d).

2.3.4 Modulation Chirp

Chirp is defined as the time variation of frequency and frequency is the time derivative of phase. From Eq. 2.13, it is known that intensity modulation is accompanied by a change in phase. The chirp then is simply the time derivative of that phase modulation $\phi(t)$:

$$\Delta\nu(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} \tag{2.21}$$

Chirp is generally quantified by various α-parameters. One such parameter is the α_{3dB}, defined as:

$$\alpha_{3dB} = 2\frac{\phi(on) - \phi(3\,dB)}{ln\,(I(on)) - ln(I\,(3\,dB))} \tag{2.22}$$

For on-off keying, changes in phase accompanying the intensity modulation is inconsequential because a photo-detector only detects intensity. However, this chirp in conjunction with chromatic dispersion can aggravate either pulse compression or broadening depending on the medium, which can affect the ability to transmit for long distances. Silicon series push-pull modulators have been shown to have a slightly negative chirp, which makes transmission more tolerant to chromatic dispersion in single-mode fiber [63].

2.4 Modulator Driving Configurations

Mach-Zehnder modulators can be driven in three main configurations: single drive, dual drive, and series push-pull. In order to maintain reverse bias for carrier depletion silicon Mach-Zehnder modulators, the n-terminal of the diode must be at a higher potential than the p-terminal. In case the diodes become forward biased, negligible change in optical transmission occur if the voltage across the diode remains below threshold. Moreover, at high data rates, a forward biased diode is too slow to react, thus slight forward biasing is still tolerable. Unlike $LiNbO_3$ modulators, the efficiency of silicon modulators is bias dependent. Silicon modulators are more efficient for low reverse bias voltages, as will be demonstrated by the measurements in Section 5.2.

2.4.1 Single and Dual Drive Modulators

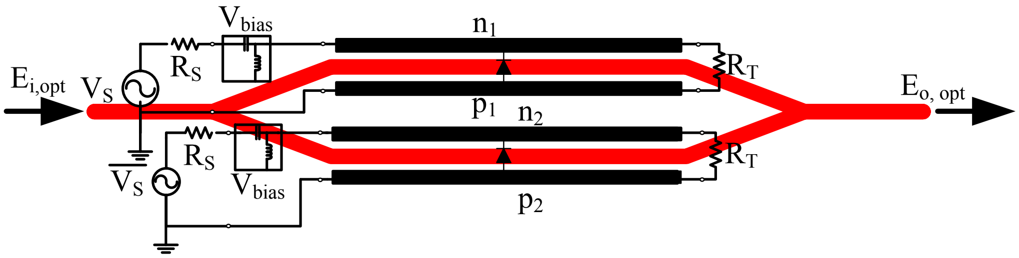
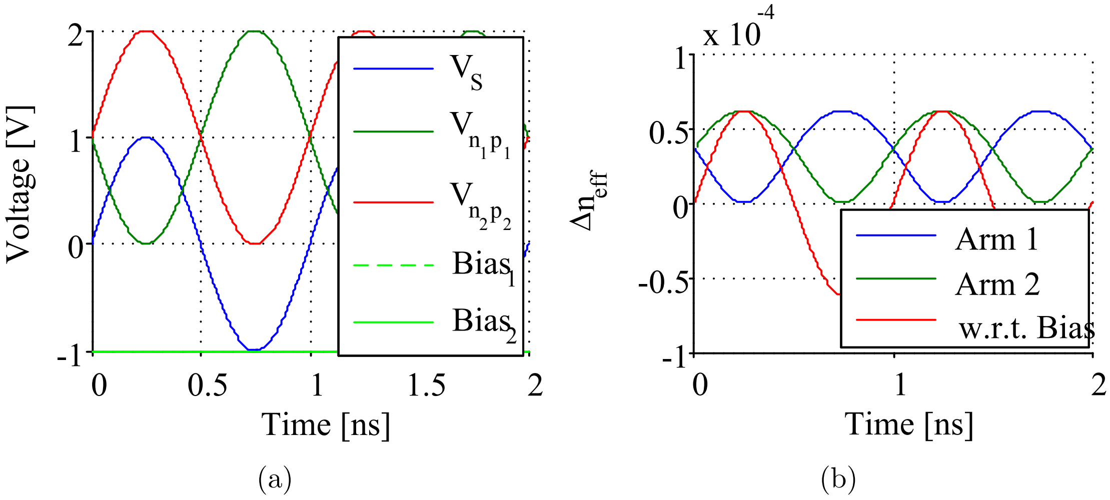


Fig. 2.10 Dual-drive modulator driving configuration.

A schematic of a dual drive modulator with its driving circuitry is shown in Fig. 2.10. A single drive modulator would be similar except that it will have only one arm of the interferometer being driven. A bias tee can be used to offset the RF signal to ensure that $V_n > V_p$ and even greater bias voltage can be used for larger small-signal bandwidth. The magnitude of the minimum bias voltage required is equal to the peak-peak drive voltage. The driving voltages for a dual-drive modulator are complementary as shown in Fig. 2.11(a) and the same bias voltage can be used on both arms. The corresponding change in effective index using data from a silicon modulator is plotted in Fig. 2.11(b). Because of the non-linearity of the carrier depletion modulation, it is observed that the change in effective index between the two arms is asymmetrical. $LiNbO_3$ modulators are driven in dual drive because it removes chirp by obtaining equal but opposite index changes in the two arms, which does not occur perfectly in silicon. Dual drive operation also approximately (but not exactly due to the non-linearity) halves the drive voltage compared to a single-drive modulator because one arm can have a $+\pi/2$ phase shift while the other has a $-\pi/2$ shift. Relative to the effective index at the bias voltage, the change in Δn_{eff} with this driving scheme would almost be equal to that obtained with twice the drive voltage on a single arm, as depicted in Fig. 2.11(b). Despite two driving voltages needed for dual drive, the halving of the drive voltage halves the dynamic power consumed by the modulator. However, the additional driver will also increase power consumption of the system.



(a) (b)

Fig. 2.11 Dual-drive modulation with a source voltage V_S of 1 V_{pp} with both arms at the same bias voltage of -1 V. Voltage across each diode is shown in (a) and the resulting index change for both arms and with respect to the bias point is shown in (b).

2.4.2 Series Push-Pull Modulator

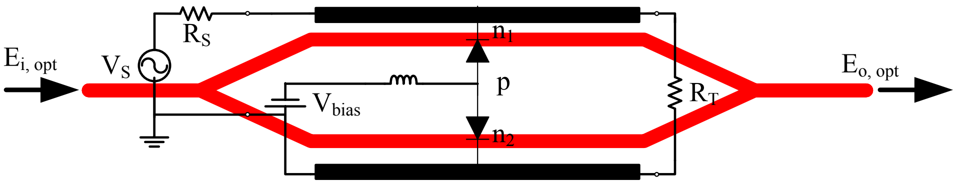


Fig. 2.12 Series push-pull modulator driving configuration.

In a series push-pull (SPP) (also called a single-drive push-pull) modulator, the two diodes of each arm of the interferometer are connected back-to-back as illustrated in Fig. 2.12 and the DC bias is applied to the common node. This results in the doubling of the junction resistance and the halving of the junction capacitance. Although the RC constant remains about the same, this will lower microwave loss as explained later in Section 3.4. In addition, with this configuration, push-pull operation can be obtained with a single modulator driver and electrode pair. Ideally, this push-pull would also remove modulator chirp, but because the carrier depletion effect is non-linear, the change in effective index is not symmetrical as shown in Fig. 2.13(b). It should be noted that due to fabrication variations one arm of an MZM can be much more effective at modulation than the other arm making push-pull even less effective at eliminating chirp. However, a negative chirp, if present, has been found to be beneficial for propagation through fibers in the C-band [63].

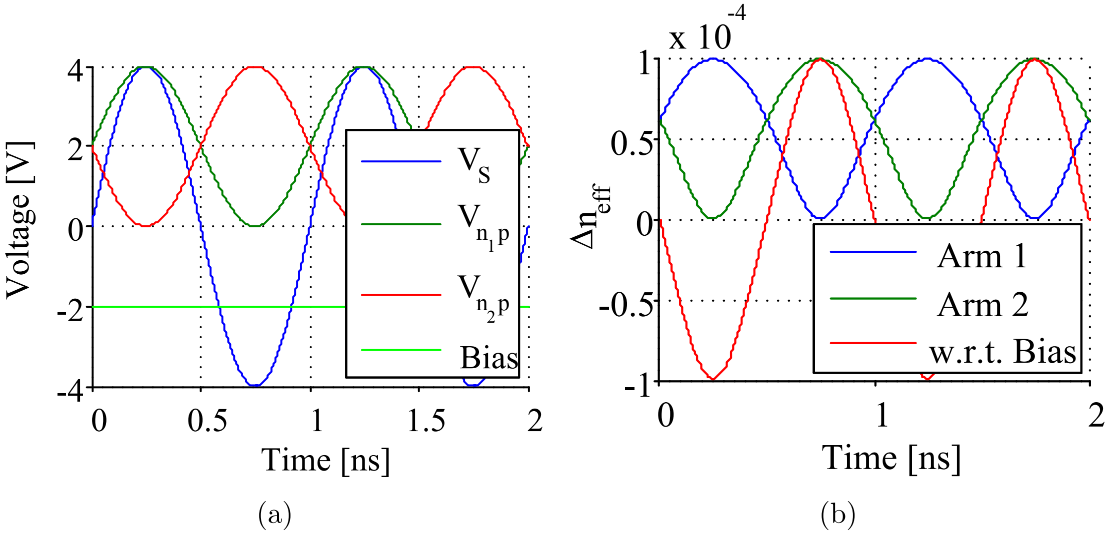
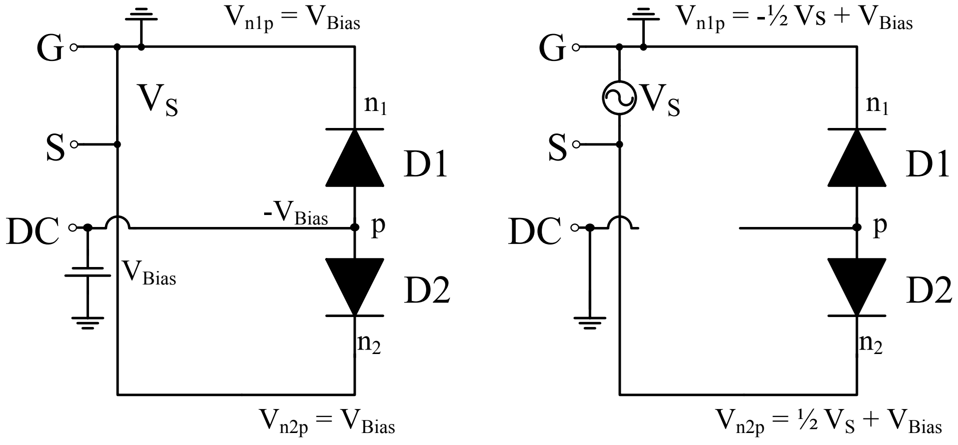


Fig. 2.13 Series push-pull modulation with a source voltage V_S of 8 V_{pp} with a bias voltage of $-2\ V$. Voltage across each diode is shown in (a) and the resulting index change for both arms and with respect to the bias point is shown in (b).

In SPP modulators, a bias tee is not needed as the RF voltage out of the modulator driver is not offset by the DC bias. It is however necessary to have an inductance between the DC bias voltage source and the common diode node [63]. At DC, the inductor behaves as short circuit and the bias voltage is applied to the common node to maintain reverse bias as shown in Fig. 2.14(a). In the AC regime, the inductance is approximately an open circuit as shown in Fig. 2.14(b). The final voltage is obtained by applying the superposition principle. If it is assumed that both diodes of the interferometer are identical, then the RF voltage is equally split amongst the two diodes. As the RF is split in half, the minimum magnitude of the bias voltage to stay in reverse bias is $V_{s,pp}/4$. Since in this particular illustration the n-side of diode 1 is grounded, the AC component of the reverse bias voltage V_{n1p} will have the opposite polarity of V_{n2p} creating a push-pull drive.

With an insufficient inductance for the bias pad, the low frequency components of the RF signal would get shorted. This can degrade SNR, increase jitter, and cause baseline wander. The need for a long inductance trace and a DC pad slightly increases the footprint of a SPP modulator compared to a single-drive modulator.



(a) (b)

Fig. 2.14 Circuit models for series push-pull modulators in the (a) DC regime and (b) AC regime with the net voltages across the diodes listed in the figure.

2.4.3 Power Estimates

In traveling wave modulators, power is mostly consumed from the resistive losses of the transmission line and the termination impedance. The power estimates of traveling wave modulators generally assume that the power consumed by the charge-discharge cycle of the p-n junction resistance-capacitance circuit is negligible. On the other hand, in lumped modulators without a termination, the power consumed by the modulator is primarily due to the charging and discharging of the junction capacitance. It is also important to note that throughout the thesis, it is the power consumed by the modulator that is estimated. The power consumed by the entire system, which would for example include the data source, RF amplifiers, laser, DAC, ADC, photo-receiver, and other equipment are not included.

For single and dual drive modulators, the bias voltage also drops across the transmission line and the termination resistor, thus the static power consumption can be significant. Therefore, it is desirable to operate these modulator configurations at low reverse bias voltages. The static power can be determined by Ohm's law as $P_{static} = V_{bias}^2/R_T$, where R_T is the termination resistance. In series push-pull modulators, the bias voltage does not drop across the transmission line or the termination resistance. The current supplied by the bias voltage in an SPP modulator was observed to be below the limit of our current meter ($<1\,\mu A$). To eliminate the static power consumption in single or dual-drive modulators, a DC block or a capacitor can be placed between the transmission line and the termination impedance [45, 46].

The dynamic power is estimated assuming a rectangular driving voltage with equiprobable ones and zeros. This leads to the expression for the RMS power of a 50% duty cycle square wave: $P_{dynamic} = (V_{pp}/2)^2/R_T$. It is assumed that the impedance of the modulator is matched to the source and termination.

The power consumption for additional active phase shifters for biasing should also be included in the total power consumption. The total power is then multiplied by a bit period to obtain the average energy consumed per bit. In addition, it should be noted that a SPP modulator does not require an inverted data source and one modulator driver less than a dual-drive modulator.

A summary of the advantages and disadvantages of different modulator driving configurations is provided in Table 2.1.

Table 2.1 Comparison of different modulator driving configurations.

	V_{drive}[a]	Static power[b]	Dynamic Power	Bias	Chirp[c]	Microwave Loss	Footprint
Single Drive	V_π	$\frac{V_{bias}^2}{R_T}$	$\frac{(V_{pi}/2)^2}{R_T}$	Bias-T	Yes	Higher	Small
Dual Drive	$2{\times}V_{\pi/2}$	$2{\times}\frac{V_{bias}^2}{R_T}$	$0.5\frac{(V_{pi}/2)^2}{R_T}$	Bias-T	No	Higher	Large
Series Push-pull	V_π	$< \mu W$	$\frac{(V_{pi}/2)^2}{R_T}$	On-chip	No	Lower	Medium

[a] Assuming a full phase shift is desired.

[b] Assuming identical biasing for both arms of the dual-drive modulator. Also assuming no DC-block is used.

[c] Assuming ideal diodes and symmetrical MZM for chirp cancellation.

2.5 Modulation in Silicon

Optical modulation is achieved by a change in either the real or imaginary (extinction coefficient) part of the refractive index. Changes in the extinction coefficient modulate the absorbance of light and can allow optical modulators to be created without interferometry as in electro-absorption modulators. In silicon, however, the most widely used method is the modulation of the real part of the refractive index in a MZI or a ring resonator configuration. The change in refractive index of the material with an electrical stimulus results in a change in phase as light travels through it as indicated by Eq. 2.1. A MZI subsequently converts the phase shift into intensity modulation.

The complex refractive index $\overline{n}$ can be expressed as:

$$\overline{n} = n + jk \tag{2.23}$$

Where, n is the real part, and k is the extinction coefficient. The absorption coefficient α from the extinction coefficient is given by:

$$\alpha = \frac{4\pi k}{\lambda_0} \tag{2.24}$$

The real part of the refractive index n can be calculated from the absorption coefficient (or vice versa) by the Kramers-Kronig relation. Universally, the Kramers-Kronig relations link the real and imaginary parts of a complex function. The complex function must be without singularities and branch cuts. In other words, it must be analytic, or infinitely differentiable. The relations are applicable to physical systems as they are causal and that implies analyticity. For calculating the change in index from a change in absorption, the relation is expressed as:

$$\Delta n(\omega) = \frac{c_0}{\pi} PV \int_0^{\infty} \frac{\Delta\alpha(\omega')}{\omega'^2 - \omega^2}\, \mathrm{d}\omega' \tag{2.25}$$

Where, c_0 is the speed of light in vacuum, ω is the angular frequency at which the refractive index is to be calculated, and the integration variable $\omega^{'}$ is the angular frequency for which the absorption spectrum data is to be integrated. A singularity exists at $\omega = \omega^{'}$, and the Cauchy principle value, denoted by PV, must be taken. Hence, to calculate the refractive index at frequency ω, the absorption spectrum is integrated over the entire frequency range such that singularities are avoided. In practice, it is much easier to measure the absorption spectrum as oppose to directly measuring the change in refractive index. Thus, the refractive index is usually numerically evaluated with the Kramers-Kronig relations from absorption spectra measurements. The change in absorption spectrum, $\Delta\alpha(\omega')$ of Eq. 2.25, from an applied electrical field E can be written as:

$$\Delta\alpha(\omega', E) = \alpha(\omega', E) - \alpha(\omega', 0) \tag{2.26}$$

In the case of a change in absorption due to change in carrier concentration ΔN:

$$\Delta\alpha(\omega', \Delta N) = \alpha(\omega', \Delta N) - \alpha(\omega', 0) \tag{2.27}$$

Perturbations in the complex refractive index $\overline{n}$ of a material can be induced by the application of a electrical field (electro-optic effects) or by modulating the free carrier concentration. There are several different electro-optic effects, of which, the most widely studied are the Pockels, Kerr, and the Franz-Keldysh effect (electro-absorption). The Pockels effect, which induces a change in refractive index n directly proportional to the applied electric field, is present in non-centrosymmetric crystals. Unstrained silicon is centrosymmetric, hence it has no Pockels effect unless it is deliberately strained, for example, by the growth of silicon nitride [64]. The Kerr effect, which changes the refractive index n quadratically with an applied electric field, is present in silicon. It is, however, found to be weaker than the plasma-dispersion effect [65]. Near 1550 nm, an applied electric field of 10^5 V/cm, results in a Δn of approximately -9×10^{-7}, far below the requirement for a π phase shift of a reasonably long modulator with practical voltages (Fig. 2.1). The Franz-Keldysh effect arises due to perturbation of the energy bands upon the application of an electric field. This leads to a change in the bandgap, and manifests as a change in the optical absorption spectrum of silicon. The Franz-Keldysh effect also introduces a change in refractive index n (electro-refraction), and a Δn of 1.7×10^{-6} for an applied electric field of 10^5 V/cm was measured near 1550 nm[65]. It should be noted that the dielectric breakdown of crystalline silicon with intrinsic doping concentration is in the same order of magnitude, approximately 4.5×10^5 V/cm.

In silicon, the strongest refractive index change comes from the plasma dispersion effect, in which light is scattered by mobile carriers. The presence of free carriers results in shifts in the absorption spectrum towards the short and long wavelengths, as well as free carrier absorption (FCA). The change in FCA for a change in carrier concentration of holes ΔN_h and electrons ΔN_e relative to the intrinsic concentration in a material is predicted by the Drude-Lorenz equation:

$$\Delta\alpha = \frac{e^3\lambda_0^2}{4\pi^2 c_0^3 \epsilon_0 n}\left(\frac{\Delta N_e}{\mu_e (m^*_{ce})^2} + \frac{\Delta N_h}{\mu_h (m^*_{ch})^2}\right) \tag{2.28}$$

Where e is the electronic charge, λ_0 is the free space wavelength, c_0 is the speed of light in vacuum, ϵ_0 is the permittivity of free space, ΔN_e and ΔN_h are the free electron and hole concentrations respectively, μ_e and μ_h are the mobility of the electrons and holes respectively, and m^*_{ce} and m^*_{ch} is the effective mass of electrons and holes respectively. Similarly, the theoretical expression for the change in refractive index can be described by:

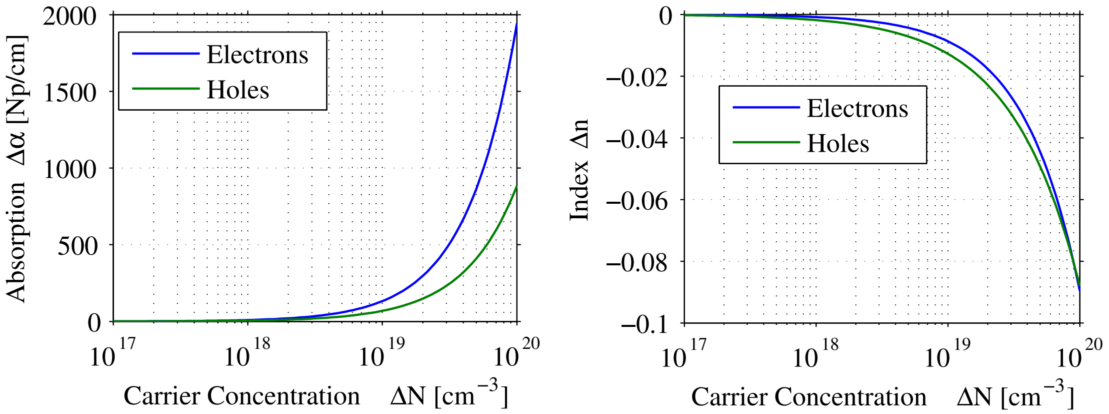
$$\Delta n = \frac{-e^2\lambda_0^2}{8\pi^2 c_0^2 \epsilon_0 n}\left(\frac{\Delta N_e}{m^*_{ce}} + \frac{\Delta N_h}{m^*_{ch}}\right) \tag{2.29}$$

As hinted by the Kramers-Kronig dispersion relation (Eq. 2.25), it is not possible to change the refractive index without also having an effect on the optical loss. An increase in effective index will typically be accompanied by an increase in optical loss. The effect on the refractive index due to changes in carrier concentration was determined in [65], and more recently in [66]. The measurements indicated a discrepancy in the theoretically predicted changes in the refractive index from the simple Drude-Lorenz model. In [66], a power-law model was fitted to the index change obtained by applying the Kramers-Kronig relation to the measured values of absorption for free carrier concentrations ranging from 10^{17} to $10^{20} cm^{-3}$. At 1550 nm, the change in attenuation and refractive index of doped silicon is expressed as:

$$\Delta\alpha = \Delta\alpha_e + \Delta\alpha_h = 8.88 \times 10^{-21}(\Delta N_e)^{1.167} + 5.84 \times 10^{-20}(\Delta N_h)^{1.109} \tag{2.30}$$

$$\Delta n = \Delta n_e + \Delta n_h = -[5.40 \times 10^{-22}(\Delta N_e)^{1.011} + 1.53 \times 10^{-18}(\Delta N_h)^{0.838}] \tag{2.31}$$

As Eq. 2.31 indicates, the refractive index decreases when carriers are injected, and increases when carriers are depleted. The trend for the attenuation as a function of mobile carrier concentration is plotted in Fig. 2.15(a), and in Fig. 2.15(b) for the refractive index. It is apparent that an increase in hole concentration results in a smaller increase in attenuation than when the electron concentration is increased. A change in hole concentration also causes a slightly larger change in index. It is also observed that attenuation and the change in index increases substantially for doping concentrations greater than approximately 5×10^{18} cm^{-3}.

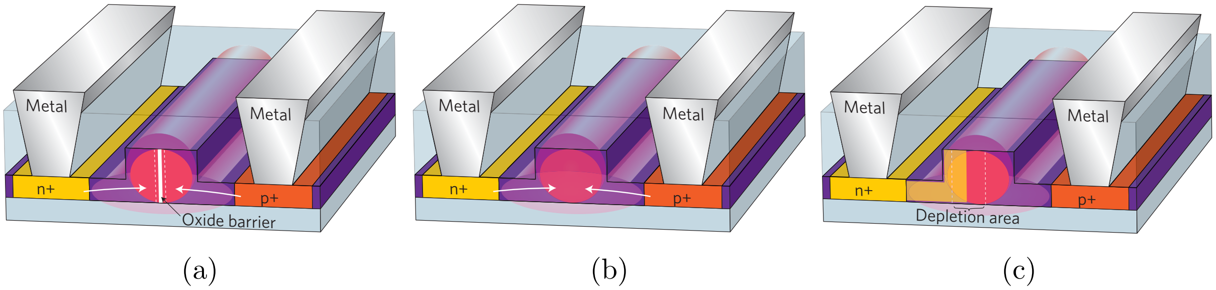


(a) (b)

Fig. 2.15 Change in (a) attenuation, and (b) refractive index in bulk silicon for different free carrier concentration.

2.6 p-n Junction Design

The carrier concentration for the plasma dispersion effect in silicon can be modulated by various structures. The most commonly studied silicon optical modulator structures are based on carrier accumulation, carrier injection, or carrier depletion. Each structure must be able to guide light while allowing to modulate the carrier concentration in the region of the optical mode.



(a) (b) (c)

Fig. 2.16 Methods of plasma dispersion modulation in silicon: (a) carrier accumulation, (b) carrier injection, and (c) carrier depletion. Figure reproduced from [67] with permission.

In carrier accumulation, a metal-insulator-semiconductor (MIS) stack is formed. The metal can also be a poly-silicon layer or a highly doped semiconductor and the insulator is typically oxide. The semiconductor which also guides the optical field is doped to have mobile carriers. A carrier accumulation structure is illustrated in Fig. 2.16(a). In a MIS structure with an n-type doped semiconductor, a positive potential applied at the metal would result in holes accumulated at the metal-oxide interface. This accumulation of holes leads to an accumulation of electrons at the oxide-semiconductor interface. The accumulation of charges in the waveguide region modifies the refractive index, thus creating modulation. Carrier accumulation structures can also be formed by inserting an insulator in a lateral p-n junction. Another structure is the Polysilicon Gate Oxide Insulator Silicon Capacitor (SISCAP) structure which is a vertical stack of polysilicon, oxide, and semiconductor [68]. The optical mode is partially present in the polysilicon and the semiconductor. Carrier accumulation has a relatively fast response because it involves the movement of the majority carriers and recombination is avoided. MIS structures can typically also operate in depletion and inversion mode,

Carrier injection and carrier depletion (which is different from the depletion regime of MIS) modulators use a diode-waveguide structure as illustrated in Fig. 2.16(b) and Fig. 2.16(c) respectively. The mode of operation, either carrier injection or depletion, is determined by the biasing of the diodes.

2.6.1 Reverse and Forward Bias Diodes

When the diode-waveguide structures is operated in forward bias, carriers are injected into the waveguide. In reverse bias operation, carriers are depleted. The small signal equivalent circuit model for a reverse and forward biased p-n junction is shown in Fig. 2.17 [69]. Here, c_j represents the junction capacitance, c_d the diffusion capacitance, G_d the conductance, and r_b the resistance up to the edge of the depletion region (including the contact resistance) for the n and p regions.

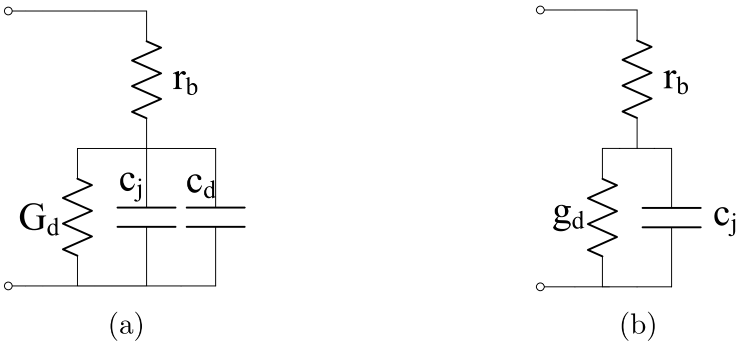


Fig. 2.17 Small signal equivalent model of a p-n junction in (a) forward bias, and (b) reverse bias.

In forward bias, the change in refractive index occurs due to minority carrier injection, and the effect is relatively strong resulting in highly efficient modulation. Low $V_\pi L_\pi$ of 0.28 V-cm [70] and of 0.075 V-cm [61] have been demonstrated. The increase in carrier concentration increases loss and this reduces the extinction ratio, as discussed in Section 2.2. Moreover, the diffusion capacitance, which represents the storage of excess minority carriers, is non-negligible. Because of this carrier life-time, the modulation speeds are very slow. High bit-rates can only be achieved with pre-emphasis or by reducing the lifetime by introducing impurities to create trap states. In a forward bias modulator, the modulation is very weak below the diode threshold (also refereed to as the cut-in or knee voltage). However, just a small voltage beyond threshold injects enough carriers to obtain a π phase shift.

When reverse biased, there is no charge injection and therefore no diffusion capacitance. This leads to a much higher intrinsic bandwidth (f_{3dBPN}) of the junction. The modulation is performed by the extraction of minority carriers representing the depletion region which overlaps with the optical mode.

$$f_{3dBPN} \approx \frac{1}{2\pi r_b c_j} \tag{2.32}$$

The junction capacitance arises from the depletion region, which is also present in the forward biased case but is much smaller in reverse biased diodes. Theoretically, the depletion width increases with increasing reverse bias with a square root dependence.

$$W_d \approx \sqrt{\frac{2\epsilon_s}{q}\left(\frac{1}{N_A}+\frac{1}{N_D}\right)(V_{bi}-V_a)} \tag{2.33}$$

Where, ϵ_s is the permittivity of the semiconductor, N_A and N_D are the acceptor and donor concentrations, V_{bi} is the built-in voltage, and V_a is the applied voltage. Therefore, the junction capacitance ($c_j = 1/W_d$), ignoring the fringe fields, is inversely proportional to the square root of the applied bias. A more comprehensive model of the junction capacitance is described in [71]. Because of this dependence, the value of the capacitance c_j and the resistance r_b in the small signal model are dependent on the bias conditions. The junction capacitance also increases with increasing concentration of the donor (N_D) or acceptor (N_A) dopants.

The equivalent circuit models can be used to model the p-n load in the telegraphist model of traveling wave electrodes as will be described in Section 3.4. Unfortunately, for reverse biased modulators, relatively large voltage swings are needed in order to achieve modulation. Since the junction capacitance is voltage dependent, this means the transmission line's characteristics will also depend on the bias voltage.

In a p-n junction, the intrinsic bandwidth is determined from the single pole RC circuit. For a greater bandwidth, the capacitance can be reduced at the junction by intentionally placing a small gap between the p and n interface to reduce an excess doping that occurs from the lateral scattering during implantation [72]. The resistance of the junction can be reduced by increasing the doping concentration; however, depending on the distance from the optical mode, this will lead to an increase in optical loss.

2.6.2 Diode-waveguide structures

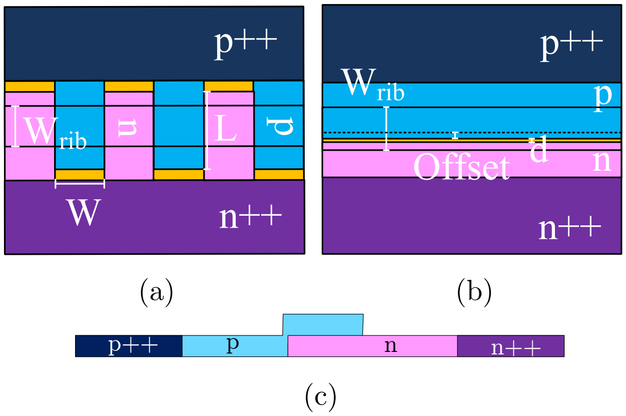


Fig. 2.18 Examples of diode-waveguide structures for modulation. (a) Interdigitated (top view), (b) lateral p-n (top view), and (c) vertical (cross-section) junctions.

Optical modulation can be performed with lateral, interdigitated, or vertical p-n junctions as shown in Fig. 2.18.

Interdigitated p-n junction

The modulation efficiency of the interdigitated (or interleaved) p-n junction is determined by the density of the junctions along the length of the modulator, which in turn is controlled by the width (W), and the overlap (L) between the n and p fingers. It was shown that the interdigitated structure offers better modulation efficiency, a $V_\pi L_\pi$ lower by 40% [37], than the lateral p-n junction. This is because the interdigitated p-n junction completely overlaps with the optical mode. This, however comes at the cost of higher capacitance (efficiency and capacitance are related) which decrease the junction bandwidth. The interdigitated junction had 3.6 times more capacitance, and the capacitance increases linearly with the overlap length. Furthermore, higher capacitance also increases the microwave loss as will be discussed in Section 3.4. Hence, in terms of speed, lateral doping profiles was determined to be much better than the interdigitated junctions [37].

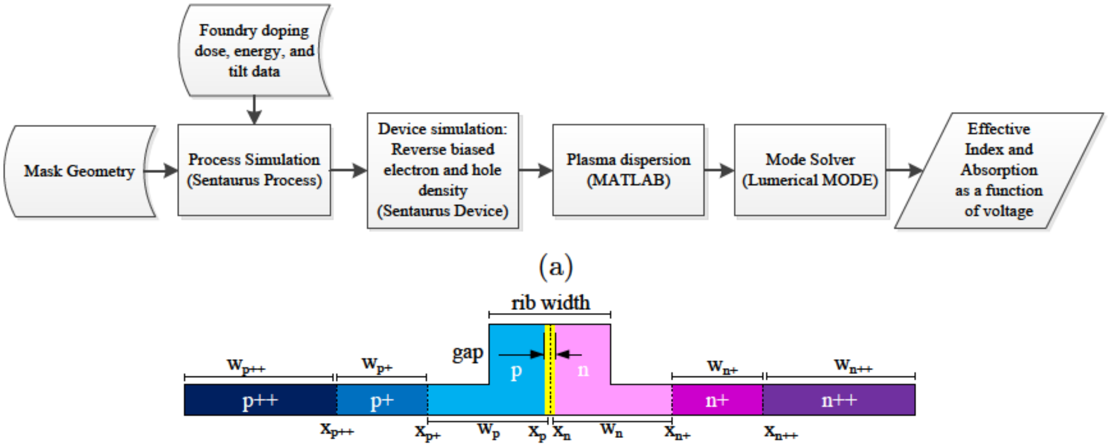
Lateral p-n junction

For the case of the lateral junction, better modulation efficiency was achieved with a 40 nm spacing (d) between the p and n doping. At the same time, an optimal spacing (d) can reduce the capacitance because it lowers the number of carriers at the junction edge [37, 73]. In addition, an offset can be introduced to increase the optical field overlap with the p-type dopant. This results in a better modulation efficiency as the p-type semiconductor has lower losses and stronger index change as described in Section 2.5.

Vertical p-n junction

A vertical p-n junction has a greater overlap with the modulation region and the optical mode, thus it has a higher efficiency than a horizontal p-n junction [73]. In [73], the vertical p-n junction had a $V_\pi L_\pi$ that was 2.28 times smaller. However, the vertical p-n junctions have a much smaller bandwidth due to a high RC constant. The vertical p-n junction was measured to have a resistance that was four times greater, and a capacitance that was 1.9 times larger.

2.6.3 Lateral p-n Junction Simulations



(a)

(b)

Fig. 2.19 (a) Simulation flow for different lateral p-n junction geometries. (b) Lateral p-n junction with variable design parameters.

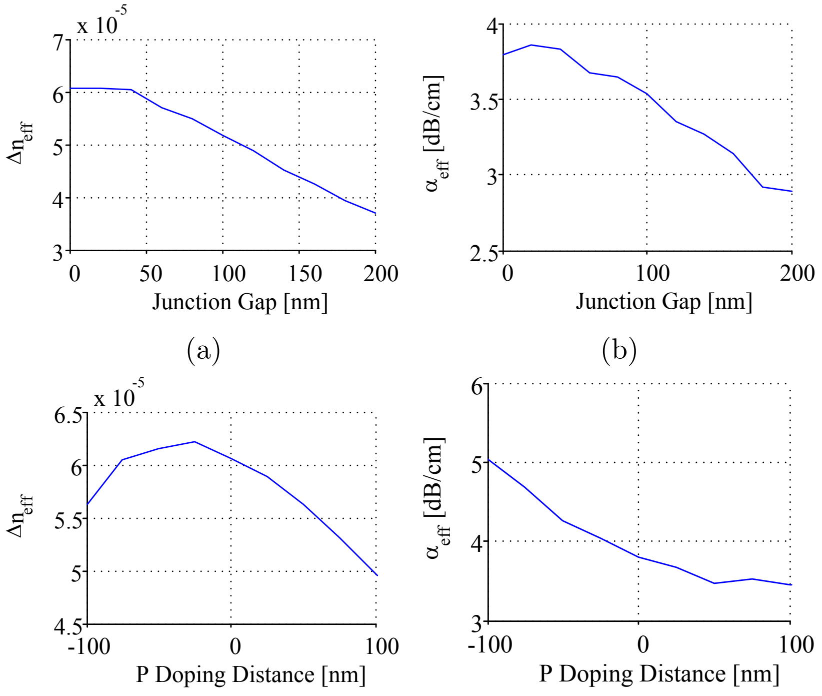
The OpSIS-IME process that is used to fabricate the device had six implant levels for modulators: n++, n+, n, p, p+, and p++. The peak concentrations for this process are: 5×10^{20}, 3×10^{18}, 5×10^{17}, 7×10^{17}, 2×10^{18}, and 1.7×10^{20}, respectively [34, 74].

The simulation methodology is shown in Fig. 2.19(a). A process simulation was first carried out using Synopsys Sentaurus Process to determine the effective doping concentration and their distribution. This was performed using the implant and annealing conditions provided by the foundry. The process simulation used Synopsys's advanced calibration parameters. A device simulation was then performed with the Synopsys Sentaurus Device tool to obtain the distribution of free carriers in the waveguide for different bias voltages. Sentaurus Device solves the drift-diffusion equations for electron and hole densities coupled with the Poisson equation for the electrostatic potential, and Fermi statistics. Other activated physical models in this simulation were incomplete ionization, Shockley-Read-Hall recombination, and doping-dependent mobilities. The change in the index of refraction from the free carriers was calculated by applying Eq. 2.30 and Eq. 2.31 in MATLAB. The index profile was then imported, for each voltage, in Lumerical MODE to calculate the optical effective index.

Although the simulations could not be validated in this work, good agreement between simulations with this method and measurements were shown in [61, 75]. It should also be noted that mask geometry information provided in the simulator was exactly as would be drawn in the layout, and there was no other empirical fitting involved. Numerous simulations were performed to study the impact of varying the dimensions and positions of each dopant type as annotated in Fig. 2.19(b). However, without additional control in the process (e.g., dopant concentration), only few simulated trends are worth noting and none were shown to give a substantial improvement in performance.

Changes in the position of the advanced (x_{p++} and x_{n++}) and intermediate (x_{p+} and x_{n+}) dopants lead to negligible changes in the effective index (thus, negligible changes in $\Delta\phi$, $V_\pi L_\pi$, and junction capacitance). The largest changes were observed in the optical attenuation and the resistance of the junction. As each position increased, optical loss decreased and the resistance increased. For the intermediate doping levels, reduction in optical loss became insignificant after about 450 nm from the center of the waveguide. Similarly, for the advanced doping levels, reduction in optical loss became insignificant after about 1000 nm from the center of the waveguide.

In another set of simulations, a gap was introduced at the junction and was varied from 0 nm to 200 nm. In summary, there was almost negligible changes in effective index for a gap up to 40 nm. Beyond a gap of 40 nm, modulation decreased as shown in Fig. 2.20(a). As the gap increases, there is a reduction in optical loss, as plotted in Fig. 2.20(b). However, for a gap of 0 to 40 nm, the decrease in loss is negligible. The depletion capacitance, in the region of 0 to 40 nm peaks in the middle, and is lowest at the edges. The p-n junction series resistance increases substantially due to the increase in the size of the intrinsic region.



(a) (b)

(c) (d)

Fig. 2.20 Simulated p-n junction (a) phase shift and (b) attenuation for a junction with a gap, and (c) phase shift and (d) attenuation for different offsets of the p-type semiconductor (negative axis distance indicates an offset in to the n region). All simulations shown are at 1.5 V reverse bias.

Afterwards, the effect of changing the offset for the p-dopant was studied. The best effective index was observed when the p doping had a 25 nm offset from the center as shown in Fig. 2.20(c). This is close to the resolution control that the process allows for implant masks. Up until about 75 nm, the effect on the index is greater or equal to the case without an offset. Fig. 2.20(d) shows that the simulated optical loss increases contrary to what is expected theoretically. The p-n junction series resistance increased with the offset because the n-type dopant has lower resistance.

Chapter 3

Traveling-Wave Electrode Theory and Design

3.1 Introduction

THIS chapter presents concepts and simulations for the design of a traveling wave Mach-Zehnder modulator (TWMZM). The next section demonstrates the necessity of distributed analysis as opposed to simple lumped element modeling. Then, simulations are presented for coplanar strips transmission lines without p-n junctions using analytic equations and a finite element model (FEM) solver. Next, the effect of the p-n junction on the transmission lines is described. Afterwards, the electro-optic modulation response is described. The electrical and optical solutions are subsequently combined to determine the overall electro-optic performance of the device.

3.2 Distributed Analysis

A distributed element analysis, instead of a lumped element analysis, needs to be performed for long devices. A common rule of thumb is that the device length should be longer than approximately 1/10th of the wavelength in the guiding medium λ_g. Alternatively, transmission line effects need to be considered when the rise time of the signal is smaller or approximately equal to the time taken for a signal to reach the end of the device and back (one round-trip). The guide wavelength can be calculated from the effective relative permittivity of the guiding medium ϵ_{re} and the frequency of the wave f, as:

$$\lambda_g = \frac{c_o}{f\sqrt{\epsilon_{re}}} \tag{3.1}$$

Then, the critical length $L_{critical}$ determining when distributed analysis should be performed is:

$$L_{critical} = \frac{\lambda_g}{10} \tag{3.2}$$

The effective relative permittivity, for a medium composed of two different layers and no magnetic material, can be approximated by taking the average of the capacitances in the substrate and the superstrate. This simplifies to an arithmetic mean of the dielectric layers:

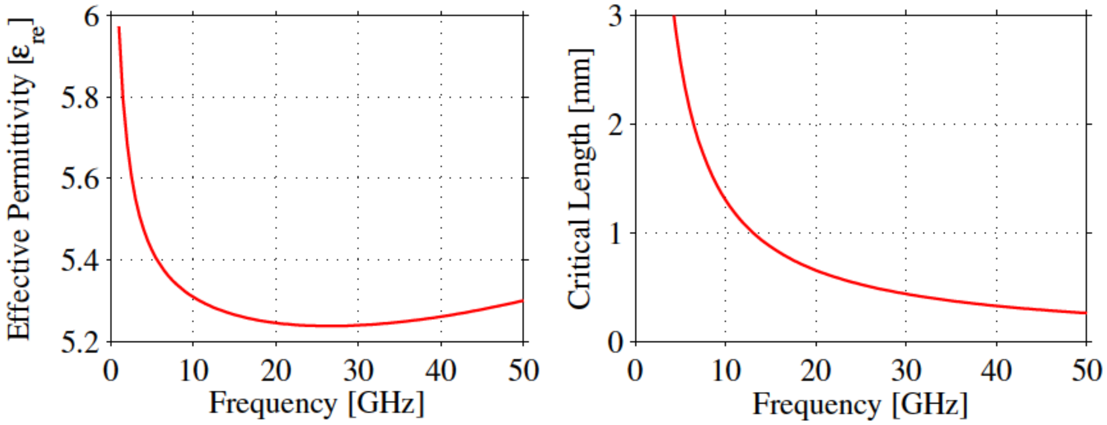
$$\epsilon_{re} = \frac{C_1 + C_2}{2C_1} = \frac{\epsilon_r + \epsilon_u}{2} \tag{3.3}$$

Where, ϵ_r is the relative permittivity of the substrate and ϵ_u is the relative permittivity of the upper medium. If it is assumed that the microwave in an SOI wafer travels half in air and half in the silicon substrate, then the permittivity can be approximated to:

$$\epsilon_{re} = \frac{\epsilon_{air} + \epsilon_{si}}{2} = \frac{1 + 11.7}{2} = 6.35 \tag{3.4}$$

It should be noted that this permittivity is a very rough estimate of the actual effective permittivity and a more accurate permittivity is obtained with a field solver. This approximate permittivity ignores dispersion and its dependence on the cross-section of the transmission line and the shape of the conductors. A FEM numerical simulation from a software such as ANSYS HFSS gives a more accurate effective permittivity. The simulated effective permittivity for an unloaded transmission line is plotted in Fig. 3.1(a) and the derived critical length is plotted in Fig. 3.1(b). At higher frequencies the permittivity is smaller than predicted by Eq. 3.4. It also varies little for frequencies above 9 GHz, indicating that the unloaded transmission line is more dispersive for lower frequencies.

In order to achieve a π-phase shift with low drive voltages, carrier depletion based modulators need to be relatively long and transmission line analysis must be performed. From Fig. 3.1(b), it is observed that transmission line analysis must be used for devices longer than 1 mm with frequencies above 12 GHz. With a capacitive load, such as the p-n junction in silicon modulators, the permittivity increases resulting in an even smaller

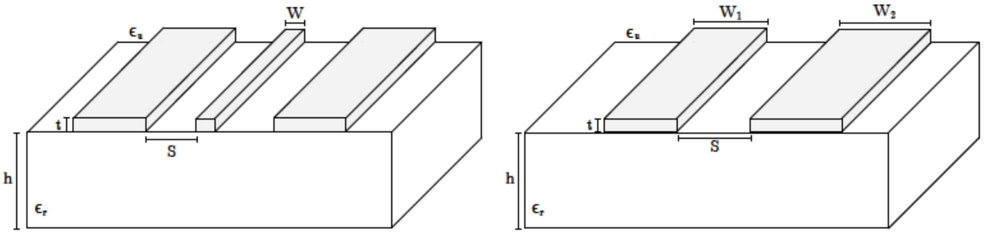


(a) (b)

Fig. 3.1 FEM simulation of (a) effective permittivity, and (b) critical length for a CPS electrode configuration.

guiding wavelength. This decreases the critical length and distributed analysis is necessary for shorter loaded transmission lines compared to unloaded transmission lines.

3.3 Coplanar Strips and Coplanar Waveguide



(a) (b)

Fig. 3.2 Schematic of a (a) coplanar waveguide, and (b) coplanar strips transmission lines.

For MZMs, the two most suitable transmission lines are the coplanar waveguide (CPW) and coplanar strips (also called coplanar striplines) (CPS) illustrated in Fig. 3.2. These transmission lines are planar, thus can be created in CMOS compatible processes. In addition, they both allow for lumped and/or distributed elements to be connected in shunt and in series to the transmission line. The drive voltage of silicon MZMs can be reduced if there is a long electro-optical interaction as discussed in Section 2.2. For this, the diode-waveguide structure appears along the transmission line (in shunt) and the modulator itself is the transmission line.

3.3.1 Coplanar Waveguide

In a CPW, the design parameters are the space S between the ground conductors and the signal conductor, and the width of the signal conductor W, as illustrated in Fig. 3.2(a). The ground conductors should ideally be infinite, or in reality as large as possible. An asymmetric CPW (ACPW) can also be created by using different widths between the ground and signal traces. Such an asymmetry can be used to increase the characteristic impedance of the transmission line [76].

The CPW is a transmission line with a fundamental transverse electro-magnetic (TEM) propagation mode. This TEM mode has no cut-off frequency. Waves propagating as TEM modes have their electric and magnetic fields in a plane perpendicular to the direction of propagation of the wave. When there are multiple different dielectric layers in a material system, e.g., oxide and silicon in an SOI substrate, a pure TEM mode cannot exist. A pure TEM mode cannot exist in a system with multiple dielectrics because a wave will have faster propagation in one medium relative to the other. This skews the plane containing the electric and magnetic fields towards the longitudinal direction. At low frequencies, about tens of GHz [77], the fundamental mode is approximately TEM and is referred to as quasi-TEM. At higher frequencies, if dispersion is strong, the mode can become TE like [77]. CPW can be used in a ground-signal-ground (GSG) for single drive modulators, or an unshielded signal-ground-signal (SGS) or GSGSG configurations for dual drive TWMZMs.

In a CPW, to reduce losses, air-bridges, i.e., open circuits for the signal trace, can be introduced to short out parasitic modes. Alternatively, the two ground traces can be connected together every $\lambda/4$ or less. More importantly, this must be performed if the two arms of the CPW have an asymmetrical shunt load, e.g., optical waveguide/p-n junctions in only one side of the GS CPW pair. Otherwise, dips at lower frequencies can manifest in the forward gain response [22].

3.3.2 Coplanar Strips

In CPS, the design parameters are the distance between the two conductors S, and the widths of the conductors W_1 and W_2, as shown in Fig. 3.2(b).

CPS have become one of the most commonly used transmission line for planar optical modulators in various technologies such silicon, InP, and GaAs. A CPS without a ground plane has parasitic dielectric slab waveguide modes, mainly the TE0 and TM0 mode [78]. This causes detrimental effects because the fundamental (TEM) mode strongly couples to the parasitic TE0 mode [78]. Compared to CPW, CPS are believed to operate with the quasi-TEM mode for frequencies higher than CPW transmission lines.

Aside from the geometry, the fundamental difference between a CPW and CPS is that the CPS is usually used as a differential (balanced) transmission line, whereas a CPW is an unbalanced transmission line. In a CPW, the potential reference of the signal trace is the grounds adjacent to it. When the CPS is used as a balanced line with the GSSG configuration, i.e., differential signaling the potential is with respect to the two signals. The difference of the two signals is unaffected by interference common to the two signal traces and this makes the CPS inherently immune to unwanted interference. However, in integrated series push-pull photonic modulator designs, the CPS are not operated as a balanced line because the SPP modulators are single-ended, i.e., only a GS or SG configuration rather than a GSSG configuration.

Furthermore, the CPW structure has a signal trace placed in between two ground traces. The ground traces act as a shield that reduces the external interference from affecting the signal and also prevents the signal from introducing interference to other nearby components. A CPS has higher radiation losses as the signal trace is not shielded by a ground trace on one side. Therefore, CPS designs need to maintain a reasonable space between devices to avoid cross-talk [79].

Typically, fan-out boards use CPW or microstrip transmission lines at the connector interface. Transitions from a CPW-to-CPS (baluns/unbals) can be created to interface the CPW transmission line on the circuit board to the on-chip CPS modulator.

3.3.3 Analytical Equations for CPW and CPS Transmission Lines

Analytically, transmission lines can be analyzed by two main methods, static (e.g., quasi-static), and dynamic (e.g., full-wave analysis). The dynamic analysis is applicable to all modes and range of frequencies, thus it computes the characteristic impedance and a propagation constant that is dependent on frequency [80]. The static or quasi-static approach is much simpler to calculate but it is theoretically only valid for DC and TEM modes. Despite this, it has been used to design very high speed transmission lines with success. Equations for the quasi-static analysis using a conformal mapping technique which maps the capacitance of the transmission line structure to parallel plate capacitances are available in various textbooks.

Several models have been developed of which some consider the finite thickness of the substrate and non-zero thickness of the metal strips. Qualitatively, a wider trace width decreases the characteristic impedance and a wider spacing from ground increases the characteristic impedance.

An analytical model for CPS is provided in [76]. In this model, the thickness of the metal strips t is included by defining an effective gap S_e and strip width W_e from a parameter Δ. The substrate is assumed to be infinitely thick. This formulation for the symmetric CPS ($W = W_1 = W_2$) is reproduced here:

$$\Delta = \frac{1.25t}{\pi}\left[1 + ln\left(\frac{4\pi W}{t}\right)\right] \tag{3.5}$$

$$S_e = S - \Delta \tag{3.6}$$

$$W_e = W + \Delta \tag{3.7}$$

The aspect ratio k_1 of the gap and strip widths and its complement k_1' are defined as:

$$k_1 = \frac{a}{b} = \frac{S}{S + 2W} \tag{3.8}$$

$$k_1' = \sqrt{1 - k_1^2} \tag{3.9}$$

The effective aspect ratio k_e is:

$$k_e = \frac{S_e}{S_e + 2W_e} \tag{3.10}$$

Its complement is:

$$k_e^{'} = \sqrt{1 - k_e^2} \tag{3.11}$$

The effective dielectric constant $\epsilon_{re}^{'}$ is obtained by modifying the capacitance term due to the metallization thickness, this is valid for $\epsilon_r \geq 9$ and $t/W < 0.1$:

$$\epsilon_{re}^{'} = \epsilon_{re} - \frac{1.4(\epsilon_{re} - 1)\frac{t}{S}}{|\frac{K(k_1)}{K(k_1^{'})}| + 1.4\frac{t}{S}} \tag{3.12}$$

The function $K(\blacksquare)$ is the complete elliptic integral of the first kind, and its argument is the elliptic modulus. The function can be evaluated as a series. Finally, the characteristic impedance for the CPS is calculated as:

$$Z_{0cs} = \frac{120\pi}{\sqrt{\epsilon_{re}^{'}}} \frac{K(k_e)}{K'(k_e)} \tag{3.13}$$

Another analytical model, presented in [76], is for an asymmetric CPS with a finite substrate thickness. The effect of the finite substrate thickness is considered by a modified expression for the effective relative permittivity. The solution for the characteristic impedance of an asymmetric CPS with a finite substrate thickness (Z_{0csf}) can be obtained by solving the following equations:

$$\epsilon_{re} = 1 + \frac{\epsilon_r - 1}{2} \frac{K(k_4)}{K'(k_4)} \frac{K^{'}(k_3)}{K(k_3)} \tag{3.14}$$

$$k_3 = \sqrt{\frac{W_1}{W_1 + S} \frac{W_2}{W_2 + S}} \tag{3.15}$$

$$k_4 = \sqrt{\frac{sinh(\frac{\pi W_1}{2h})}{sinh(\frac{\pi(W_1+S)}{2h})} \frac{sinh(\frac{\pi W_2}{2h})}{sinh(\frac{\pi(W_2+S)}{2h})}} \tag{3.16}$$

$$Z_{0csf} = \frac{60\pi}{\sqrt{\epsilon_{re}}} \frac{K^{'}(k_3)}{K(k_3)} \tag{3.17}$$

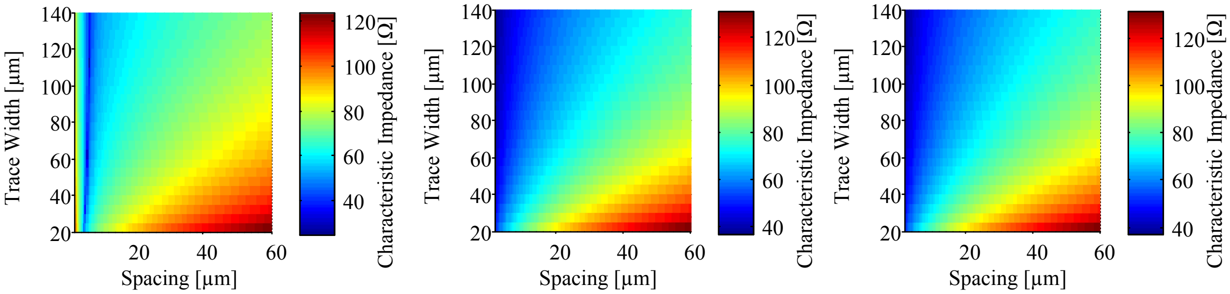
The final analytical expressions for cases with an infinite substrate, infinitesimal metal thickness, and air superstrate for symmetric CPW, CPS, and an ACPS are listed in Table 3.1.

Table 3.1 Analytical expressions for different types of CPW and CPS electrodes.

	Effective Permittivity	Characteristic Impedance	
CPW	$\epsilon_{re} = \frac{\epsilon_r + 1}{2}$	$Z_{0cp} = \frac{30\pi}{\epsilon_{re}} \frac{K'(k_1)}{K(k_1)}$	$k_1 = \frac{S}{S+2W}$
CPS	$\epsilon_{re} = \frac{\epsilon_r + 1}{2}$	$Z_{0cs} = \frac{120\pi}{\epsilon_{re}} \frac{K(k_1)}{K'(k_1)}$	$k_1 = \frac{S}{S+2W}$
ACPS	$\epsilon_{re} = \frac{\epsilon_r + 1}{2}$	$Z_{0acs} = \frac{60\pi}{\epsilon_{re}} \frac{K(k_2)}{K'(k_2)}$	$k_2' = \sqrt{\frac{W_1}{W_1+S} \frac{W_2}{W_2+S}}$

3.3.4 Unloaded SOI Transmission Line

This section begins by using the equations presented in Section 3.3.3 for the calculation of the characteristic impedance. The effective dielectric constant used is 6.35 as calculated in Section 3.2. It is to be noted that this approximation ignores dispersion of the microwave, the effect of the cross section of the transmission line, shape of the conductors, and is generally only valid for low frequencies.



(a) (b) (c)

Fig. 3.3 Analytical characteristic impedance of a coplanar strip transmission line for varying trace widths and spacing for the case of (a) 2 µm metal lines and an infinitely thick substrate, (b) infinitesimal metal line and an infinitely thick substrate, and (c) infinitesimal metal line and a finite 700 µm thick substrate.

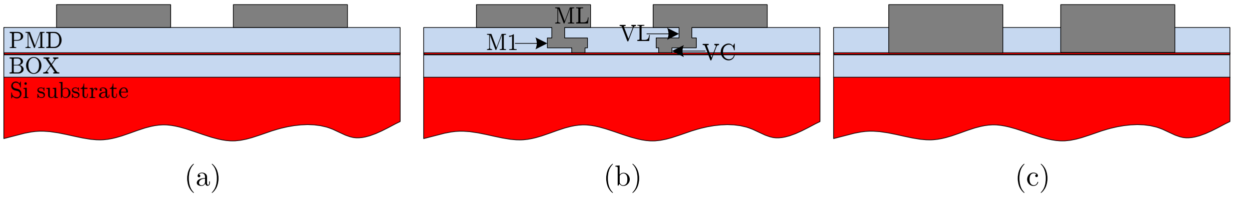
Fig. 3.3 illustrates the effect of changing the ground and signal trace widths and the spacing between them on the characteristic impedance of the CPS transmission lines. Fig. 3.3(a) calculated this dependence for the case with 2 µm thick metal lines and an infinite substrate using Eqns. 3.5-3.13. Fig. 3.3(b) calculated this dependence for the case with infinitely thin metal lines and an infinite substrate using the equations for the CPS show in Table 3.1. Fig. 3.3(c) calculated this dependence for the case with infinitely thin metal lines and a

700 µm finite substrate using Eqns. 3.14-3.17.

It is observed that the characteristic impedance increases as the spacing increases and the trace widths decrease. This agrees with the qualitative prediction. In addition, comparing Fig. 3.3(b) and Fig. 3.3(c), there are only slight differences between the case with the infinitely thick substrate and the one with the finite substrate, indicating that according to the analytical model, a 700 µm substrate is thick enough to be treated as an infinite substrate.

Fig. 3.3(a) shows the effect of having a finite metal thickness. For a regime with spacing less than about 5 µm, the characteristic impedance is much higher than the case with infinitesimally thin metal layers of Fig. 3.3(b) and Fig. 3.3(c). However, for spacing larger than 5 µm, the characteristic impedance behaves similar to the case with no metal thickness.

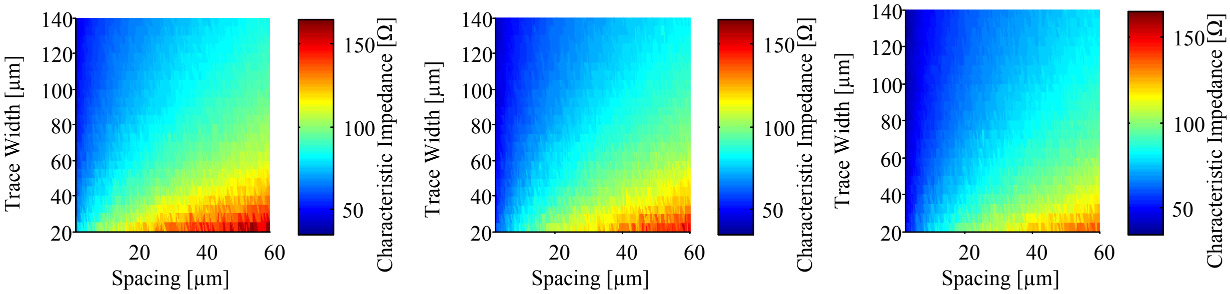
Repeating the simulation with a 3-D full-wave electromagnetic field solver, for instance ANSYS HFSS, would consider effects that are ignored by the analytical model such as the dielectric loss, fringe fields, radiation loss, frequency dependence, and non TEM modes. Since the definitions of microwave propagation constant and the characteristic impedance assume uniform transmission lines, HFSS only solves these transmission line quantities based on the cross-section at the interface of the simulation ports. The following will present simulations performed at 30 GHz using ANSYS HFSS.



(a) (b) (c)

Fig. 3.4 Cross-sections of simulated CPS transmission lines. Three cases were considered: (a) CPS with last metal layer only, (b) CPS with entire metal stack including the first metal, last via, and silicon contact vias, and (c) CPS with entire metal stack of the same width as the top metal layer.

Fig. 3.4 shows the cross-section at the ports of the simulation model used in HFSS. To study the effects of the vias and the first metal layer (M1), three separate studies were done. In all the studies only the width of the last metal layer (ML) and the spacing between them was varied. The simulated structure consisted of 700 µm thick substrate with a resistivity of 750 Ω-cm, a 2 µm buried-oxide-layer (BOX), a 220 nm silicon top layer, and oxide as the pre-metal dielectric (PMD). The top metal layer was 2 µm thick. The other dimensions and material properties mirrored the one used in the fabrication process and the wafer layers. The first case, illustrated in Fig. 3.4(a), considered only the top metal layer (ML). In the second case, depicted by Fig. 3.4(b), the whole metal stack was modeled. The location of the last via (VL), first metal, silicon contact via (VC) were changed to maintain contact with the top metal as the spacing between the top metal electrodes was changed. The widths of the last via and first metal stack remained constant. The third case, diagrammed in Fig. 3.4(c), also included the whole metal stack, but the widths of all the metal layers and vias were made equal to obtain a larger electrode cross-section.

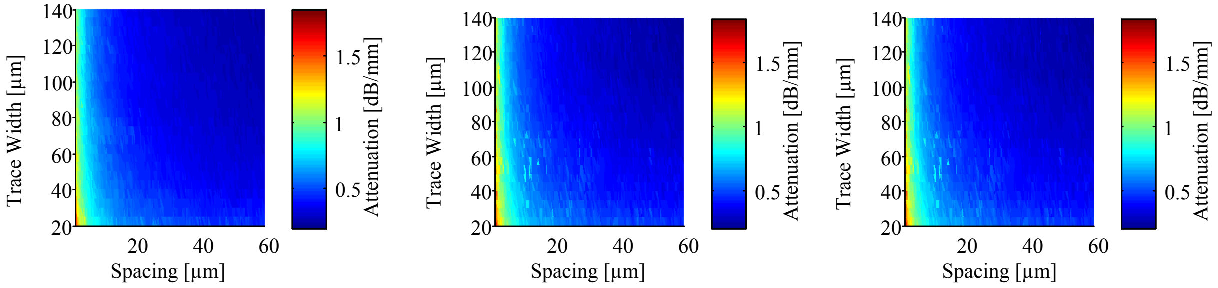


(a) (b) (c)

Fig. 3.5 FEM simulation of coplanar strips characteristic impedance at 30 GHz for three cases: (a) last metal only, (b) entire metal stack, and (c) entire metal stack of equal widths.

Fig. 3.5 shows the characteristic impedance from HFSS simulations for these three cases. The impedance for the case with the ML only (case (a)) is higher than predicted by the analytical equations (Fig. 3.3). The trend, however, is the same: the impedance decreases with increasing trace width and increases with electrode spacing. In the cases with the full metal stack, the characteristic impedance is smaller, and also slightly smaller than the analytical prediction. When comparing the cases with the full metal stack, it is observed that the characteristic impedance is slightly higher for case (b).

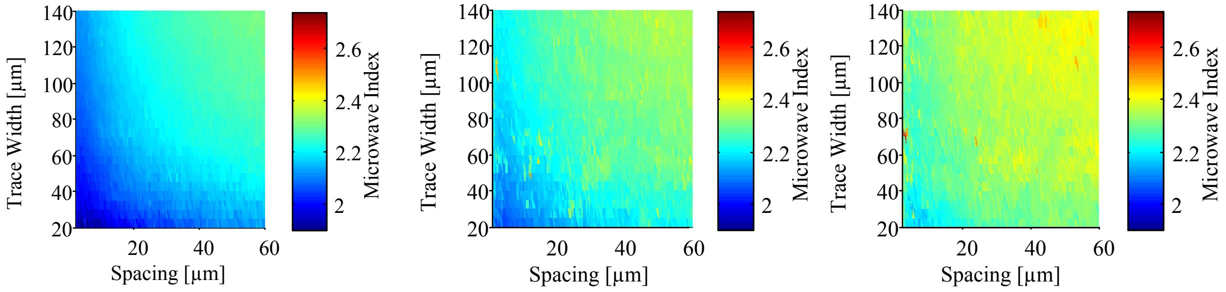
Fig. 3.6 shows the propagation attenuation from HFSS simulations. In the case with the top metal layer only, the attenuation decreases with increasing trace width and also decreases with increasing spacing. It is observed that the attenuation dominates for smaller spacing regardless of the trace width. Thus, it is important to maintain at least a spacing of approximately 10 µm to obtain low losses. In addition, simulations show that the attenuation does not change much with larger cross-sectional area. This is explained by assuming that the conductor loss is dominant. Then at 30 GHz, because of the skin effect only a



(a) (b) (c)

Fig. 3.6 FEM simulation of coplanar strips attenuation at 30 GHz for three cases: (a) last metal only, (b) entire metal stack, and (c) entire metal stack of equal widths.

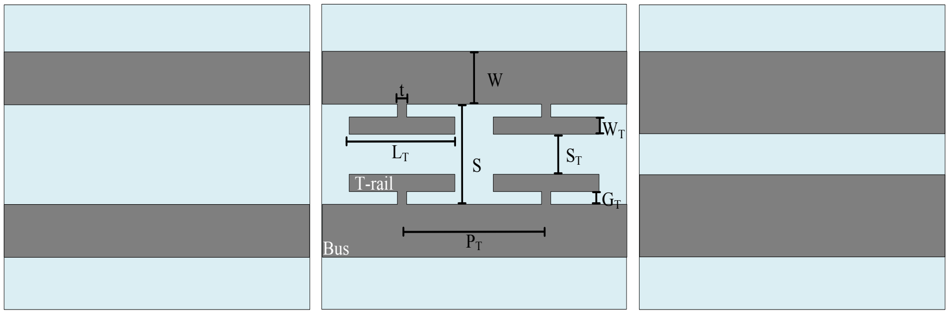
small portion of the outer cross-section of the electrode is used for conduction. Beyond a certain point, additional area does not contribute in lowering the attenuation. Considering the characteristic impedance simulations of Fig. 3.5, when attempting to increase the characteristic impedance, it is more advantageous to increase the electrode spacing as this will lower loss. On the other hand, when reducing the characteristic impedance, it is better to increase the trace width. When the entire metal stack is simulated, the attenuation is found to be slightly larger at small spacing values.



(a) (b) (c)

Fig. 3.7 FEM simulation of coplanar strips effective index at 30 GHz for three cases: (a) last metal only, (b) entire metal stack, and (c) entire metal stack of equal widths.

Fig. 3.7 shows the effective index obtained from HFSS simulations for the three different cases. The simulations indicate that increasing the spacing and/or the trace width generally increases the effective index for all three cases. The effective index is the lowest for the case with the top metal layer only with an approximate range of 2.0-2.3. For case (b), the effective index varies between 2.1-2.3. The effective index is the highest for case (c) ranging from 2.2-2.4. Furthermore, the effective index in case (c) varies less with respect to the trace spacing and width compared to the other two cases.



(a) (b) (c)

Fig. 3.8 Different CPS electrode designs: (a) narrow CPS, (b) slow-wave CPS, and (c) wider CPS transmission lines.

In the literature, CPS structures with 'T' rails (or slow-wave electrodes) have been studied to increase the effective index of the transmission line [26, 81, 82, 83]. To study the effect of 'T'-rails on CPS transmission lines, three more cases of electrode structures were simulated. These structures are shown in Fig. 3.8. The first case is of a regular CPS electrodes with width and spacing matching that of the bus of the 'T'-rails transmission line. The third case is a wider CPS transmission line with the spacing equal to S_T and the width equal to $W + W_T + G_T$. For the following simulations, only the top metal layer was modeled. In order to extract the microwave attenuation, microwave index, and characteristic impedance, the S-parameters simulated from HFSS were converted to ABCD parameters using the following equations [84]:

$$A = \frac{\left(Z_{01}^{*} + S_{11}Z_{01}\right)\left(1 - S_{22}\right) + S_{12}S_{21}Z_{01}}{2S_{21}\sqrt{R_{01}R_{02}}} \tag{3.18}$$

$$B = \frac{\left(Z_{01}^{*} + S_{11}Z_{01}\right)\left(Z_{02}^{*} + S_{22}Z_{02}\right) - S_{12}S_{21}Z_{01}Z_{02}}{2S_{21}\sqrt{R_{01}R_{02}}} \tag{3.19}$$

$$C = \frac{\left(1 - S_{11}\right)\left(1 - S_{22}\right) - S_{12}S_{21}}{2S_{21}\sqrt{R_{01}R_{02}}} \tag{3.20}$$

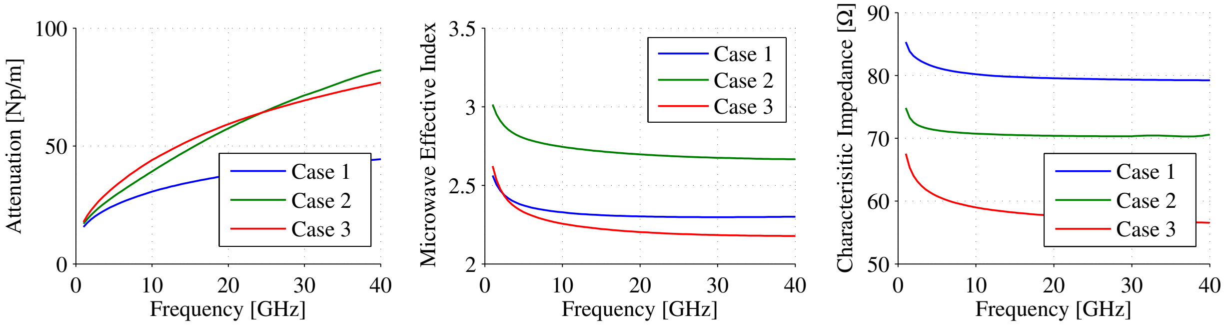
$$D = \frac{(1 - S_{11})\left(Z_{02}^{*} + S_{22} Z_{02}\right) + S_{12} S_{21} Z_{02}}{2 S_{21} \sqrt{R_{01} R_{02}}} \tag{3.21}$$

Where, S_{ij} are the two-port S-parameters, R_{0i} and Z_{0i} are respectively the normalizing resistance and impedance at the i^{th} port (typically 50 Ω), and * denotes the complex conjugate operation. After obtaining the ABCD parameters of the transmission line, the propagation constant and the characteristic impedance were calculated as:

$$\gamma_{ul} = \alpha_{ul} + j\beta_{ul} = \frac{cosh^{-1}(A)}{l} \tag{3.22}$$

$$Z_{0ul} = \sqrt{\frac{B}{C}} \tag{3.23}$$

Where, l is the length of the transmission line, γ_{ul} is the propagation constant of the unloaded line, α_{ul} is the attenuation coefficient in $[Np/m]$, β_{ul} is the phase constant in $[rad/m]$, and Z_{0ul} is the characteristic impedance of the unloaded transmission line.



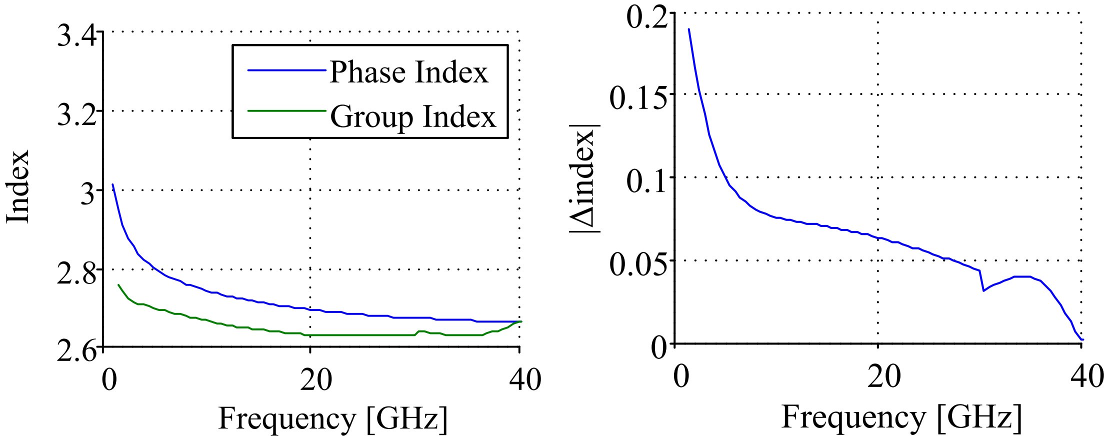
(a) (b) (c)

Fig. 3.9 Simulated (a) microwave attenuation, (b) microwave index, and (c) characteristic impedance of CPS transmission lines

The microwave attenuation for the three cases is shown in Fig. 3.9(a). Simulations show that the slow-wave CPS transmission line has a similar attenuation to the case of the wider CPS transmission line with a small spacing. Both these cases have a larger attenuation than the narrower CPS transmission line with a larger spacing, thus agreeing with the earlier conclusion that a larger spacing reduces the loss of unloaded electrodes. However, it should be noted that with the p-n junction loading in a SPP configuration, having a large spacing does not necessarily lead to lower loss, as will be explained later in Section 3.4.

For the microwave index, plotted in Fig. 3.9(b), the slow-wave CPS transmission line indeed has a higher index than the other two cases. The index is also relatively constant for high frequencies, therefore dispersion in this SOI technology for the unloaded transmission lines is more significant at relatively low frequencies. When compared to the large parameter sweep of Fig. 3.7, it is also evident that the slow-wave design has a higher index than possible with any trace width or spacing of regular CPS lines.

The characteristic impedance of the slow-wave transmission line is approximately the average of the characteristic impedance of the other two cases, as shown by the simulation results in Fig. 3.9(c).



(a) (b)

Fig. 3.10 The calculated group and phase index of the slow-wave CPS design is shown in (a), and in (b) the absolute value of their difference is plotted as a function of frequency.

For velocity matching, as discussed further in Section 3.5, it is the group velocity of the microwave and optical wave that needs to be matched. Ideally, a TEM microwave has no dispersion and the group index equals the phase index. This is often the assumption made in TWMZM designs. However, as the simulations indicate, there is relatively strong dispersion at lower frequencies. In Fig. 3.10(a), the calculated phase index and the group index is plotted. The group index is calculated from the phase index as: $n_{\mu,g} = n_{\mu,ul} + \omega\frac{dn_{\mu,ul}}{d\omega}$. Their absolute difference is plotted in Fig. 3.10(b), and it is observed that beyond 10 GHz, the difference between the two indices is less than 0.08. The difference is higher for lower frequencies, as expected. For simplicity, it will be assumed through out the thesis that the microwave group and phase index are approximately the same.

Telegraphist Transmission Line Model

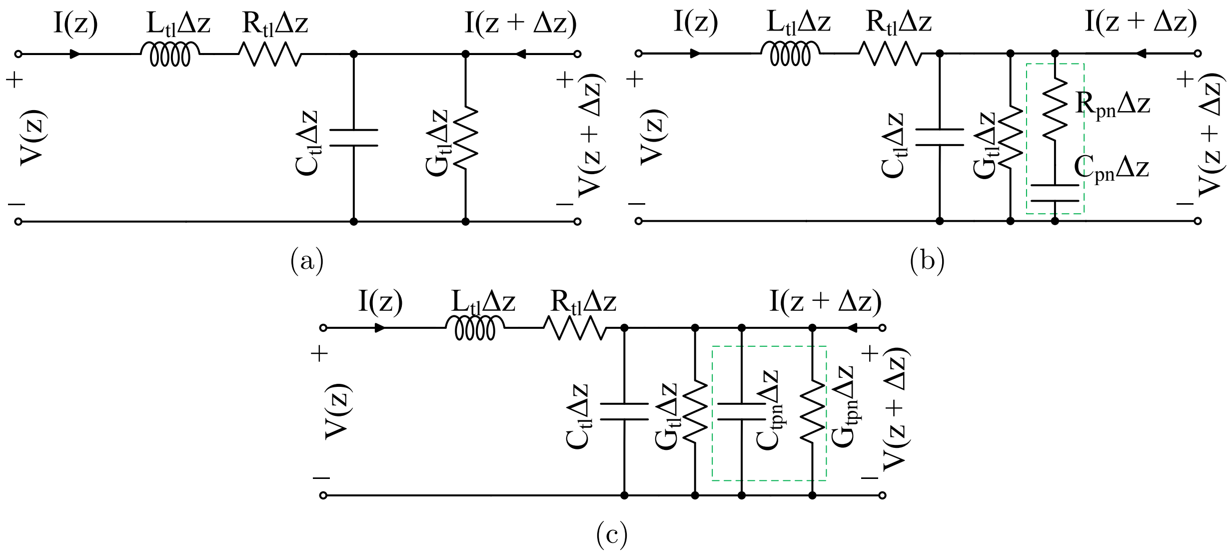


Fig. 3.11 Equivalent RLCG model of a transmission line that is (a) unloaded, (b) loaded with a p-n junction, and (c) loaded with a p-n junction transformed into a parallel circuit.

From basic transmission line and electromagnetic wave theory, a lossy transmission line can be represented by a lumped element model as shown in Fig. 3.11(a). This model is applicable for TEM modes and is a good approximation for quasi-TEM modes. The R $[\Omega/m]$, G $[S/m]$, L $[H/m]$, and C $[F/m]$ are the resistance, conductance, inductance, and capacitance per unit length, respectively. In the equivalent model, the capacitance is implied from the electric field established by a potential difference between the two conductors of the transmission line. The inductance represents the creation of a magnetic field due to the current flowing on the line. The resistance is the loss in the conductors. The conductance models the loss in the dielectric between the conductors. All components have values per unit length, and Δz represents a length of the transmission line that is much smaller than the effective wavelength of the propagating wave to allow for the lumped element analysis ($\Delta z \ll \lambda_{eff}$). It should also be noted that the RLGC parameters are also frequency dependent, in particular the resistance and the conductance. The inductance and capacitance typically become constant at high frequencies.

For an unloaded lossy transmission line (here, unloaded is used to refer to a transmission line without the p-n junction), the microwave propagation constant is:

$$\gamma_{ul} = \alpha_{ul} + j\beta_{ul} = \sqrt{(R_{tl} + j\omega L_{tl})(G_{tl} + j\omega C_{tl})} \tag{3.24}$$

Where, α_{ul} $[Np/m]$ is the attenuation factor, β_{ul} $[rad/m]$ is the phase factor, and ω is the angular frequency.

The characteristic impedance, which is the ratio of the forward moving voltage and the backward propagating current waves at each position of the line, is found from:

$$Z_{0ul} = \sqrt{\frac{R_{tl} + j\omega L_{tl}}{G_{tl} + j\omega C_{tl}}} \tag{3.25}$$

Due to the complex variables in the square root of the propagation constant, it is difficult to obtain a closed-form expression for α_{ul} and β_{ul} in terms of the RLGC parameters. Simplified expression can be obtained by considering special cases, which can be divided into a lossless or lossy regime of the transmission line. Lossy transmission lines can be further divided into different frequency regimes or a low-loss case.

For the lossless case, the lossy elements are ignored ($R_{tl} = G_{tl} = 0$). The RLGC model simplifies to a model with only L_{tl} and C_{tl}, and $\alpha_{ul} = 0$. The expression for the characteristic impedance and propagation constant for lossless lines are listed in column 2 of Table 3.2.

At very low frequencies, $\omega \approx 0$, the RLGC model simplifies to a model with only R_{tl} and G_{tl}, and $\beta_{ul} = 0$. The expression for the characteristic impedance and propagation constant for lossy lines at low frequencies are listed in column 3 of Table 3.2.

Under low-loss conditions and at very high frequencies, the following three inequalities are satisfied: (1) $R_{tl} \ll \omega L_{tl}$, (2) $G_{tl} \ll \omega C_{tl}$, and (3) $R_{tl}G_{tl} \ll \omega^2 L_{tl}C_{tl}$. Low-loss conditions are assumed to be met by all practical transmission lines as otherwise the transmission line would be useless. With these conditions, the expression for the characteristic impedance and propagation constant can be simplified. The characteristic impedance simplifies to the lossless case:

$$Z_{0ul} = \sqrt{\frac{L_{tl}}{C_{tl}}} \tag{3.26}$$

This characteristic impedance can be substituted into the simplified expression for the

Table 3.2 Propagation constant and characteristic impedance of lines in different regimes.

		Lossy		
	Loss-less	Low Frequency	Intermediate Frequency	High Frequency
Characteristic Impedance (Z_{0ul})	$\sqrt{\frac{L}{C}}$	$\sqrt{\frac{R}{G}}$	$\frac{1-j}{\sqrt{2}}\sqrt{\frac{R}{\omega C}}$	$\sqrt{\frac{L}{C}}$
Propagation Constant (γ_{ul})	$j\omega\sqrt{LC}$	$\sqrt{RG}$	$\frac{1+j}{\sqrt{2}}\sqrt{\omega CR}$	$\alpha_c(f) + \alpha_d(f) + j\omega\sqrt{LC}$

low-loss propagation constant to obtain:

$$\gamma_{ul} = \overbrace{\frac{1}{2}\left(\underbrace{\frac{R_{tl}}{Z_{0ul}}}_{\text{conductor loss}} + \underbrace{G_{tl}Z_{0ul}}_{\text{dielectric loss}}\right)}^{\alpha_{ul}\ [Np/m]} + j\,\overbrace{\omega\sqrt{L_{tl}C_{tl}}}^{\beta_{ul}\ [rad/m]} \tag{3.27}$$

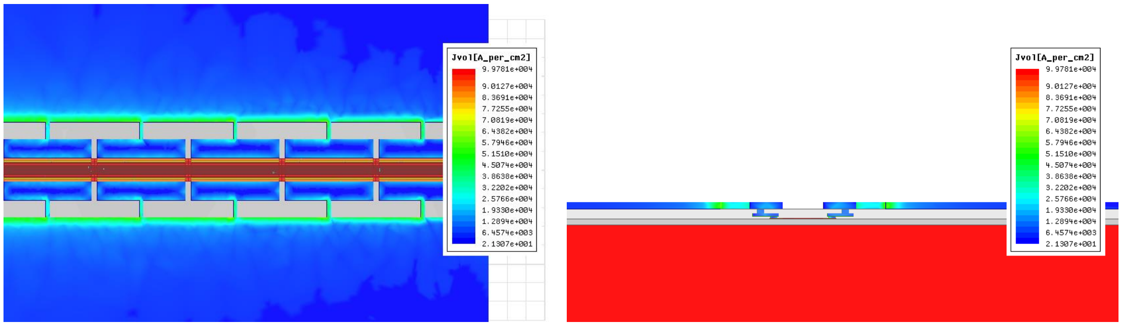
The microwave phase index for the unloaded transmission line in the lossless and low-loss (high-frequency) regimes is given by:

$$n_{\mu,ul} = \frac{c_0}{v_{\mu,ul}} = c_0\sqrt{L_{tl}C_{tl}} = c_0 C_{tl} Z_{0ul} = c_0 \frac{L_{tl}}{Z_{0ul}} \tag{3.28}$$

The RLGC components in the telegraphist model are also frequency dependent. The resistance of the transmission line R_{tl} changes with frequency due to the skin effect. At DC, the current is distributed uniformly over the entire cross-section of the conductor. With the application of an AC current, a changing magnetic field is created inside the conductor. This changing magnetic field induces a current that opposes the driving current. Since the magnetic field is strongest at the center of the conductor, the current is most canceled at the center and the effective current density is highest at the extremities of the conductor. An alternative interpretation is based on the path of least impedance. As frequencies increase, so does the impedance from the inductance of the line. The self-inductance is largest at the center of a conductor than at the edges. Since current prefers the path of least impedance, current density will be greater at the edge of the conductor. The skin depth (δ) is defined as the depth where current density drops to $1/e$ (63%) of the value at the conductor's surface. The expression for skin depth in good conductors is:

$$\delta = \sqrt{\frac{1}{\sigma\pi\mu_0\mu_r f}} \tag{3.29}$$

Where, σ is the conductivity of the conductor, μ_0 and μ_r are respectively the free-space and relative magnetic permeability, and f is the frequency. Almost all of the current will be within five times the skin depth, indicating that for high frequencies, conductors with a larger area will have little effect in reducing the conductor loss. This interpretation also agrees with the simulated attenuation shown in Fig. 3.6, where little changes were seen in attenuation (in units of dB/mm) for various widths and spacing of CPS lines. Furthermore, in transmission lines, the current is mainly located at the edges facing the signal and reference conductors. This is corroborated by the FEM simulations of the slow-wave electrode shown in Fig. 3.12. Since, conductor loss is also the highest where the current is the densest, using larger electrode widths is not beneficial for reducing loss at these high frequencies.



(a) (b)

Fig. 3.12 Simulation of current volume distribution in the electrodes and vias at 30 GHz, (a) top view, and (b) side view. The current is denser at the edges that face each conductor. It is mostly concentrated on the edge of the bus electrodes and the bridge of the 'T' shaped extensions.

The total resistance is the sum of the resistance in the signal and return paths, which depend on the cross-section of the transmission line. Therefore, the actual values of the DC (R_{DC}) and AC (R_{AC}) resistance should be obtained from a field solver. Because of the skin effect, it is observed that the AC resistance increases with the square root of frequency as:

$$R_{tl} = R_{DC} + R_{AC}\sqrt{f} \tag{3.30}$$

Using a conductor with a higher conductivity will result in a smaller skin depth and a lower resistance. For example, a round wire at 1 GHz made of aluminum would have approximately 25% more resistance (thus, more loss) than a wire made with copper.

Using Eq. 3.27 and 3.30, the conductor loss for the unloaded transmission line assuming constant C_{tl} and L_{tl} can be expressed as:

$$\alpha_{cul} = \frac{1}{2}\left(\frac{R_{DC} + R_{AC}\sqrt{f}}{Z_{0ul}}\right) = \frac{1}{2}\left(R_{DC} + R_{AC}\sqrt{f}\right)\sqrt{\frac{C_{tl}}{L_{tl}}} \tag{3.31}$$

Since the current distribution changes due to the skin effect, the inductance will also change. Sources of inductance in a transmission line are the self (internal) inductance of the signal and return traces, and the inductance external to the conductors from magnetic fields in the dielectrics. The external inductance is proportional to the total current flowing in the conductors and is frequency independent. At higher frequencies, because of the skin effect, the current accumulates at the edges of the conductor and this reduces the the self-inductance of the conductors. Thus, the frequency dependence of the inductance can be expressed as:

$$L_{tl} = L_{\infty} + \frac{R_{AC}}{2\pi\sqrt{f}} \tag{3.32}$$

Where, L_{∞} is the external inductance. The external inductance dominates at high frequencies.

The loss through the dielectric C_{tl} is modeled by introducing the conductance G_{tl}. In the lossless model, G_{tl} is ignored. The dielectric loss is a material property and is represented by the imaginary part $\epsilon^{''}$ of the dielectric constant:

$$\epsilon(\omega) = \epsilon'(\omega) - j\epsilon''(\omega) \tag{3.33}$$

For lossy mediums, the dielectric loss can also be known from its conductivity as:

$$\epsilon'' = \frac{\sigma}{\omega\epsilon_0} \tag{3.34}$$

The loss tangent is the ratio of the imaginary and real parts of the complex permittivity:

$$tan(\delta) = \frac{\epsilon''}{\epsilon'} \tag{3.35}$$

The current ($I = C\frac{dV}{dt}$) through a capacitor with a dielectric that has a complex permittivity can be broken into a real and imaginary part. The real part represents the resistive loss and is modeled as G_{tl}, while the imaginary part is simply the lossless C_{tl} from displacement currents. The conductance can then be expressed as:

$$G_{tl} = G_0 + 2\pi f C_{tl} tan(\delta) \tag{3.36}$$

Where, G_0 is the DC dielectric loss. Thus, from Eq. 3.36 and Eq. 3.27, the dielectric loss for the unloaded transmission line can be written as:

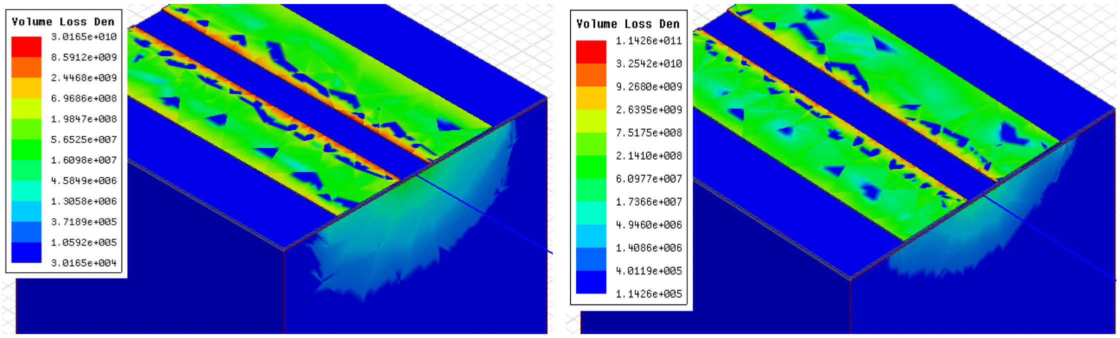
$$\alpha_{dul} = \frac{1}{2}\left[G_0 + 2\pi f C_{tl} tan(\delta)\right] Z_{0ul} = \frac{1}{2}\left[G_0 Z_{0ul} + \frac{2\pi tan(\delta)\sqrt{\epsilon_{eff}}}{c_0} f\right] \tag{3.37}$$

The dielectric loss is mostly independent of the geometry. However, the effective index changes with geometry of the transmission line and this will affect the dielectric loss. Typically, the loss tangent is nearly constant with frequency. The overall dielectric loss, however, is linearly dependent on frequency.

In SOI, the substrate and dielectric loss are much smaller than compared to bulk CMOS technologies. Losses can also be lowered with high resistive silicon substrates. The effect of the substrate loss is also much more significant in microstrip transmission line as opposed to the coplanar transmission lines. For microstrips, often the ground conductor is below the substrate and fields cross the lossy substrate to terminate on that ground conductor. Simulations from HFSS for unloaded CPS electrodes in a 750 Ω-cm silicon substrates is shown in Fig. 3.13 for 10 GHz and 30 GHz, respectively. It is observed that the loss in the conductors, especially at the edges, can be significantly higher than the loss in the substrate. The conductor loss increases with higher frequencies and the contribution of the substrate loss to the total loss decreases.

In silicon modulators, as will be shown in the next section, the high resistance of the p-n junction from low doping will introduce additional dielectric loss.

The total loss of the unloaded transmission line is given by the sum of the conductor



(a) (b)

Fig. 3.13 Simulated volume loss density (log scale) of the substrate and conductor at (a) 10 GHz and (b) 30 GHz.

loss (Eq. 3.31) and the dielectric loss (Eq. 3.37): $\alpha_{ul} = \alpha_{cul} + \alpha_{dul}$.

The RLGC parameters for the HFSS simulations of the CPS structures shown in Fig. 3.8 were obtained by applying the following equations to the characteristic impedance Z_{0ul} and propagation constant γ_{ul} extracted from the simulated S-parameters.

$$R_{tl} = \Re\left\{Z_{0ul}\gamma_{ul}\right\} \tag{3.38}$$

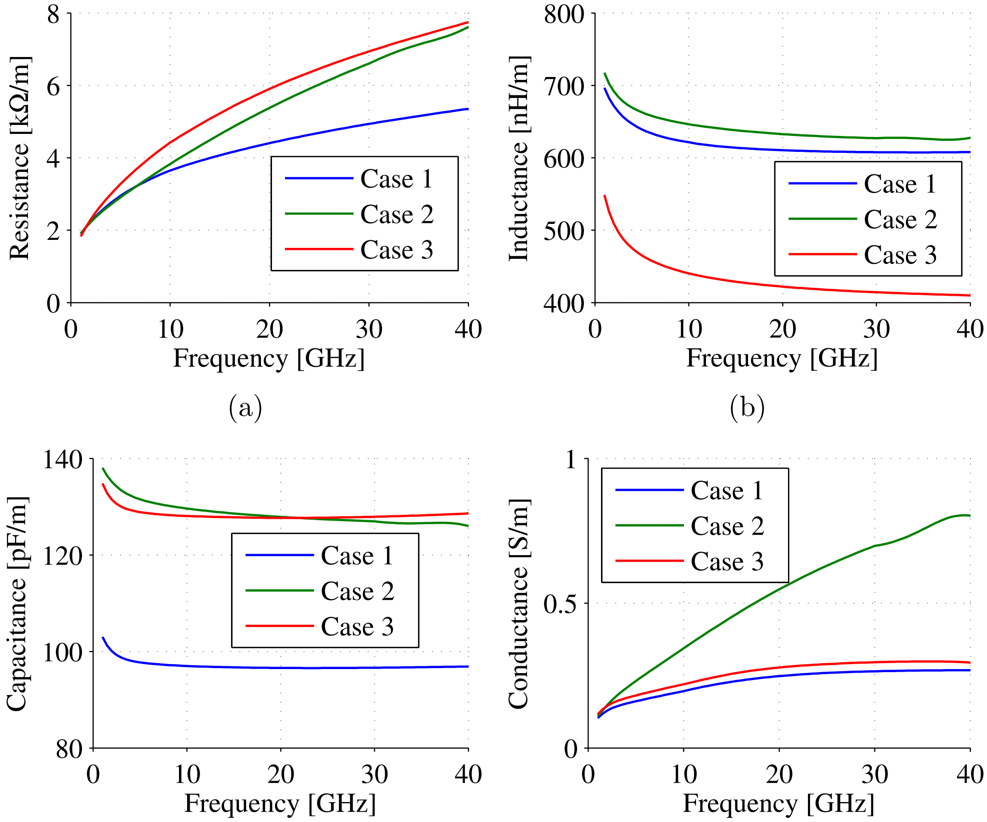
$$L_{tl} = \frac{\Im\left\{Z_{0ul}\gamma_{ul}\right\}}{\omega} \tag{3.39}$$

$$G_{tl} = \Re\left\{\frac{Z_{0ul}}{\gamma_{ul}}\right\} \tag{3.40}$$

$$C_{tl} = \frac{\Im\left\{\frac{Z_{0ul}}{\gamma_{ul}}\right\}}{\omega} \tag{3.41}$$

The RLGC parameters for the three cases are shown in Fig. 3.14. These simulation results indicate the the CPS transmission line with ‘T’-rails has an inductance that is similar to the narrower CPS transmission line, but at the same time, it has the high capacitance of the CPS transmission line with the smaller spacing between the electrodes. Since this structure has the high capacitance and the high inductance of the other two structures, it has a higher microwave index and allows for a slower propagating wave.

The extracted conductance of the transmission line with the ‘T’-rails is much higher than the other two cases, as shown in Fig. 3.14(c). However, its resistance shown in Fig. 3.14(a) and its attenuation shown in Fig. 3.9(a) are similar to the third case, indicating again that



(a) (b)

(c) (d)

Fig. 3.14 RLGC parameters for the three CPS cases shown in Fig. 3.8.

the attenuation is dominated by the resistive (conductor) losses and not the conductance (dielectric) losses for the case of the unloaded CPS transmission line.

The 'T'-rails transmission line and the wider CPS also have lower characteristic impedance and this also increases the conductor loss according to Eq. 3.31 (assuming Z_{0ul} is in the low-loss regime, and thus independent of R_{tl} and G_{tl}). For example, going from a characteristic impedance of 70 Ω to 58 Ω increases the conductor loss by 20.7%. This will, however, also decrease the dielectric loss by the same percentage.

From Fig. 3.14, it is also observed that at high frequencies, the inductance (agreeing with Eq. 3.32) and capacitance are relatively constant. In the low-loss or high frequency regime, this also indicates that the dispersion is relatively lower at high frequencies compared to the lower frequencies. The inductance of the slow-wave CPS electrode is also dependent on the thickness of the stem denoted by t in Fig. 4.2. From additional simulations, it was observed that the index of the slow-wave CPS increased to that of the narrower CPS (case a) as the stem of the T-rail decreased in width. Moreover, a higher inductance and thus a larger slowing factor, can also be obtained by using a smaller width W and/or by increasing the distance between the bus and the inner edge of the 'T'-rails (G_T). This larger inductance would also increase the characteristic impedance, as described by Eq. 3.25, and this would need to be compensated by increasing the capacitance.

3.4 Junction Loaded Transmission Line

In carrier-depletion modulation, the p-n junction is operated in reverse bias. Under reverse bias and small-signal operation, the p-n junction can be modeled as shown in Fig. 2.17(b). For simplicity, the conductance of the depletion region is ignored when loading the transmission line. Because the depletion capacitance changes with increasing reverse bias voltage, the response of the transmission line is also dependent on the biasing voltage. The transmission line loaded with the p-n junction can be modeled by simply inserting the series R_{pn} and C_{pn} elements as shown in Fig. 3.11(b) [39, 85]. The p-n junction resistance and capacitance have distributed units of [Ω-length] and [F/length], respectively.

In order to simplify the loaded telegraphist model, the series R_{pn} and C_{pn} elements can be transformed into an equivalent parallel RC circuit, as shown in Fig. 3.11(c). This is accomplished with the following transformation:

$$Q = \frac{1}{\omega C_{pn} R_{pn}} \tag{3.42}$$

$$R_{tpn} = R_{pn}(1 + Q^2) \tag{3.43}$$

$$C_{tpn} = C_{pn}\frac{Q^2}{1 + Q^2} \tag{3.44}$$

With this transformation, the resistance and capacitance of the junction can be easily merged with the conductance and capacitance of the unloaded transmission line. Then, the characteristic impedance of the junction loaded line Z_{0l} and the propagation constant γ_l is:

$$Z_{0l} = \sqrt{\frac{R_{tl} + j\omega L_{tl}}{(G_{tl} + R_{tpn}^{-1}) + j\omega(C_{tl} + C_{tpn})}} \tag{3.45}$$

$$\gamma_l = \alpha_l + j\beta_l = \sqrt{\left(R_{tl} + j\omega L_{tl}\right)\left([G_{tl} + R_{tpn}^{-1}] + j\omega[C_{tl} + C_{tpn}]\right)} \tag{3.46}$$

Similar to the unloaded transmission line case, the loaded characteristic impedance, propagation constant, and phase index under low-loss conditions become:

$$Z_l = \sqrt{\frac{L_{tl}}{C_{tl} + C_{tpn}}} \tag{3.47}$$

$$\gamma_l = \overbrace{\frac{1}{2}\left(\underbrace{\frac{R_{tl}}{Z_l}}_{\substack{\text{conductor}\\\text{loss}}} + \underbrace{\left(G_{tl} + R_{tpn}^{-1}\right) Z_l}_{\substack{\text{dielectric}\\\text{loss}}}\right)}^{\alpha_l\ [Np/m]} + j\,\overbrace{\omega\sqrt{L_{tl}\left(C_{tl} + C_{tpn}\right)}}^{\beta_l\ [rad/m]} \tag{3.48}$$

$$n_{\mu,l} = c_0\sqrt{L_{tl}\left(C_{tl} + C_{tpn}\right)} \tag{3.49}$$

In these equations, it is assumed that loading the transmission line did not change the inductance. This is true if the currents in the silicon do not flow along the transmission line and only flow orthogonally. In this case, no self-inductance would be present in silicon [85].

With the p-n junction loading, compared to the unloaded case, both the conductor loss and the dielectric loss change due to the new characteristic impedance. From Eq. 3.46, it is observed that the dielectric loss has an additional term from the silicon p-n junction. In order to simplify the expression for the conductor and dielectric losses, it is assumed that the frequency of operation is much smaller than the junction bandwidth: $\omega \ll \omega_{pn}$, with $\omega_{pn} = 1/(R_{pn}C_{pn})$. With this inequality, Q of Eq. 3.42 is assumed to be much greater than unity. Then, the transformed resistance (Eq. 3.43) and capacitance (Eq. 3.44) of the junction is approximated as:

$$R_{tpn} \approx \frac{1}{\omega^2 C_{pn}^2 R_{pn}} \tag{3.50}$$

$$C_{tpn} \approx C_{pn} \tag{3.51}$$

With the assumptions of Eq. 3.50 and 3.51, and low-loss conditions, the effect of the p-n junction loading on the characteristic impedance and microwave index can be summarized by Eq. 3.45 and Eq. 3.49. The junction loading reduces the characteristic impedance and increases the microwave index.

With the same assumptions, the loaded transmission line conductor loss, and the silicon dielectric loss can be expressed as:

$$\alpha_{cl} = \frac{1}{2}\left(R_{DC} + R_{AC}\sqrt{f}\right)\sqrt{\frac{C_{tl} + C_{pn}}{L_{tl}}} \tag{3.52}$$

$$\alpha_{Sil} = \frac{1}{2}\left(4\pi^2 f^2 C_{pn}^2 R_{pn}\right)\sqrt{\frac{L_{tl}}{C_{tl} + C_{pn}}} \tag{3.53}$$

From Eqn. 3.52, the conductor loss increases with $\sqrt{f}$, while the dielectric loss (Eq. 3.53) increases quadratically with frequency. In addition, the dielectric loss also increases quadratically with the p-n junction capacitance and linearly with p-n junction resistance. Therefore, to reduce loss, it is advantageous to reduce the junction capacitance instead of reducing the resistance. This can be achieved by placing two diodes in a series, such as in a SPP configuration, which would roughly half the capacitance and double the resistance. Assuming identical Z_{0l}, this would mean a SPP modulator can have 50% lower silicon dielectric loss compared to a single or dual drive modulator. In the situation that no changes are made to keep the Z_{0l} identical for both cases, then the benefit will lower. As an example, if it is assumed that $C_{pn} \approx 2C_{tl}$, then a SPP configuration would have 39% lower silicon dielectric loss. The SPP structure also has the advantage that it reduces capacitance without affecting the modulation efficiency. If the junction capacitance of a single drive modulator was to be reduced, it would certainly increase the $V_\pi L_\pi$ figure-of-merit.

By looking at Eq. 3.48, it is also observed that with a lower characteristic impedance, the conductor loss can be increased while decreasing the dielectric loss. This is beneficial as the conductor loss is much smaller than the loaded dielectric loss resulting in a lower overall loss. However, with this, the transmission line would no longer be matched to the common 50 Ω sources and terminations.

In Fig. 3.15, the total microwave loss, obtained from simulations in HFSS is plotted. It is observed that, as predicted by the equations, the loss is higher for the case with a single p-n junction. The loss is significantly lower for the unloaded transmission line. The loss of the case with two p-n junctions in series falls in between the other two cases. The difference

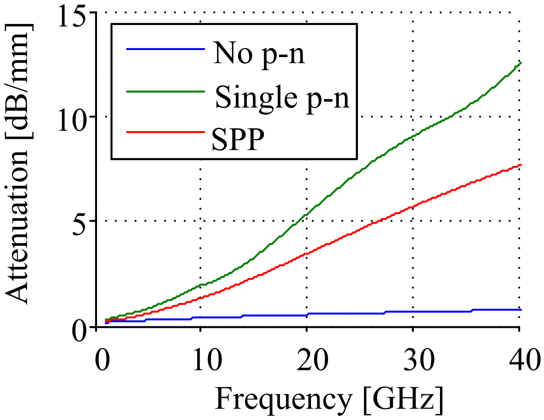
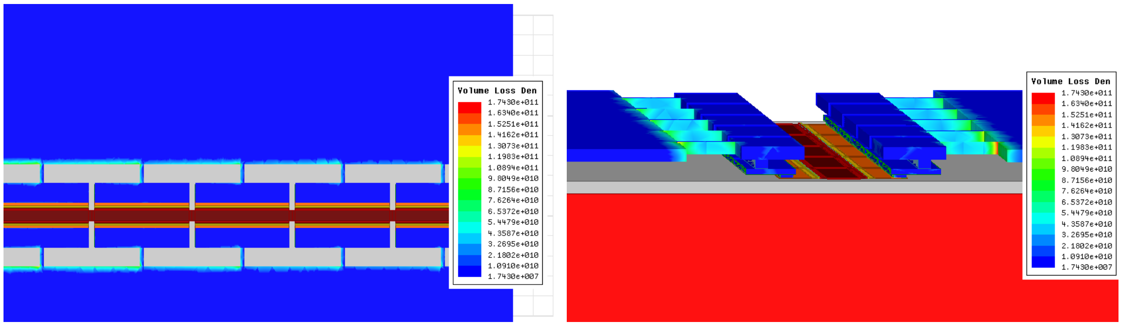


Fig. 3.15 FEM simulations of microwave loss for cases of transmission line without a p-n junction, a single p-n junction, and two p-n junction in series (for SPP).

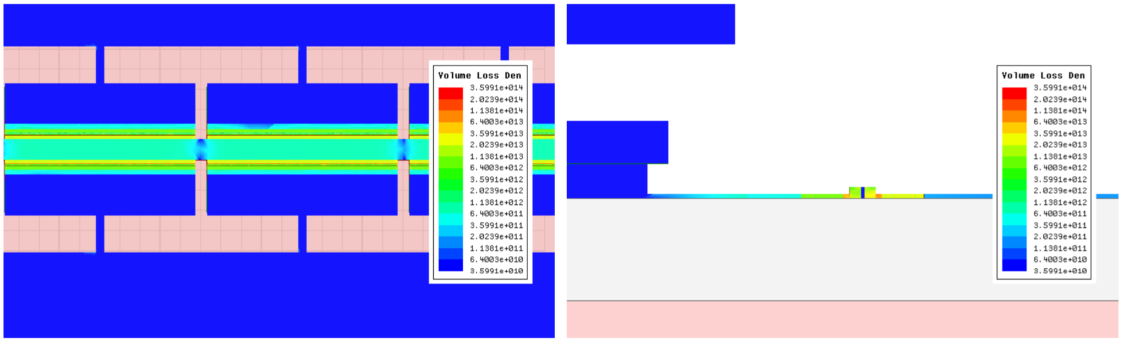
in loss is significant. For example, considering the 6.4 dB/mm point, this corresponds to 22 GHz for the single p-n case, but 33 GHz for the SPP case. For a 1 mm long device, this would correspond to its electro-optic bandwidth under the assumption of perfect velocity and impedance match. Fig. 3.15 also shows that the silicon dielectric loss is dominant, because the simulated attenuation is the total loss, but the trend follows that of Eq. 3.53.



(a) (b)

Fig. 3.16 Volume density loss in the metal and vias at 30 GHz, (a) top view, and (b) side view.

In Fig. 3.16, the volume loss density in the metal and vias, obtained from simulations in HFSS is shown. It is observed that most of the metal loss happens at the bridge of the ‘T’ shaped extensions. The other location of loss is the edges of the bus electrode. In Fig. 3.17, the volume loss density of the metals, vias, and all layers of silicon is shown (on a logarithmic scale). It is observed that the silicon p-n junction loss is significantly greater than the loss in the metal. Furthermore, the loss in the silicon is higher in the low doped (high resistance) regions. The highest loss occurs in the 90 nm N and P doped slabs. As discussed earlier, a larger spacing between electrodes would lower the conductor loss for unloaded transmission lines. In a SPP configuration, however, increasing the spacing would also increase the size of the highly doped common semi-conductor, which would lead to an increase in microwave loss.



(a) (b)

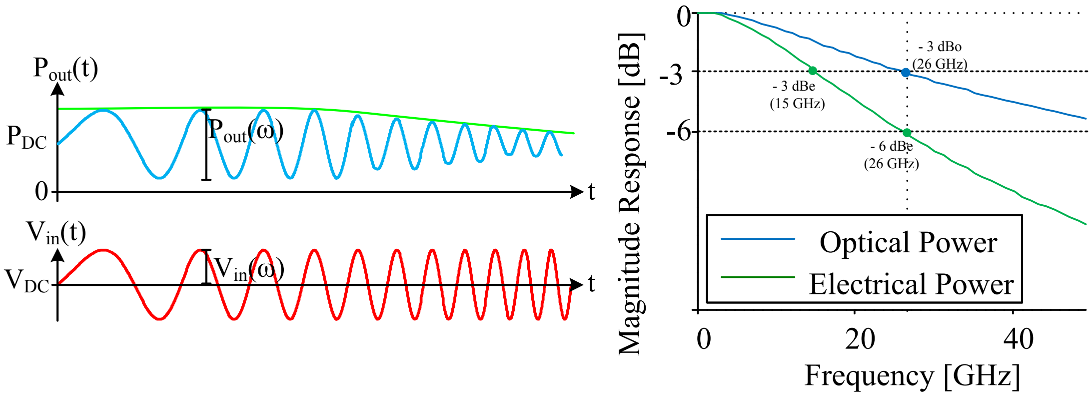
Fig. 3.17 Volume density loss (logarithmic scale) in the metal, vias, and silicon at 30 GHz, (a) top view, and (b) side view.

3.5 Electro-optic Small Signal Modulation Response

The small signal modulation response $M(\omega)$ is defined as the optical modulation power for a given electrical input voltage. It is the measure of how efficient a modulator is at converting electrical input into modulated power as a function of frequency. The efficiency generally drops as frequencies increase as illustrated by the example in Fig. 3.18(a), in which the light blue curve represents the envelope of the instantaneous optical power. The modulated power is the depth of this envelope. The modulation response is defined as:

$$M(\omega) = \frac{P_{out}(\omega)}{V_{in}(\omega)} \tag{3.54}$$

Where, P_{out} is the phasor of the modulated optical power, and V_{in} is the phasor of the driving voltage.



(a) (b)

Fig. 3.18 Graphical depiction of (a) small signal modulation response, and (b) electrical and optical bandwidth definitions. In (a), the amplitude of the input voltage (red) remains constant over a range of frequency while the corresponding modulated optical power (blue) decreases in amplitude for higher input frequencies, thus exhibiting a low-pass E-O modulation response.

Generally, only the magnitude of the small signal modulation response (the responsivity) $m(\omega)$ is of interest. It is also typically normalized to DC or a low-frequency point. However, in the context of equalization, the response should be normalized to the maximum.

$$m(\omega) = \frac{|M(\omega)|}{|M(0)|} \approx \left|\frac{\Delta\phi(\omega)}{\Delta\phi(0)}\right| \tag{3.55}$$

The modulation response (not normalized, Eq. 3.54) is the quantity that is measured by a typical lightwave component analyzers (LCA) when characterizing an electro-optic device. Therefore, a modulator biased far from the quadrature point will have a very low reading on the vertical scale of the LCA as the modulation depth can be very small. The value of the modulation response (magnitude of Eq. 3.54) also depends on the input power at the receiver, but the shape of the response and the normalized response (Eq. 3.55) is generally independent of the received power.

More specifically, Eq. 3.55 is the optical definition of the modulation response. The distinction between optical and electrical bandwidth is illustrated by Fig. 3.18(b), where the green curve plots the electrical power ($20log_{10}$ of PD current) and the blue curve the optical power ($10log_{10}$ of PD current). The optical -3 dB bandwidth (-3 dBo) is the point where the optical power falls to 50%. This corresponds to the level where the current of a detector also drops by 50% because current is directly proportional to measured optical power. The electrical power, however, is quadratically related to the current. The electrical -3 dB bandwidth (dBe) is the point at which the electrical power falls to 50%. The -3 dBo bandwidth, thus corresponds to the -6 dBe bandwidth.

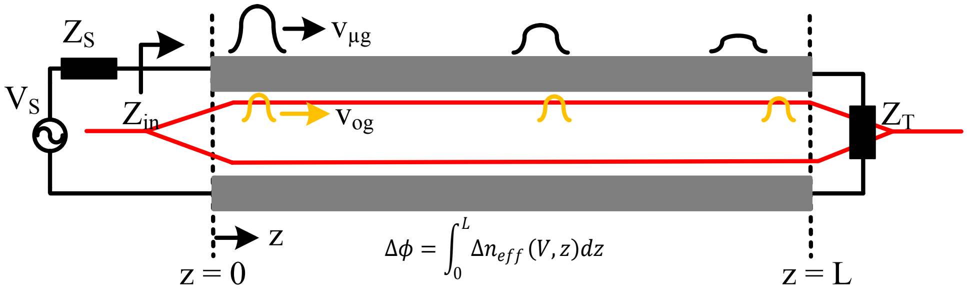


Fig. 3.19 Traveling wave electro-optic interaction. The net phase shift experienced by the optical wave is the integral of the change in effective index at each point of the transmission line. The $\Delta n_{eff}(V,z)$ depends on the attenuation of the microwave pulses (black) and group velocity mismatch with the optical pulses (orange) as they propagate along the transmission line.

If the modulator is biased at quadrature and a small voltage is applied, then it can be assumed that the phase shift varies linearly with an applied voltage. With this assumption, the small signal response can be thought of as being equivalent to the frequency response of the phase shift experienced by an optical wave. The total change in phase experienced by the optical wave can be obtained by integrating over the length of the modulator the change in index experienced by the optical pulse at an infinitesimal point. The traveling wave electro-optic interactions affect the amount of index change along the modulator because it is dependent on the microwave attenuation and the matching of their velocities as illustrated in Fig. 3.19. As the microwave propagates, it attenuates and the index change experienced by the optical wave reduces. In addition, a walk-off between the optical wave and microwave leads to a different part of the optical wave experiencing the change in index leading to inter-symbol interference.

The final modulation response of a $LiNbO_3$ TWMZM was derived in [86], as:

$$m(\omega) = \frac{R_T + R_S}{R_T} \left| \frac{Z_{in}}{Z_{in} + Z_S} \right| \left| \frac{(Z_T + Z_0)F(u_+) + (Z_T - Z_0)F(u_-)}{(Z_T + Z_0)e^{\gamma_m L} + (Z_T - Z_0)e^{-\gamma_m L}} \right| \tag{3.56}$$

Where, Z_T, Z_S, and Z_{in} are respectively the impedance at DC of the load, the source, and the impedance of the transmission line. The other parameters are:

$$F(u) = \frac{1 - exp(u)}{u} \tag{3.57}$$

$$u_{\pm}(\omega) = \pm\alpha_m L + j\frac{\omega}{c_0}(\pm n_{\mu g} - n_{og})L \tag{3.58}$$

Where, α_m is the microwave loss, L is the E-O interacting length of the MZM, $n_{\mu g}$ is the microwave group index, and n_{og} is the optical group index.

There are three main aspects to consider for optimizing the E-O response of the TWMZM: impedance matching, velocity matching, and microwave loss.

3.5.1 Impedance Mismatch

Impedance matching is necessary to reduce reflections at various interfaces which can cause inter-symbol-interference and also to obtain maximum power transfer to the modulator. At the source interface, the transmission line should be matched to the source impedance, and at the end of the modulator, it must be matched to the termination impedance. The modulator's characteristic impedance is primarily determined by the driver's impedance, because it is difficult to create impedance matching circuits that are as broadband as the modulator. Broadband on-chip terminations can be created with a custom impedance as attested by the measurements in Section 5.3. Typical RF drivers are designed for 50 Ω impedance, though some commercially available drivers can have a lower impedance.

An impedance mismatch appears as ripples in the frequency response, which can result in a larger bandwidth due to peaking. In [87], simulations with a 50 Ω source impedance and a smaller termination impedance resulted in a larger bandwidth at the expense of a smaller voltage drop across the p-n junction at lower frequencies. At higher frequencies, the voltage drop was almost identical to a proper termination. The larger bandwidth was believed to be due to the reflections from the mismatch creating some sort of pre-emphasis on the RF signal [87]. This bandwidth refers to a response that was normalized at a lower frequency. Relative to the peak, the content at other frequencies are more attenuated.

At the extremes, an open-circuit termination leads to a narrow band response. A short circuit termination has zero transmission at DC, but has a resonance peak at a higher frequency. The effect of impedance match can be studied by solving Eqn. 3.56 with the assumption of perfect velocity match and no microwave attenuation. In this case, the modulation response simplifies to [86]:

$$m(\omega) = \left| \frac{(1 - \Gamma_T \Gamma_S)[1 + \Gamma_T F(U_-)]}{1 - \Gamma_T \Gamma_S e^{-2j\beta_m L}} \right| \tag{3.59}$$

$$U_- = -j \frac{2\omega}{c_0} n_{og} L \tag{3.60}$$

Where, Γ is the reflection coefficient for the termination or source relative to the characteristic impedance of the transmission line.

$$\Gamma_\zeta = \frac{Z_\zeta - Z_0}{Z_\zeta + Z_0} \tag{3.61}$$

Fig. 3.20 shows the calculated response in a system with zero microwave loss and perfect velocity match. In this case, the impedance of the transmission line was varied and the source and termination impedance was fixed to 50 Ω. It is observed that an impedance mismatch results in ripples in the response. The lower frequencies have the larger ripples.

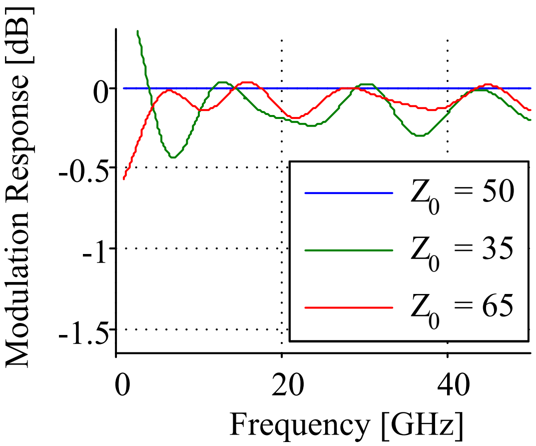


Fig. 3.20 Effect of characteristic impedance mismatch on the modulation response with perfect velocity mismatch and no microwave loss.

With an impedance mismatch at the interface of the source and the transmission line, only a certain percentage of the voltage is transferred to the transmission line. This is calculated by voltage division:

$$V_{in} = \frac{Z_{in}}{Z_S + Z_{in}} V_S \tag{3.62}$$

3.5.2 Microwave Loss

The effect of microwave loss can be studied by solving Eq. 3.56 and assuming that the transmission line is perfectly impedance matched ($Z_L = Z_T = Z_0$) and velocity matched. Then, the expression for $m(\omega)$ becomes [86]:

$$m(\omega) = e^{-\alpha_m(f)\frac{L}{2}} \left| \frac{sinh(\alpha_m(f)\frac{L}{2})}{\alpha_m(f)\frac{L}{2}} \right| \tag{3.63}$$

Solving this for the -3 dBo leads to $\alpha_m(f_{3dBo})L = 6.897$ dB, and for the -3 dBe point, it results in $\alpha_m(f_{3dBe})L = 6.339$ dB. This indicates that for a loss-limited modulator, the -3 dBe and -3 dBo point will correspond to the 6.4 and 6.9 dB of microwave loss, respectively. Since the loss is linearly proportional to length, the -3 dB bandwidth of loss-limited TWMZMs is inversely proportional to the length.

The dependence of the CPS geometry on the microwave attenuation and the sources of microwave loss of the loaded and unloaded junction were presented in Sections 3.3.4-3.4.

3.5.3 Velocity Mismatch

The effect of velocity mismatch on bandwidth can be studied by solving Eq. 3.56 with the assumption that the transmission line is perfectly matched $Z_T = Z_S = Z_0$, and that it is is lossless. Under these circumstances, only the component of the forward traveling wave contributes to the modulation response. The expression for $m(\omega)$ becomes [86]:

$$m(\omega) = \left| sinc\left(\frac{\pi f \Delta n_{mo} L}{c_0}\right) \right| \tag{3.64}$$

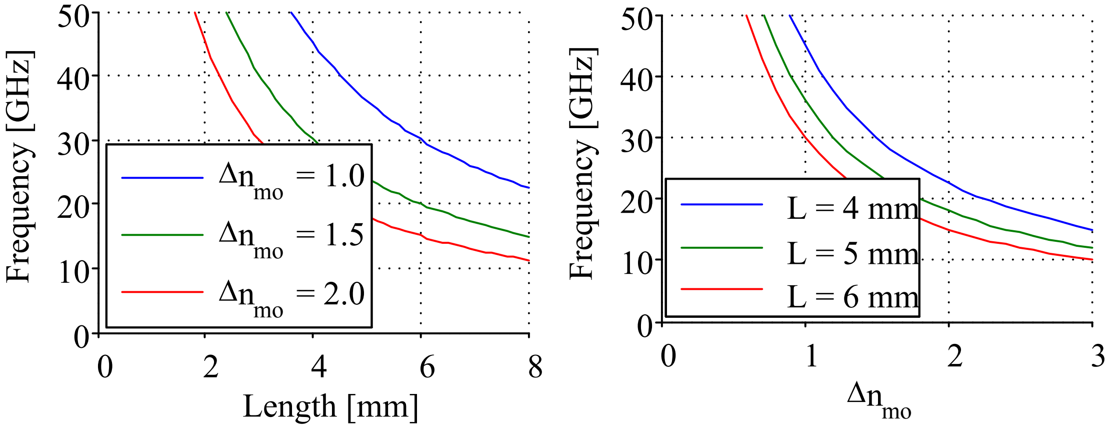
Where, Δn_{mo} is the difference in microwave and optical group index. When solved for the optical 3dB bandwidth, it results in:

$$f_{3dBo} = \frac{1.89 c_0}{\pi |n_{\mu g} - n_{og}| L} \, [Hz] \tag{3.65}$$

This equation demonstrates that the velocity mismatch is less significant for short electrodes. The ideally achievable bandwidth increases with decreasing length L of the TWMZM .

With counter propagation, for example in Michelson interferometers (or reflective MZMs), the effect on the bandwidth due to the mismatch can be calculated by using a positive sign instead of a negative one, i.e., the mismatch adds: $\Delta n_{mo} = |n_{\mu g} + n_{og}|$

In Fig. 3.21, the effect of the velocity mismatch is shown. For a 4 mm long modulator, an index mismatch of 1 will theoretically allow for a bandwidth of 45 GHz. The graphs also show that as the index mismatch increases, the dependency on bandwidth and length of the modulator becomes weaker.



(a) (b)

Fig. 3.21 Effect of velocity mismatch on the optical -3 dB bandwidth, shown for (a) different index mismatch, and (b) for different lengths of the TWMZM.

The simulated group index of the TE mode in rib and channel waveguides were presented in Section 2.2.1. The simulated phase index of the microwave was shown in Sections 3.3.4-3.4 for different cases of CPS lines. However, it is the group velocities of the microwave and optical waves that need to be matched as depicted by Fig. 3.19. This is because the optical wave is no longer a single frequency wave after it interacts with the microwave and becomes a phase modulated packet traveling at the group velocity [88]. If the dispersion of the transmission line is insignificant, then the phase index and group index are almost equal. This is the case for TEM microwaves [88]. The phase index of the microwave can then be used to match the optical group velocity. Otherwise, the group index can be approximated from the phase index. With the phase velocity of the loaded line ($v_{\mu l} = c_0/n_{\mu,l}$), the group velocity of the loaded line ($v_{g\mu l} = c_0/n_{g\mu,l}$) can be obtained as:

$$v_{g\mu l} = \frac{\partial \omega}{\partial \beta_l} = v_{\mu l} + \beta_l \frac{dv_{\mu l}}{d\beta_l} \tag{3.66}$$

Since the resistance and the capacitance of the p-n junction varies with voltage, this creates a transmission line whose microwave index also changes with the applied bias voltage. Hence, it is not possible to achieve perfect velocity matching for all bias voltages.

3.6 Segmented p-n Junction MZM Design

Segmented capacitive loading TWMZM designs have been commonly used for GaAs and InP platforms. Rather than having a continuous p-n junction for the entire length of the silicon TWMZM, by segmenting the p-n junction, the junction capacitance can be reduced. This will, however, increase the junction resistance in the lumped model of each segment, therefore maintaining the RC constant. In III-V modulators, segmented loading can be done to add capacitance and slow-down the microwave, however, in silicon, this segmentation can only be used to reduce the capacitive loading and speed up the microwave. This is because, in silicon modulators, the lateral p-n junction is the capacitor that is used to do the modulation and without segmentation, the transmission line will be the most loaded.

The amount of segmentation is defined by a fill factor f. The fill factor determines how much of a segment is filled with the p-n junction. The periods of the segments should be chosen to be small to allow for lumped element modeling for simulations, and also to push the Bragg frequency higher (discussed in Section 3.8).

The fill factor is easiest to calculate using the low-loss (or high-frequency) equations for the characteristic impedance Z_{0l} (Eq. 3.47) and microwave index $n_{\mu,l}$ (Eq. 3.49) of the loaded transmission line as derived in Section 3.4. As desired for velocity matching, $n_{\mu,l}$ is made equal to the optical group index n_{og}. Then, combining $n_{\mu,l}$ with the expression for Z_{0l} and assuming that the p-n junction does not change the inductance of the transmission line, the condition on the unloaded inductance L_{tl} of the line becomes [89]:

$$L_{tl} = \frac{n_{og} Z_{0l}}{c_0} \tag{3.67}$$

This expression for the unloaded inductance L_{tl} can then be substituted into the low-loss unloaded transmission line characteristic impedance Z_{0ul} (Eq. 3.26), to yield:

$$Z_{0ul}^2 C_{tl} = \frac{n_{og} Z_{0l}}{c_0} \tag{3.68}$$

By substituting the expression for the unloaded microwave index $n_{\mu,ul}$ (Eq. 3.28) for Z_{0ul} and simplifying, this leads to an expression for Z_{0ul} in terms of the optical group velocity and the loaded characteristic impedance:

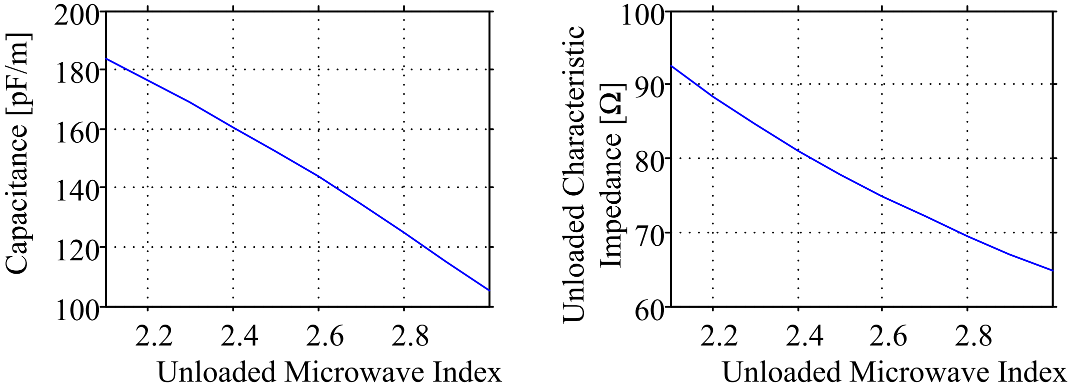
$$Z_{0ul}n_{\mu,ul} = n_{og}Z_{0l} \Rightarrow Z_{0ul} = \frac{n_{og}Z_{0l}}{n_{\mu,ul}} \tag{3.69}$$

Moreover, with $n_{\mu,l}$ set to n_{og}, Z_{0l} can be rearranged to give the desired capacitance of the p-n junction $C_{pn,d}$ for simultaneous impedance and velocity matching:

$$C_{pn,d} = \frac{n_{og}^2 - n_{\mu,ul}^2}{c_0 Z_{0l} n_{og}} \tag{3.70}$$

This also requires the assumption of Eq. 3.51 to hold.

From the simulations of the optical group index in a rib waveguide of Section 2.2.1, n_{og} is estimated to be 3.89 at 1550 nm. The target loaded impedance Z_{0l} is 50 Ω. The required unloaded characteristic impedance loading and capacitance can then be calculated using Eq. 3.69 and Eq. 3.70, respectively. These were calculated for different $n_{\mu,ul}$ and is plotted in Fig. 3.22.



(a) (b)

Fig. 3.22 Calculated requirements for the low-loss case of (a) junction capacitance, and (b) unloaded characteristic impedance for simultaneous matching of impedance and velocity.

By segmenting the junction, the effective p-n junction capacitance can be lowered to achieve the desired loading capacitance. The fill factor f can be determined from the actual junction capacitance $C_{pn,a}$ and the desired capacitance. For a SPP modulator, the actual junction capacitance can be approximated to half.

$$f = \frac{c_{pn,d}}{c_{pn,a}} \tag{3.71}$$

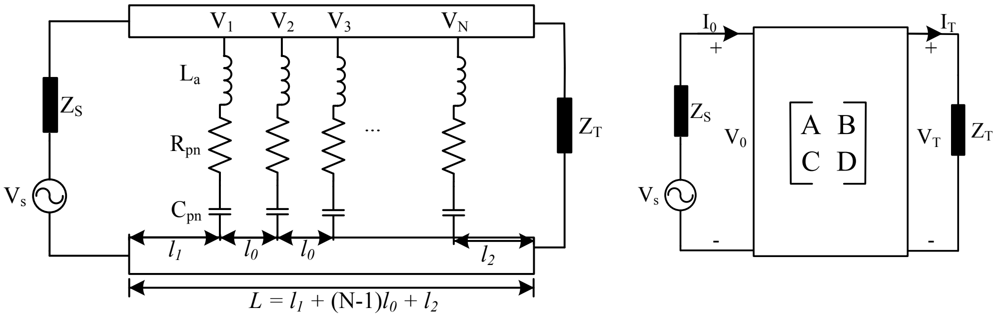
For example, if $C_{pn,a}$ is 225 pF/m, at a $n_{\mu,ul}$ of 2.3, the loading required is about 168 pF/m. Then for a single drive or a dual drive modulator, the fill factor should be chosen as 74.7%. For a SPP modulator, $C_{pn,a}$ halves resulting in a fill factor of 150%. This fill factor is unrealistic. The large requirement in fill factor signifies that the capacitance loading for SPP due to the p-n junction is insufficient to slow down the microwave. To achieve this, slow-wave electrode structures such as the CPS with ‘T’ shaped extensions or floating metal lines need to be employed.

By segmenting the p-n junction, the total active length of the phase shifter is also reduced. This would increase V_π at DC. At higher frequencies, if the bandwidth is extended due to the proper impedance and velocity match from the segmentation, then the voltage will actually be more effective at doing the modulation, even though the active length has been reduced.

3.7 Cascaded Matrix Model

The effect of the p-n junction loading on an unloaded-transmission line can also be studied by using a cascaded matrix model. In this model, transmission matrices (ABCD) of small segments of the TWMZM are multiplied to obtain the overall response. The segments must be small enough to satisfy the length condition required by the lumped telegraphist model described in Section 3.3.4. The frequency analysis analytical model using this method was described by Li et al. in [89]. It also considers the electro-optic interaction and simplifies combining the interactions of Sections 3.3.4 and 3.5. This model was validated for GaAs TWMZMs in [89] and is repeated here.

The p-n junction loaded section of the TWMZM is divided into N uniform segments of length l_0 as illustrated in Fig. 3.23(a). Each segment n is modeled by two ABCD matrices, which are obtained from lumped element models. One ABCD matrix (T_0) represents the unloaded transmission line and is obtained by converting the unloaded propagation constant γ_{ul} and characteristic impedance Z_{0ul} to ABCD parameters as:



(a) (b)

Fig. 3.23 Cascaded matrix modeling of TWMZMs as described in [89]. Subfigure (a) depicts the division of the TWMZMs into N active segments, and (b) shows the entire TWMZM modeled as a single ABCD matrix. Figures adapted from [89].

$$T_0 = \begin{bmatrix} cosh(\gamma_{ul} l_0) & Z_{0ul} sinh(\gamma_{ul} l_0) \\ Z_{0ul}^{-1} sinh(\gamma_{ul} l_0) & cosh(\gamma_{ul} l_0) \end{bmatrix} \tag{3.72}$$

The other ABCD matrix (T_a) represents the p-n junction by a series L_a, R_a, and C_a circuit. Where, L_a would represent any additional inductance present between the p-n junction and the transmission line. The variables R_a and C_a are the resistance R_{pn} and capacitance C_{pn} of the junction adjusted with the fill factor f for segmented designs as discussed in Section 3.6. For a continuous junction, the fill factor can be set to unity. These three quantities are in lumped units (i.e., not per unit length).

$$T_a = \begin{bmatrix} 1 & 0 \\ \left\{ R_a + j\left(\omega L_a - \frac{1}{\omega C_a}\right) \right\}^{-1} & 1 \end{bmatrix} \tag{3.73}$$

The ABCD matrices, unlike S-parameters, can be cascaded to obtain the ABCD matrix of the total transmission line T_T. By matrix multiplication:

$$T_T = T_1 (T_a T_0)^{N-1} T_a T_2 \tag{3.74}$$

Where, T_1 and T_2 are the ABCD matrices for de-embedding the unloaded transmission lines from the edges of the bond pads to the start of the loaded transmission line. In Fig. 3.23(a), these segments are shown with a length of l_1 and l_2, respectively. With the

ABCD matrix of the entire circuit, the voltages and currents at the input and output are expressed as:

$$\begin{bmatrix} V_0 \\ I_0 \end{bmatrix} = \begin{bmatrix} T_T \end{bmatrix} \begin{bmatrix} V_T \\ I_T \end{bmatrix} \tag{3.75}$$

Knowing that the source voltage is: $V_S = V_0 + I_0 Z_0$, the expressions for the current I_T and voltage V_T at the termination can be obtained as:

$$I_T = \frac{V_S}{AZ_T + B + CZ_S Z_T + DZ_S} \tag{3.76}$$

$$V_T = I_T Z_T = \frac{V_S Z_T}{AZ_T + B + CZ_S Z_T + DZ_S} \tag{3.77}$$

Once the I-V at the termination is known, the voltage at each segment n can be determined successively, starting from the end of the transmission line and moving to the beginning. For carrier depletion MZMs, the voltage across the p-n junction capacitance is related to the phase shift and it determines the modulation response as states in Eq. 3.55. The voltage across the capacitance of segment n is found from V_n by voltage division:

$$V_{cn} = V_n \left[\frac{1}{1 - \omega^2 L_a C_a + j\omega R_a C_a} \right] \tag{3.78}$$

Finally, the modulation response is obtained by summing all the modulation voltages and multiplying it by a phase factor determined by the group velocity of the optical wave. The frequency response is normalized to $V_S/2$, which is the voltage that would be given by the source to the transmission line with perfect impedance match (Eq. 3.62).

$$m(f) = \left| \underbrace{\frac{2}{NV_S}}_{\substack{\text{Normalization} \\ \text{constant}}} \sum_{n=1}^{N} \overbrace{V_n \left(\frac{1}{1 - \omega^2 L_a C_a + j\omega R_a C_a} \right)}^{\text{Voltage across } n^{th}\ C_a} \underbrace{e^{j\frac{\omega(l_1 + (n-1)l_0)}{v_o}}}_{\substack{\text{microwave} \\ \text{phase factor}}} \right|^2 \tag{3.79}$$

The parameters of the unloaded transmission line used in the computation of Eq. 3.72 are frequency dependent as demonstrated in Section 3.3.4. The propagation constant and characteristic impedance of the loaded line can be determined by applying Eq. 3.22 and Eq.3.23 to T_T.

3.8 Bragg Frequency and Periodic Structures

The presence of a periodic structure gives rise to the Bragg effect, which can result in a stop-band and pass-band in the frequency response. The Bragg frequency is higher for structures with a smaller period. This cut-off frequency can be estimated using the microwave phase velocity of the junction-loaded transmission line $v_{\mu l}$ and the period p [90]:

$$f_{cutoff} \approx \frac{v_{\mu l}}{2p} \tag{3.80}$$

Therefore, it is desired to have an electrode design with a small repeating period.

Chapter 4

Device Design and Layout

4.1 Introduction

THIS chapter provides a summary of the design choices used in the fabricated device with respect to the discussions and results provided in the previous chapters of this thesis.

4.2 p-n Junction

With the foundry process used to manufacture this device (OpSIS-IME), it was not possible to design vertical p-n junction, or p-n junctions with compensated doping. Therefore, carrier-depletion based lateral p-n junctions were chosen. This also allowed for high speed operations and lower capacitance compared to the other possibilities as explained in Section 2.6.

The modulator was fabricated in a silicon-on-insulator (SOI) wafer with a top silicon thickness of 220 nm and a 750 Ω-cm silicon substrate. The diode-waveguide structures were formed by doping rib waveguides that are 500 nm wide and 220 nm high on a 90 nm slab for operation in the C-band. The p-n junction doping concentrations and clearance were chosen to be identical to that reported in [22]. This made it possible to compare the impact of the electrode design and the series push-pull (SPP) driving configuration against a dual-drive modulator without T-shaped extensions. The cross-section of the SPP p-n junction and the dimensions used in the fabricated design is shown in Fig. 4.1.

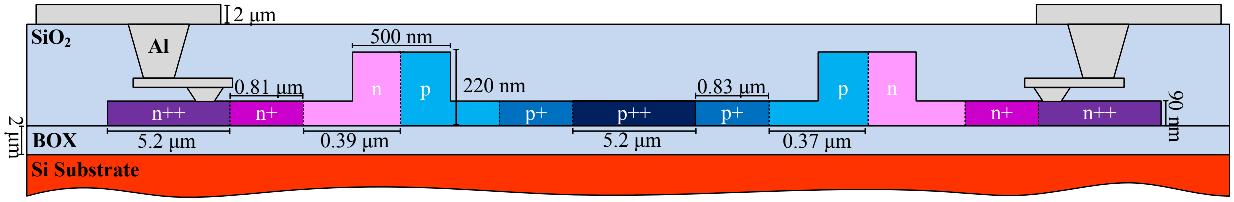


Fig. 4.1 Cross-section schematic of the SPP TWMZM design on an SOI wafer (drawing not to scale).

4.3 Transmission line

The SPP configuration was chosen as the design because of its smaller footprint, lower microwave loss, and the ability to operate at high bias voltages with negligible static power consumption. A more detailed explanation and comparison was provided in Section 2.4. By placing two diodes in series, the capacitance is halved but the series resistance is doubled, thus the RC constant remains about the same. The reduction in capacitance, however, reduces the microwave loss, as discussed in Section 3.4, which can help increase the bandwidth. Alternatively, loss can be lowered by reducing Z_{0l}, which would decreases the dominant dielectric loss in exchange of higher conductor loss as was done in [22]. This TWMZM targeted an impedance of 50 Ω for compatibility with standard 50 Ω microwave drivers. The reduction in the capacitance regrettably increases the phase velocity of the microwave, thus increasing the mismatch between the optical and electrical waves. To compensate for this effect, coplanar strips electrodes using ‘T’-rails to increase the index of the microwave were implemented. The electrodes used in the layout of the device was as described and simulated in Section 3.4.

The electrodes are fabricated in the top 2-µm thick aluminum metal layer with the dimensions indicated in the sub-caption of Fig. 4.2. In our design, a spacing of 12.6 µm was used between the ‘T’-shaped segments (S_T) in order to minimize conductor loss. A larger spacing would ensure a lower loss from unloaded electrodes but this would increase loss from a larger width of the common dopant (p++ of Fig. 4.1) and also need larger trace widths for the same characteristic impedance.

In our design, t was chosen to be 2 µm, the smallest allowed by the fabrication design rules, to obtain the highest index increase as possible.

Because the capacitive loading due to the p-n junction does not increase significantly with the area of the pad ($W_T \times L_T$), W_T was chosen to be 10 µm as required by the design rules for the metal and via stack underneath the pad.

In our design, the p-n junction was striated to reduce the current flowing in the highly doped silicon instead of the metal transmission lines [22], but according to the simulations in [85], this should not be necessary. A period P_T of 50 µm and a gap of 3 µm was used between the 'T' segments, resulting in a L_T value of 47 µm and a fill factor of 94%. A small period was used to ensure that the Bragg cut-off frequency is beyond the desired bandwidth.

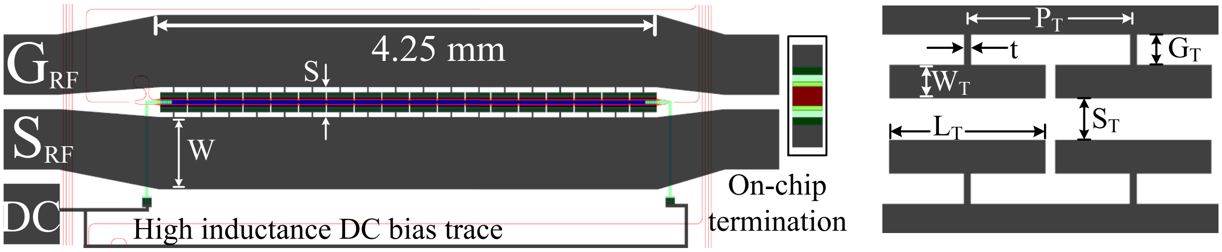


Fig. 4.2 Layout schematic of the SPP TWMZM and a magnified view of the 'T'-shaped extensions (schematic not to scale). The dimensions are: $W = 120\ \mu m$, $S = 51\ \mu m$, $t = 2\ \mu m$, $S_T = 12.6\ \mu m$, $G_T = 9.2\ \mu m$, $W_T = 10\ \mu m$, $L_T = 47\ \mu m$, and $P_T = 50\ \mu m$.

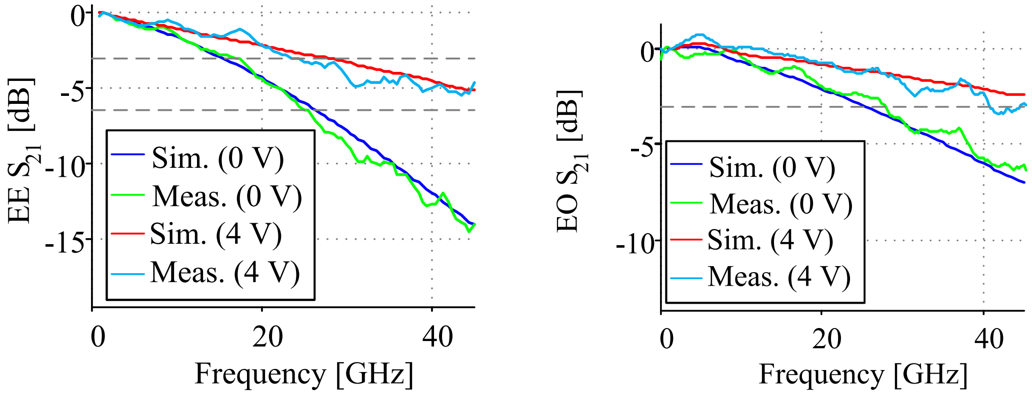
4.4 Modulator Length

The trade-offs with the modulator length was discussed throughout the thesis. A longer modulator will lead to a smaller V_π at DC. Long modulators would, however, result in a smaller E-O bandwidth due to higher losses and also due to a larger impact of velocity mismatch as presented in Section 3.5. One of the objectives of this work is to study the effects of using slow-wave electrodes on the electro-optic bandwidth, and velocity mismatch becomes more apparent for longer modulators. However, the die area was only 5 mm wide, therefore a long modulator could not be designed.

The fabricated design was 4.7 mm long from pad center to pad center. The p-n junction loaded section was 4.25 mm long with a fill factor of 94%. With this fill factor, the active phase-shifting length is 4 mm.

4.5 Electro-optic simulation

The E-E S_{21} simulation of the loaded transmission line was performed with the model and FEM simulations of Section 3.3.4-3.4. Due to variations in fabrications from mask misalignment errors and variations in doping concentration, the simulated responses did not match the measured response using previously published measured values of the p-n junction resistance and capacitance. A good agreement between the simulation and measured responses, as shown in Fig. 4.3(a), is obtained when the p-n junction parameters are used as frequency-independent adjustable parameters. This was necessary as the two diodes of the SPP TWMZM can have significantly different characteristics (as shown in Section 5.2) and it is no longer correct to assume that the SPP configuration will approximately double the resistance or halve the capacitance of the p-n junction. However, in dual-drive designs, the variations between the two arms of the MZM can be small if the same orientation of the two diodes is maintained during layout. This cannot be done for SPP designs as the two diodes need to be mirrors of each other.



(a) (b)

Fig. 4.3 Comparison of measured and simulated (a) E-E S_{21} response with R_{pn} and C_{pn} as fitting parameters, and (b) E-O S_{21} response using the electrical simulation of (a). Dashed lines mark -3 dB and/or -6.4 dB.

Using the E-E simulations, the E-O response was simulated using a similar method as in [89] and described in Section 3.7 and without using any additional fitting parameters. The simulated E-O responses are shown in Fig. 4.3(b) and a -3 dB bandwidth greater than 45 GHz was expected at 4 V reverse bias. The measured bandwidth is 41 GHz (4 GHz lower than simulated) at that bias voltage.

4.6 Device Layout

The schematic of the TWMZM was shown in Fig. 4.2. A 1×1 modulator design was chosen because of its simplicity and the availability of tested Y-branches with low insertion loss. The interferometer was formed using the compact Y-branches described in [57]. An intentional imbalance of 100 µm strip waveguide was incorporated to allow phase shift measurements and biasing of the modulator at quadrature point by tuning the wavelength as further explained in Section 2.3.3. For the purpose of this study, it was not necessary to operate the modulator at an absolute wavelength, hence carrier injection and/or thermal phase shifters were excluded from the design of this device. Light was vertically coupled to the chip using focusing grating couplers.

In order to obtain a high inductance for the bias voltage, a 5 µm wide and 5 mm long metal line was used to apply the reverse bias voltage. The line was connected to a highly doped p-type semiconductor common to both diodes of the MZM. The line was connected at both ends of the modulator in case there was a voltage drop along the length of the highly doped semi-conductor. The high inductance line is necessary to properly bias the modulator circuit and prevent the lower frequency components to be shorted to ground as explained in Section 2.4.2.

Additionally, an on-chip termination (OCT) was placed such that it was able to be connected with the modulator by ball bumping using a wire-bonder. For larger scale integration of MZMs, this would simplify RF packaging because it reduces the quantity of RF connectors necessary. The OCT was designed using Ohm's law and the sheet resistance provided by the foundry. The 50 Ω OCT is a 35-µm long and 50-µm wide n++ doped semiconductor.

The layout of the device was created using the Mentor Graphics Pyxis Layout tool and the OpSIS-IME process design kit (PDK). A parametrized cell (P-cell) of the TWMZM was created by writing a script in the Advanced Mentor Programming LanguagE (AMPLE) for Pyxis. Routing was performed with the interactive routing tool and waveguide bends were created using a Pyxis add-on developed by Lukas Chrostowski and Jonas Flueckiger from the University of British Columbia. The Design Rule Checks (DRC) was performed with Mentor Calibre.

4.7 Testability

Light was coupled in and out of the chip using vertical grating couplers. In addition to the optical I/O, two RF probes at each end of the modulator were needed for the electrical I/O. In order to be able to test the device with this requirement, a fiber array was used. This allowed for both the input and output grating couplers to be placed at one edge of the chip. The pitch of the fiber array, and therefore the pitch of the grating couplers was 127 µm.

This modulator was designed to be tested using GSSG probes with a 125 µm pitch. The RF signal was applied using one S of the pair and the DC bias with another S pad of the GSSG probe. Bondpads used were 100×100 µm with a 125 µm pitch and the top passiviation layer removed. Metal tiling was avoided near the modulator to remove any impact of floating metal pieces on the transmission line performance. Silicon tiling was also avoided in the vicinity of the modulator as a precaution. The effect of intrinsic silicon tiles was found to be negligible from further FEM simulations.

A photograph of the experimental setup is shown in Fig. 4.4.

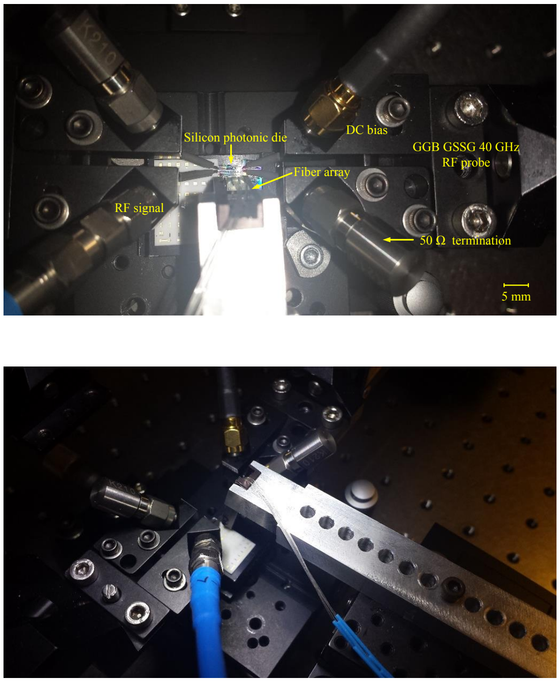


Fig. 4.4 Experimental setup.

Chapter 5

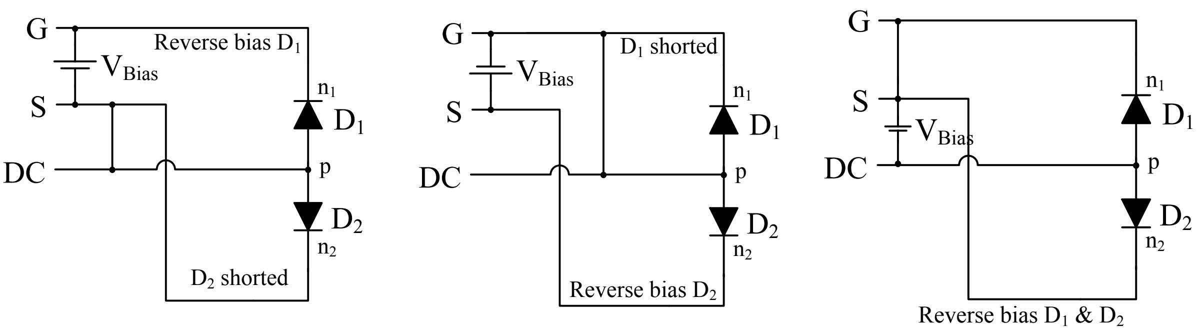
Experimental Characterization and Discussions

5.1 Introduction

This chapter presents and discusses the measured results of the silicon photonic Mach-Zehnder modulator. The measurements are categorized as DC and optical measurements, small-signal measurements, and large signal measurements. The large signal measurements were performed using a typical bit-error-rate testing configuration, followed by pre-emphasis using analog signal processing (ASP), and then by using digital-signal-processing (DSP). A detailed description of the methodology and experimental setup used in the characterization of the device is provided for each set of measurements.

5.2 DC and Optical Characterization

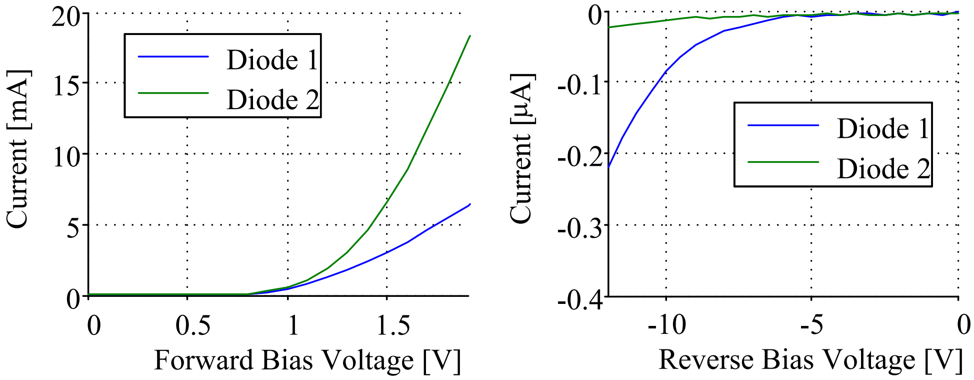
The current-voltage (I-V) relationship of the diodes were measured using a semiconductor parameter analyzer. The stimulus was applied to the terminals of the traveling-wave electrodes using a ground-signal (GS) probe as illustrated in Fig. 5.1. In order to characterize each arm of the SPP TWMZM separately, one of the diodes were shorted by applying the same potential at the p and n terminal of the diode. The experimental setup used to characterize reverse bias diode 1 with diode 2 being shorted is shown in Fig. 5.1(a). Similarly, the schematic for reverse biasing diode 2 with diode 1 shorted is shown in Fig. 5.1(b). The measured forward bias and reverse bias I-V curves are shown respectively in Figs. 5.2(a)



(a) (b) (c)

Fig. 5.1 DC configurations for testing individual diodes of a SPP TWMZM: (a) diode 1 reverse biased, (b) diode 2 reverse biased, and (c) both diodes under the same reverse bias.

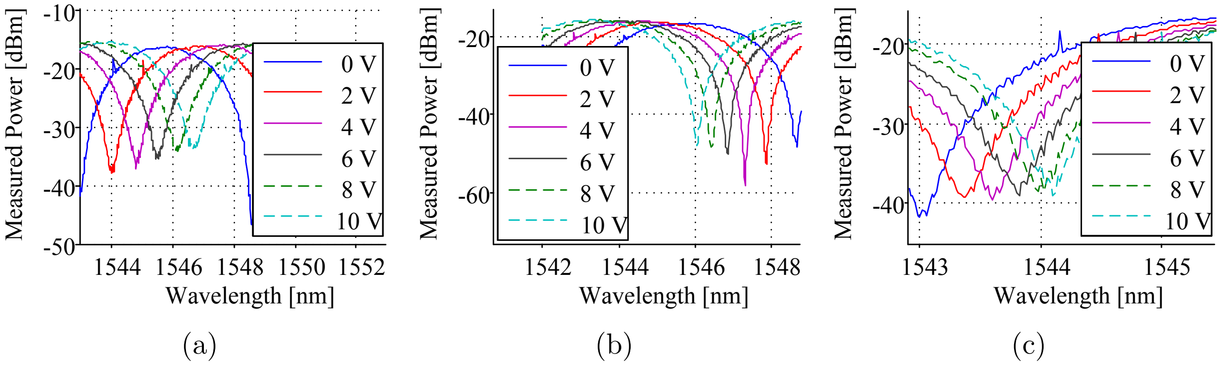
and 5.2(b). The forward bias curves indicate that both diodes begin to turn on at about 1 V. It is observed that diode 1 had a much higher resistance than diode 2. Under reverse bias operation, measurements show that diode 1 also begins to break down at -8 V, but diode 2 barely reaches the breakdown threshold at -12 V. These I-V curves indicate that at higher reverse bias voltages, the asymmetry in the modulation between the two arms of the MZM would worsen. The maximum current at reverse bias of 10 V occurs for diode 1 and is about 0.1 μA. Thus, reverse-bias leakage power can be considered to be negligible especially if considering energy/bit for high bit rates.



(a) (b)

Fig. 5.2 I-V measurements for (a) forward bias, and (b) reverse bias diodes of the MZM.

Using the same DC biasing schemes of Fig. 5.1, the transmission spectrum of diode 1, diode 2, and both diodes under the same reverse bias were measured. These spectra are shown in Fig. 5.3. From these transmission spectra, the insertion loss, the modulation phase shift, and the $V_\pi L_\pi$ FOM are determined. The maximum output of the modulator measured with an optical power meter was -15 dBm at 0 V bias. From a pair of nearby grating couplers, the fiber-to-fiber I/O coupling loss was measured to be 8.4 dB. The routing loss from the grating couplers to the modulator I/O is estimated to be 2.8 dB using the average strip waveguide loss of 2.4 dB/cm reported in [31]. Removing these contributions from the transmission spectra measurements, the insertion loss for the modulator is found to be approximately 3.8 dB.

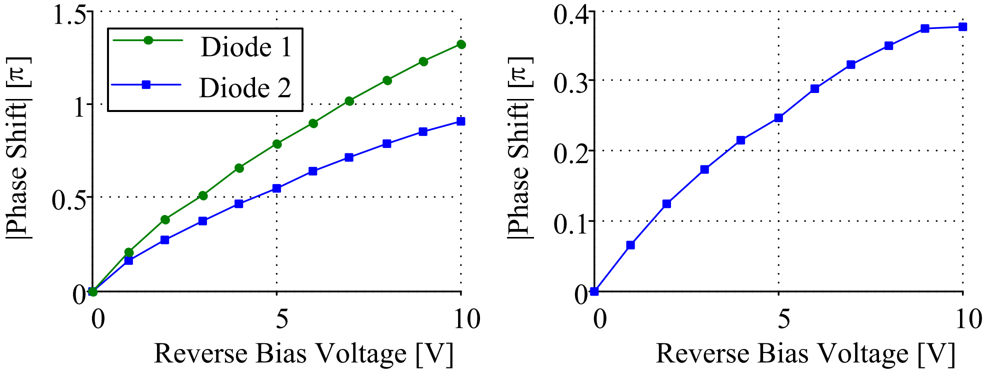


(a) (b) (c)

Fig. 5.3 Transmission spectrum measurements for (a) diode 1, (b) diode 2, and (c) both diodes under the same reverse bias voltage.

The corresponding phase shifts of each arm of the modulator under reverse bias was estimated from the measured transmission spectrum using $\Delta\phi = 2\pi\Delta\lambda/FSR$, where $\Delta\lambda$ is the wavelength shift of the spectrum relative to the 0 V reference, and FSR is the free spectral range described in Section 2.2.3. The determined phase shifts are plotted for diode 1 and diode 2 in Fig. 5.4(a), and for both diodes under the same reverse bias in Fig. 5.4(b). It is observed that the amount of phase shift is non-linear and decreases with increasing voltage. This occurs because the successive changes in the depletion region width overlap less with the optical mode. Moreover, the depletion width of the depletion region varies as the square root of applied voltage. Diode 1 attains a π phase shift at 7 V indicating a $V_\pi L$ of 2.8 V-cm. Diode 2 does not achieve a π phase shift up to a bias voltage of 10 V. Fig. 5.4(b) quantifies the asymmetry in the modulation performance of the two arms of the MZM. Ideally, no phase shift should be visible in this test case.

The DC measurements show that significant deviation is observed between two diodes that are placed only 7.6 µm apart. This is primarily believed to be due to mask misalignment during fabrication. Since junction capacitance is related to efficiency, this also indicates that the capacitance of the two diodes are not equal. Thus, for SPP designs, it is not strictly correct to assume that the capacitance will halve relative to single or dual-drive junctions. It should also be noted that this deviation can be reduced in dual-drive designs because the mask layers can be laid out in the same orientation, thus any misalignment will affect both diodes equally.



(a) (b)

Fig. 5.4 Phase shift calculated from transmission spectrum for (a) diode 1 and diode 2, and (b) both diodes under the same reverse bias voltage.

The $V_\pi L_\pi$ was calculated instead of being measured from different phase shifters of varying lengths. The values were calculated as $V_\pi L_\pi = \pi V_{applied} L_{DUT}/\Delta\phi$, where $V_{applied}$ is the applied reverse bias voltage, $\Delta\phi$ is the deduced phase shift at that voltage, and L_{DUT} is the phase shifting length of 4 mm. This calculated $V_\pi L_\pi$ is plotted in Fig. 5.5(a) for diode 1 and diode 2 in reverse bias, and for both diodes under the same reverse bias in Fig. 5.5(b). If the two diodes were identical, then the $V_\pi L_\pi$ of Fig. 5.5(b) would have been extremely high. It is again observed from Fig. 5.5(a) that the efficiency deteriorates at higher reverse bias voltages.

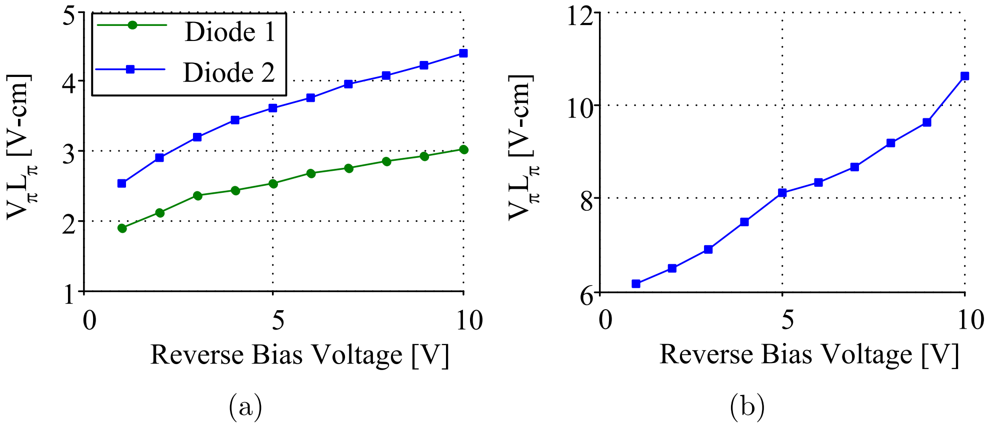


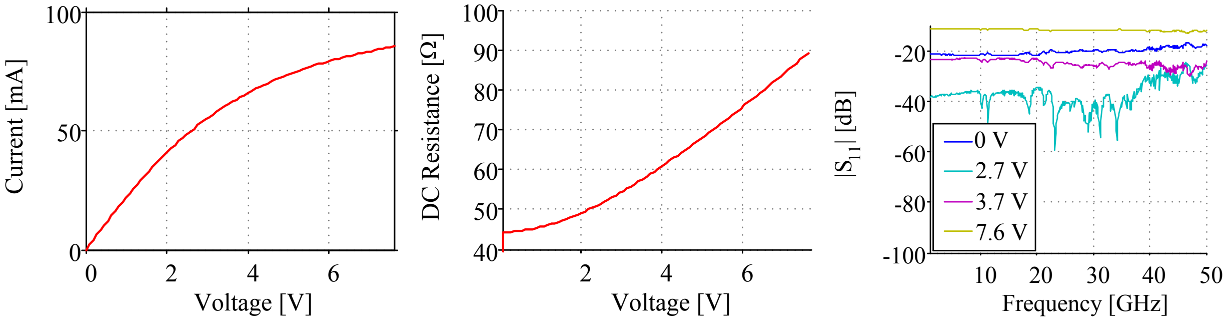
Fig. 5.5 $V_\pi L_\pi$ calculated from transmission spectrum for (a) diode 1, and diode 2, and (b) both diodes under the same reverse bias voltage.

5.3 On-chip Termination - DC and Small signal

On-chip terminations (OCT) can be convenient for larger scale integration of MZMs. OCTs can also be used to introduce peaking into the frequency response of the modulator.

The measurements in this section also show that they can be quite broadband. The measured results of an OCT designed to 50 Ω is shown in Fig. 5.6. The I-V measurements of the OCT are shown in Fig. 5.6(a) and it is observed that the resistor begins to saturate at about 3 V. The source of the current saturation is the velocity saturation of the carriers in the semiconductor, which is reached beyond a certain applied electric field. The current saturation can be pushed to higher voltages by using a wider OCT. The corresponding resistance is shown in Fig 5.6(b). This subfigure depicts that improper matching could result with very high RMS drive voltages. Impedance mismatch will also occur with high reverse bias voltages in single or dual-drive modulators.

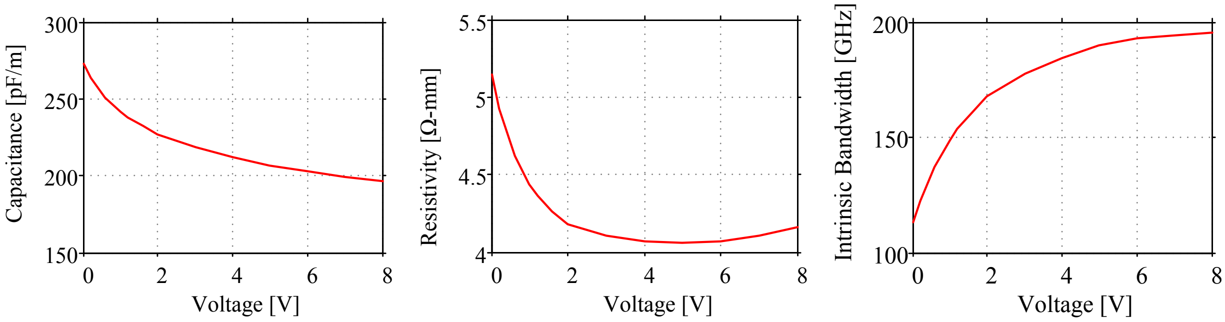
Small signal measurements were performed using a 50 GHz vector network analyzer (VNA) and the S_{11} parameter is shown in Fig. 5.6(c). The voltages in the legend were corrected to account for the resistance of the external bias tee used. It is observed that the lowest reflection was measured for 2.7 V, which corresponds to a DC resistance of 52 Ω. Surprisingly, even at 7.6 V, when the DC resistance is close to 90 Ω, the S_{11} magnitude is very low (near -11.5 dB).



(a) (b) (c)

Fig. 5.6 Measurements of an on-chip termination designed to 50 Ω. Saturation effect is observed in the (a) DC I-V measurement, and the (b) calculated DC resistance. Small signal measurements around different DC voltages is shown in (c).

5.4 p-n Junction



(a) (b) (c)

Fig. 5.7 Extracted p-n junction (a) capacitance, (b) resistance, and (c) the associated intrinsic bandwidth for different reverse bias voltages.

The small signal parameters for the p-n junction was extracted by fitting the simulated S_{11} phase and magnitude of a circuit model to the measured S_{11} parameters, as discussed in [61]. These measurements were taken from a separate test phase shifter with an identical cross-section. In this model, however, only the junction resistance and depletion capacitance was considered (c_j and r_b of the model shown in Fig. 2.17(b)). The junction capacitance decreases with increasing reverse bias because of the larger depletion width at higher reverse bias. The larger depletion width also slightly reduces the resistance because the transverse distance from the contact to the quasi-neutral region is shortened. The extracted capacitance is plotted in Fig. 5.7(a) and the extracted resistance is plotted in Fig. 5.7(b) for different reverse bias voltages. The corresponding intrinsic bandwidth is plotted in Fig. 5.7(c), and an intrinsic bandwidth greater than 100 GHz is observed.

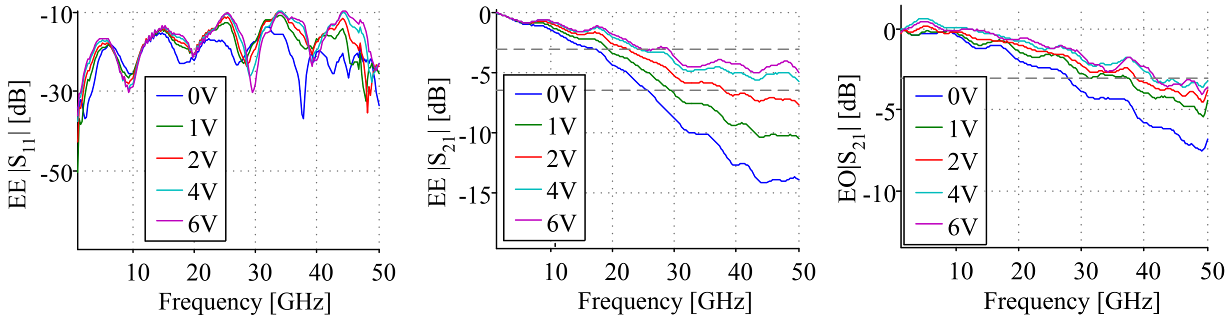
Unfortunately, a test structure with a mirrored p-n junction was not present on the layout. Therefore, the characteristics of the diode in the other arm of the MZM could not be extracted.

5.5 Small Signal Characterization

The RF and electro-optic (E-O) S-parameters were measured using an Agilent 50 GHz lightwave component analyzer (LCA) and 40 GHz GSSG probes. Although the probe is rated for 40 GHz, effects of the higher modes were not observed in the measurements up to 50 GHz. All unused signal ports were terminated with 40 GHz 50 Ω terminations. The probes and cables were calibrated using a Short-Open-Load-Through (SOLT) calibration substrate. The RF power on the LCA was set to -5.0 dBm, which corresponds to 0.356 Vpp in a 50 Ω load. This power is small enough to approximate the non-linear E-O response of carrier-depletion modulators to a linear one.

The measured S_{11} is shown in Fig. 5.8(a). At 0 V bias, the reflection magnitude is relatively low and is below -10 dB over 50 GHz. The reflection magnitude increases for higher reverse bias due to an increase in impedance mismatch resulting from the change in junction resistance and capacitance. Electrical S_{21} normalized at 1.5 GHz for different reverse bias voltages is shown in Fig. 5.8(b). The 6.4 dB point at 0 V bias is 25 GHz, at 1 V reverse bias is 30 GHz, and at 2 V reverse bias is 38 GHz. At 4 V reverse bias, the -3 dB point is 28 GHz, and the 6.4 dB point is beyond 50 GHz.

The measured E-O S_{21} responses, normalized to 1.5 GHz, are shown in Fig. 5.8(c). After each change in bias voltage, the wavelength of the laser was set to bias the modulator at quadrature point to ensure linearity and the best modulation depth. The bandwidth increases with increasing reverse-bias due to a reduction in depletion capacitance at higher reverse bias voltages. Slight peaking is also observed with increasing reverse bias, which is believed to be because of an increase in impedance mismatch. At 1 V reverse bias, the response first approaches the -3 dB point at 33 GHz but only falls below -3 dB at 37 GHz. With 2 V reverse bias, the -3 dB bandwidth increases to 38 GHz. With 4 V reverse bias,

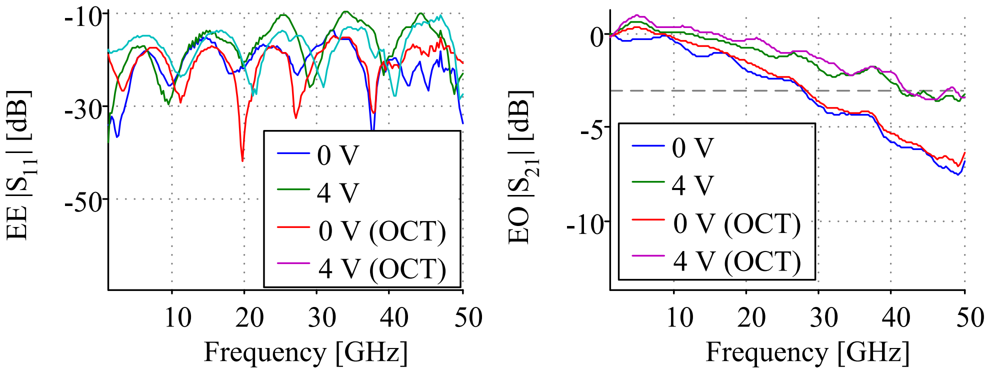


(a) (b) (c)

Fig. 5.8 S-parameter measurements: (a) electrical to electrical S_{11}, (b) electrical to electrical S_{21} normalized to 1.5 GHz (dashed lines indicate -3 dB and -6.4 dB), and (c) electrical to optical S_{21} normalized to 1.5 GHz (solid line indicates -3 dB).

the response crosses the -3 dB point for the first time at 41 GHz. The response also remains near the -3 dB line beyond 50 GHz, where it is only -3.2 dB.

Assuming an average $V_\pi L_\pi$ of 3.2 V-cm at 4 V bias, the RF V_π at 50 GHz is estimated to be 11.4 V. This is one of the lowest 50 GHz RF V_π reported to date in this technology.



(a) (b)

Fig. 5.9 S-parameter measurements with and without the on-chip termination: (a) electrical to electrical S_{11}, and (b) electrical to optical S_{21} normalized to 1.5 GHz (solid line indicates -3 dB).

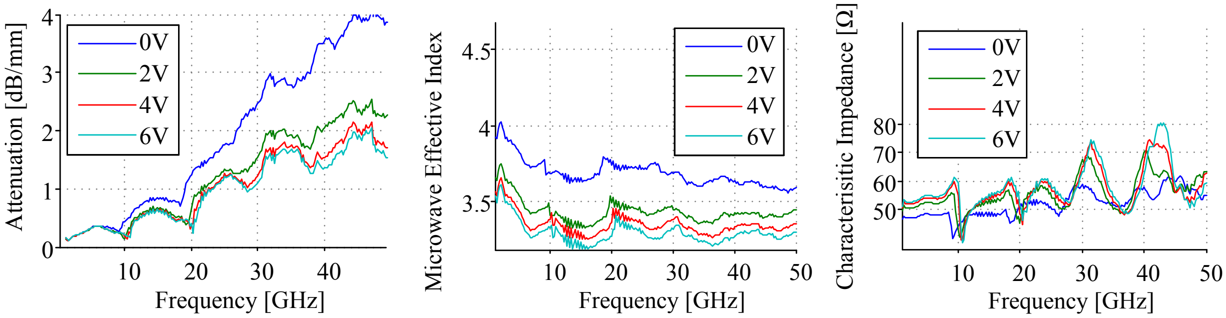
The measured E-O S_{21} responses with and without the on-chip termination, normalized to 1.5 GHz, are shown in Fig. 5.9. The measured reflection and transmission responses are almost identical with and without the on-chip termination despite the deviation in

DC resistance from 50 Ω at 4 V for the OCT. This is because the bias voltage, in a SPP TWMZM, does not drop on the termination resistance and therefore the OCT's resistance remains close to 50 Ω.

In order to verify the attenuation, index and impedance mismatch, these parameters were extracted from measured E-E S-parameters and are plotted in Fig. 5.10. The unloaded portion of the transmission line were de-embedded using simulations from HFSS. Unfortunately, the ripples in the measured S-parameters allowed only for sub-optimal extractions. The attenuation follows a quadratic dependency with frequency, thus experimentally validating the prediction of Eq. (3.53). The frequency at 1.6 dB/mm (6.4 dB/4 mm) from the extracted attenuation is within 10 GHz (this deviation is due to the ripples) of the measured E-E -6.4 dB bandwidth. The microwave index for 0 V at 30 GHz is 3.7, which is close to the simulated optical group index of 3.89 of the rib waveguide. At 4 V bias, the microwave index drops to 3.4 because of the reduced capacitance leading to a faster microwave propagation and the resulting index mismatch is 13%.

A better match, by using a design with a higher slow-wave factor, would result in the -3 dB E-O bandwidth being closer to the 6.4 dB E-E bandwidth, which is at least 50 GHz (the maximum measurable by our LCA) at 4 V bias as shown in Fig. 5.8.

The characteristic impedance is close to 50 Ω at lower frequencies but increases for higher frequencies.



(a) (b) (c)

Fig. 5.10 Extracted (a) attenuation, (b) microwave phase index, and (c) characteristic impedance from measured RF S-parameters.

To further confirm the increase in E-O bandwidth of a design with a slow-wave electrode, the E-O S_{21} response of this modulator is compared with that of a SPP TWMZM with a regular CPS transmission line (no SW). The responses are plotted in Fig. 5.11. The regular CPS TWMZM design has a spacing of 37 µm and a width of 60 µm with an active length that is 0.4 mm smaller than the modulator with the slow-wave electrodes.[1] From simulations, it is predicted than an extra 0.4 mm of junction loading will further decrease its E-O bandwidth by about 5 GHz. At 4 V bias, it is observed that the modulator without SW electrodes has higher peaking. Its response first crosses the -3 dB point at approximately 36 GHz (when normalized to 1.5 GHz). The measurement of that device also shows a parasitic mode being excited at about 48 GHz. At 50 GHz, its response is -5 dB down.

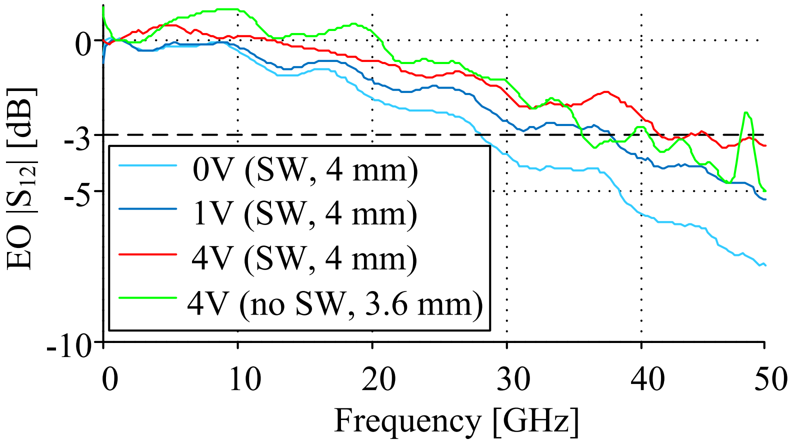


Fig. 5.11 Electro-optic response of this modulator and a shorter design without slow-wave electrodes.

5.6 Large Signal Characterization: OOK Modulation

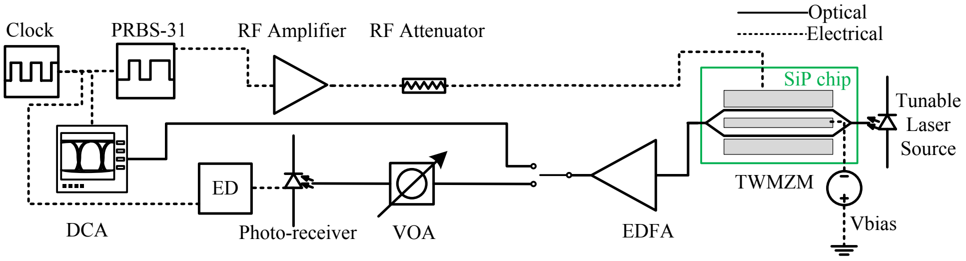


Fig. 5.12 OOK communication link.

---

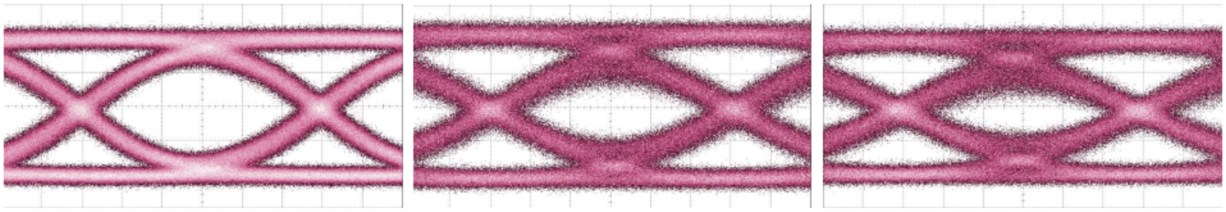
[1] The design, layout, and characterization of that device was performed by my colleague Alireza Samani.

The large signal behavior of the modulator was studied as part of a communication link illustrated in Fig. 5.12. A SHF bit pattern generator (BPG) was used to obtain a PRBS-31 signal. The signal was amplified with a 50 GHz SHF RF amplifier with an AC coupled output. The output of the amplifier was attenuated to limit the driving voltage for different test cases. A DC bias was applied via the on-chip inductance to ensure reverse bias operation. For the optical link, a tunable laser source was used to provide light near 1550 nm and bias the modulator at quadrature. Light was coupled to the silicon photonic chip (SiP) using a fiber array and on-chip vertical focusing grating couplers. For receiver sensitivity measurements, the modulated optical signal was attenuated with a variable optical attenuator (VOA) and then captured with a 35 GHz Picometrix photo-receiver. The photo-receiver was connected to an error detector (ED). Optical eye diagrams were collected by temporarily connecting the modulated light to an Agilent digital communications analyzer (DCA). The modulated optical signal had to be amplified to be able to capture eye diagrams with the DCA. An EDFA was not necessary for BER measurements, however, the absence of the EDFA was observed to induce a BER penalty of approximately 1.8 dB.

The optical eye diagrams at different bitrates with a PRBS-31 signal and a modulator driving voltage of 4.8 V_{pp} is shown in Fig. 5.13. The ER and the Q-factor, as measured by using the DCA with a 65 GHz optical sampling module, are listed in the sub-caption. For all measurements, the modulator was set at quadrature point. Generally, a Q-factor of 7 corresponds to a BER of 1×10^{-12}. The Q-factor reported by the DCA does not de-embed the effect of the DCA's optical sampling module and it is more pessimistic. Eyes with a clear opening are observed up to 60 Gbps.

It should also be noted that ER measurements are highly dependent on the test and measurement equipment and methodology. For instance, the EDFA limited the zero level to a power of -3.7 dBm, thus deteriorating the ER and the Q-factor measured by the DCA.

The increase in extinction ratio for different driving voltages is shown in Fig. 5.14(a). At low bitrates, the extinction ratio improves by 2.5 dB for a drive voltage increasing to 8.1 V_{pp} from 4.8 V_{pp}. It is also observed that the extinction ratio decreases for higher reverse bias voltage, as expected from the DC measurements. An important point to note for SPP modulators is that improper biasing can lead to very high extinction ratio with low drive voltages, as shown by the red curve in Fig. 5.14(a), where the bias pad was left floating (not connected). In this condition, only one diode conducts. Although clean eye diagrams with high ER and Q-factors are observed on the DCA, error-free operation could



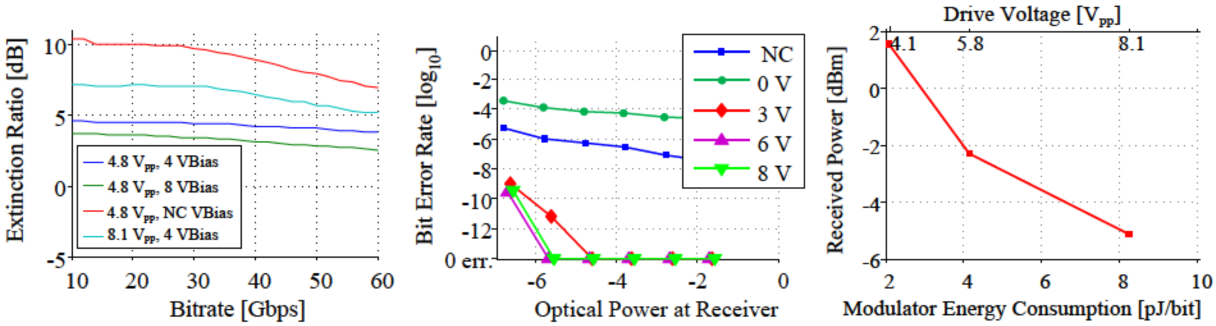
(a) ER: 4.2, Q:7.7
40 Gbps

(b) ER: 3.9, Q:3.9
56 Gbps

(c) ER: 3.8, Q:3.4
60 Gbps

Fig. 5.13 Optical eye-diagrams of MZM at different bitrates with ER in dB and Q-factor in linear units.

not be measured when the bias pad was floating as shown in Fig. 5.14(b).



(a)

(b)

(c)

Fig. 5.14 Effect of drive and bias voltages for extinction ratio and receiver sensitivity. (a) Extinction ratio for diffferent drive and bias voltages. (b) BER measurement for received power at different bias voltages and a drive voltage of 4.8 V_{pp}. (c) Modulator energy consumption for error-free operation at 40 Gbps and the required optical power at receiver.

The sensitivity curves for 40 Gbps is shown for different reverse bias voltages in Fig. 5.14(b). This was obtained with the BPG set at 400 mVpp and a -3 dB attenuator placed at the output of the amplifier. The measured drive voltage is 4.8 V_{pp} in a 50 Ω load. The 0 errors line indicate a measurement with 0 errors detected in 3 terabits, thus allowing us to claim error-free operation (BER $< 1\times10^{-12}$) with a 95% confidence level. The link was also observed to operate error-free up to 45 Gbps with a received power greater than 2 dBm (graph not shown here) and the same drive voltage. The performance of the photo-receiver deteriorates sharply beyond 40 Gbps. At 40 Gbps, error-free operation occurs with a received power of -6 dBm for a reverse bias of 6 V. A larger driving voltage would improve the sensitivity, illustrating a trade-off with power consumption. The required optical power at the photo-receiver for different drive voltages were measured. Fig. 5.14(c) shows the optical power at the receiver necessary for error-free operation at 40 Gbps and the energy consumed by the SPP TWMZM to achieve this.

The BER could not be tested after propagation through fiber due to the lack of a clock recovery circuit. The drifting optimum sampling instant prevented to perform a BER test for a duration required (3 terabits) to claim error-free operation. It should be noted that even after propagating through several kilometers of fiber, the quality of the eye on the DCA degraded only slightly. No errors are observed during BER measurements for brief amounts after syncing.

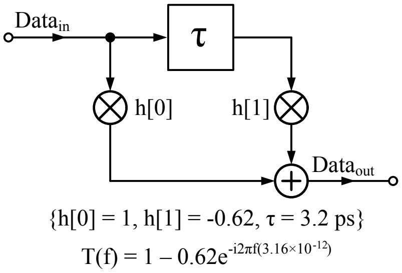
5.7 Large Signal Characterization: OOK Modulation with ASP

Signal processing can be used to achieve high bitrate operation with bandwidth limited components or a bandwidth limited channel. By increasing the strength of the higher frequency content relative to the lower frequency content, the signal after a low-pass medium would no longer have a distorted spectrum. This can be achieved using DSP and digital-to-analog converters. A simpler and low-cost method is to create the pre-emphasis signal using passive analog components.

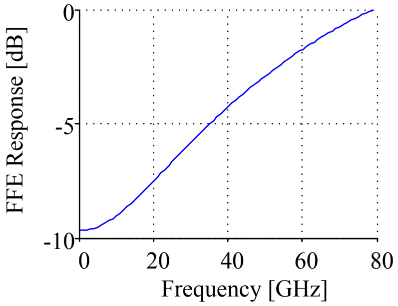
Previously this was done for forward bias MZMs [61, 70]. Here, it is repeated for reverse bias operation to compensate for the limitation of the BPG. The two-tap filter illustrated in Fig. 5.15(a) was created by using the synchronized data from one channel and the inverted data from another channel of the BPG and power combining them. The signal was then amplified. The amplitude of the data signals and the delay between them were controlled with the BPG. The magnitude and delay of the filter taps are shown in Fig. 5.15(a) and the calculated response of the filter is shown in Fig. 5.15(b). This filter is mainly trying to compensate for the low frequency response of the BPG, which is estimated to 28 GHz from rise-time measurements.

The resulting emphasized driving signal had a peak-to-peak voltage of 6.21 V_{pp} and 1.74 V_{rms}.

The optical eye diagrams at different bitrates with a PRBS-31 signal are shown in Fig. 5.16. Comparing the eyes at 60 Gbps of Fig. 5.16 and Fig. 5.13, it is observed that the eye with ASP has very little inter-symbol interference, however, it is noisier. The noise is

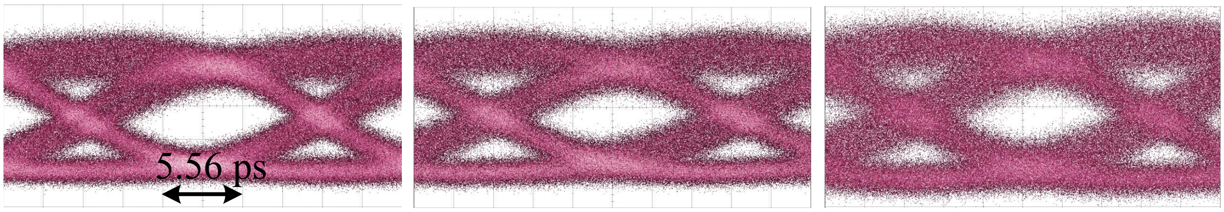


(a) 60 Gbps



(b) 70 Gbps

due to the attenuation from the equalizer used to compensate the BPG. With this passive filtering approach, clear eyes are seen up to 70 Gbps. The extinction ratio at 60 Gbps is 2.1 dB greater than observed in Fig. 5.13 and the Q-factor is slightly better. Even at 72 Gbps, the eye diagram still has a flat level indicating that the performance is not limited by the bandwidth. In fact, Fig. 5.16(a) shows about 5.6 ps of flat levels, corresponding to a bit period of 11.1 ps and 90 Gbps. Higher bitrates could not be tested as the BPG failed to generate signals beyond 72 Gbps due to limitations with the digital circuitry.



(a) ER: 5.9, Q: 3.7
60 Gbps

(b) ER: 5.5, Q: 3.5
70 Gbps

(c) ER: 3.7, Q: 2.6
72 Gbps

Fig. 5.16 Optical eye-diagrams of MZM at different bitrates with analog pre-emphasis.

The limitation of using only two-taps for the pre-emphasis filters is that the lower frequencies can be cut-off leading to baseline wander and pattern dependent errors [61]. Furthermore, with the limited taps, it is not possible to design a filter that would optimize for both noise and bandwidth. These issues can be mitigated with digital signal processing.

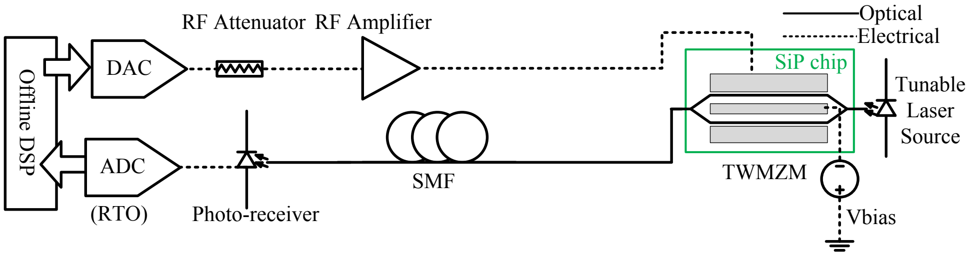


Fig. 5.17 DSP communication link.

5.8 Large Signal Characterization: OOK and PAM Modulation with DSP

This section presents the testing of the modulator using DSP with slight modifications from the method presented in [23, 24]. First, the data to be sent through the modulator was generated using offline DSP. The data pattern was uploaded to the DAC, whose output was connected to an SHF amplifier. The DAC has an 8-bit resolution and was operated at 70 gigasamples per second (GSPS). The DAC has an analog bandwidth specification of >13 GHz. For the PAM-8 modulation format only, an attenuator was placed to avoid gain compression from the SHF RF amplifier. The amplified signal was then connected to a 40 GHz GSSG probe. In this context, the 40 GHz of the probe refers to the high-pass cut-off frequency of the parasitic modes and not the insertion loss. The insertion loss of the probe at 40 GHz is 1 dB. The insertion loss of each cable is approximately 2.4 dB at 40 GHz. The DC bias voltage to the TWMZM was applied through the on-chip inductance operating as an RF choke. On the optical side, a tunable laser source was used, whose wavelength was set to bias the modulator at quadrature. The modulated signal was launched into different fiber lengths of Corning SMF-28e^{+} fiber. On the receiver side, a differential-out 35 GHz Picometrix photo-receiver was used to convert light to the electrical domain. The differential cables were connected to the ADC with skew-matched cables. A real-time oscilloscope with a 33 GHz bandwidth, 8-bit resolution, and a sampling rate of 80 GSPS was used as the analog-to-digital (ADC) converter. In this entire chain, the bandwidth limiting component is the DAC.

The sequential steps of the transmitter offline DSP, diagrammed in Fig. 5.18, is discussed next. First, the data to be transmitted is generated. Instead of generating a binary pattern and then applying a coding scheme, e.g., Gray-coding, the symbols themselves

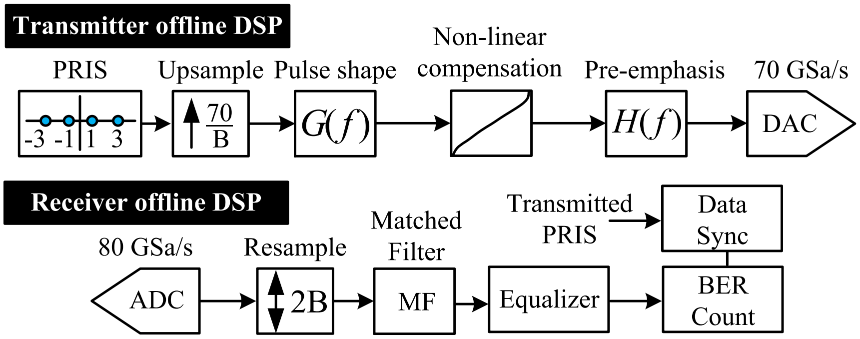


Fig. 5.18 Transmitter and receiver offline DSP.

were randomly generated as integers equally spaced apart and spanning the desired PAM modulation order. Since the DAC was operated at 70 GSPS, this pseudo-random integer sequence (PRIS) pattern had to be up-sampled by inserting $(\frac{70}{B} - 1)$ zeros, where B is the baudrate in gigabaud (Gbd). Afterwards, a raised-root cosine (RRC) pulse-shaping filter $G(f)$ with a roll-off factor parameter $\alpha \in [0, 1]$, was applied. This filter interpolates the data to the DAC's sampling rate and also limits the bandwidth of the data stream to $\frac{1}{2}B(1+\alpha)$. With a larger roll-off factor, a higher bandwidth is required but a larger ISI-free sampling window is obtained. The pulse shaped data was passed through an inverse sine non-linear compensation filter to be able to recover the shape of the data after the MZM sinusoidal transfer function. A pre-emphasis filter $H(f)$ is applied to compensate for the limitations of the entire system. The pre-emphasis filter was obtained from a least-mean-squares algorithm using a training sequence. The processed signal is then quantized to 8-bit and uploaded to the DAC's memory for transmission.

On the receiver side, the data captured by the ADC at 80 GSPS is re-sampled to twice the symbol rate $2\times B$. Then an RRC matched filter, defined at 2 samples per symbol, is applied to the data. Afterwards, the data is equalized with a finite-impulse-response (FIR) filter with coefficients determined from a training sequence. After equalization, the bit error rate was obtained by de-coding the multi-amplitude received and transmitted patterns into a binary stream and then comparing the bits at the optimum sampling point. In the testing setup used for this experiment, the DAC and ADC clocks were synchronized and no digital clock recovery was necessary. In a more realistic scenario, clock recovery with DSP can be applied as demonstrated in [23].

Since the analog performance is mainly limited by the DAC performance, this indicates

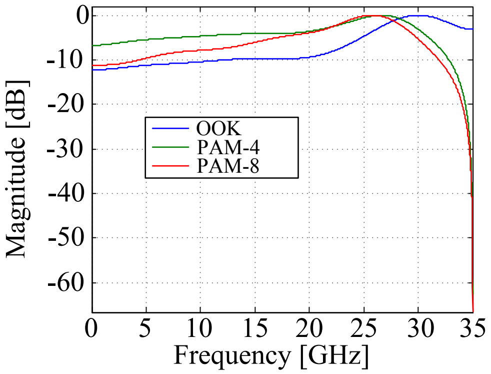
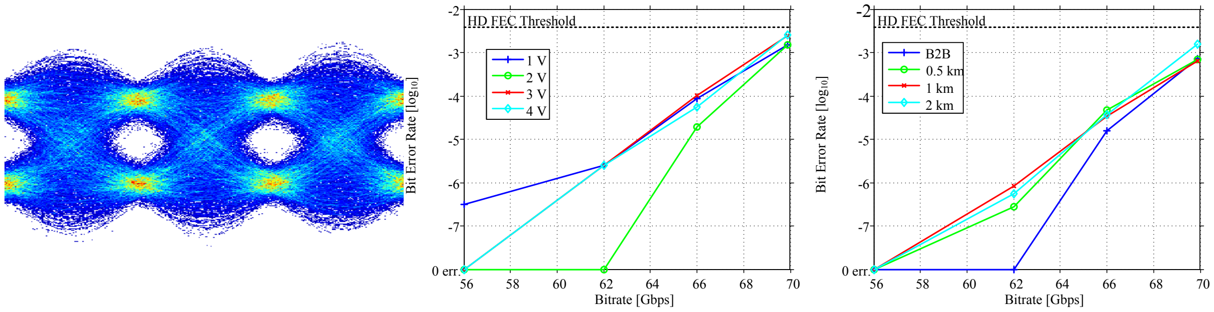


Fig. 5.19 Digital pre-emphasis filters applied at the transmitter for OOK, PAM-4, and PAM-8.

that for a communication link with DSP and error correction coding, low error rate transmission at high bitrates can be obtained with low bandwidth MZMs. However, as DAC and ADC performance will improve, using a TWMZM with a high bandwidth will be beneficial.

In Fig. 5.19, the pre-emphasis filters ($H(f)$ of Fig. 5.18) that were applied at the transmitter for OOK, PAM-4, and PAM-8 are shown. For each modulation format, the filter was optimized for transmission of 2 km but using different baudrates. For the case of PAM-2 (OOK), PAM-4, and PAM-8 the baudrate used for optimization were 70, 56, and 37, respectively. The optimum roll-off factor of PAM-2 (OOK), PAM-4, and PAM-8 were 0.9, 0.1, 0.4. For PAM-2, frequency content from 20 to 35 GHz is emphasized and the lower frequency content remained near -10 dB. For PAM-8, compensation was done from -11 dB near DC up to about 28 GHz. For PAM-8, the pre-emphasis filter appears more like a zero-forcing equalizer (ZFE) up to 25 GHz. For PAM-4, the emphasis is similar to PAM-8, except that the lower frequency components are not attenuated as much. Since the LMS algorithm was used to obtain these filters, it is not only the frequency response of the entire system that is being compensated, as would have been the case with a ZFE. The ZFE would equalize to obtain a large bandwidth and amplify noise by doing so. The LMS algorithm gives a balance of bandwidth-noise with low error rates as the objective, therefore, the pre-emphasis filters are not simply the inverse of the analog response of the DAC.



(a) (b) (c)

Fig. 5.20 PAM-2 modulation results with DSP. Shown in (a) is the eye diagram, (b) BER for different bias voltages, and (c) BER for different transmission distances.

PAM-2 Modulation

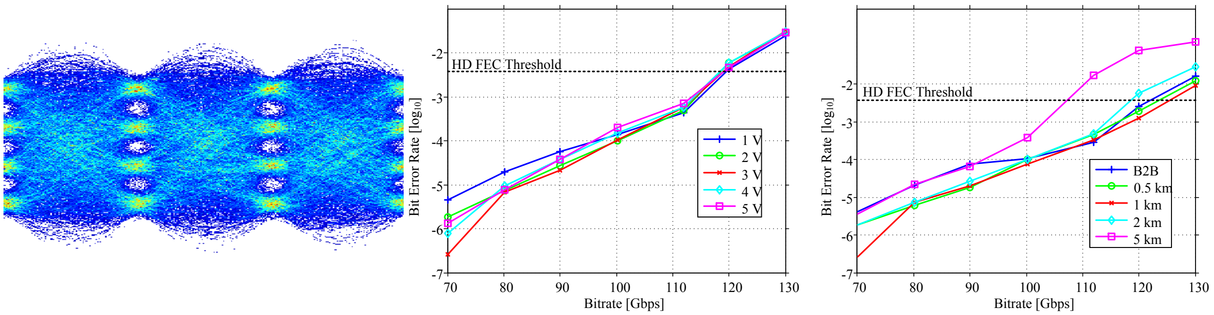
In Fig. 5.20, the results for OOK modulation with DSP are shown. In Fig. 5.20(b), the BER after 2 km of propagation for different reverse bias voltage of the modulator and transmission speeds is shown. The best performance was observed for a reverse-bias voltage of 2 V. At that bias voltage, the experiment was repeated for different fiber lengths. It is observed that there is a slight degradation in BER for the 2 km compared to other distances. This is believed to be due to the chromatic dispersion present at 1550 nm.

The ITU-T G.709 OTU4 standard defines a line rate of 111.809973 Gbps for a payload rate of 99.532800 Gbps. Of the total overhead, 6.7% is for the forward error correction (FEC) code. The ITU-T standards allow FEC coding schemes with this percentage of overhead. It is generally accepted that a pre-FEC hard-decision (HD) BER threshold of 3.8×10^{-3} results in a corrected BER of 1×10^{-15}. This can be achieved with variants of the BCH code or enhanced RS codes. Without coding, the BER at 70 Gbps for 2 km of propagation is 1.57×10^{-3}. Thus, with coding, this can be considered error-free at about 62 Gbps.

The measured drive voltage for OOK with DSP was 4.3 V_{pp} and 0.98 V_{rms}.

PAM-4 Modulation

In Fig. 5.21, the results for PAM-4 modulation with DSP are shown. In Fig. 5.21(b), the BER after 2 km of propagation for different reverse bias voltage of the modulator and transmission speeds is shown. At 112 Gbps, the best performance was observed for a



(a) (b) (c)

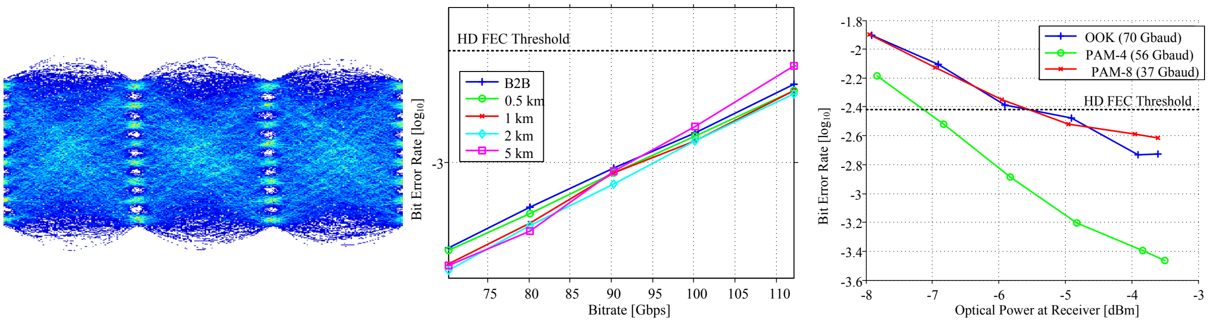
Fig. 5.21 PAM-4 modulation results with DSP. Shown in (a) is the eye diagram, (b) BER for different bias voltages, and (c) BER for different transmission distances.

reverse-bias voltage of 2 V. At that bias voltage, the experiment was repeated for different fiber lengths.

Without coding, the BER at 112 Gbps for 2 km of propagation is 4.92×10^{-4}. Thus, with coding, this can be considered error-free at about 100 Gbps.

The measured drive voltage for PAM-4 with DSP was 5.21 V_{pp} and 0.94 V_{rms}.

PAM-8 Modulation



(a) (b) (c)

Fig. 5.22 PAM-8 modulation results with DSP. Shown in (a) is the eye diagram, (b) BER for different transmission distances, and (c) the sensitivity of the receiver for the different modulation formats after 2 km of propagation.

In Fig. 5.22(b), the BER after propagation for different fiber lengths at a reverse bias voltage of -1 V is shown. Without coding, the BER at 112 Gbps for 5 km of propagation is 3.16×10^{-3}. Thus, with coding, this can be considered error-free at about 100 Gbps. PAM-8 operation was observed to operate below FEC threshold for 5 km of fiber as oppose to PAM-4. This may be partly due to the larger roll-off factor used with the PAM-8 modulation format making it more tolerant to ISI from pulse broadening. The RMS and peak-peak drive voltage for PAM-8 modulation format was also higher than for the PAM-4 case.

The measured drive voltage for PAM-8 with DSP was 6.73 $\mathrm{V_{pp}}$ and 1.32 $\mathrm{V_{rms}}$.

In Fig. 5.22(c), the receiver sensitivity of the three cases are compared after 2 km of propagation. It is observed that for PAM-8 with FEC, error-free operation is obtained for received optical power greater than -5.4 dBm. The optical power requirement for PAM-4 is lower. With coding, PAM-4 transmission at 100 Gbps can be error-free for received powers larger than -7.1 dBm.

Compared to the previously stated BER measurements, the measurements shown in Fig. 5.22(c) are optimized with respect to the power at the receiver and are therefore lower. At a received power of -3.6 dBm and after 2 km of propagation, the BERs for OOK (70 Gbps), PAM-4 (112 Gbps), and PAM-8 (112 Gbps) are 1.57×10^{-3}, 3.42×10^{-4}, and 2.43×10^{-3}, respectively.

5.9 Modulator Energy Estimate

Modulator power energy consumption for OOK NRZ bitstreams was derived in Section 2.4.3. The power consumed by pre-emphasis signals from analog or digital signal processing is highly pattern dependent. For DSP, the drive voltage is is sensitive to the roll-off factor and the clipping ratio applied at the transmitter side. A simple approach to estimate the power consumption is to use the RMS voltage and assume that power is mainly consumed by the termination resistor. Then the dynamic power can be estimated as $P_{dynamic} = V_{rms}^2/R_T$, where R_T is the termination impedance. The RMS voltage was measured by capturing a long pattern of the PRBS signal with a 33 GHz real-time oscilloscope and were mentioned in the previous section. Because of the SPP structure, static power from the bias voltage is ignored.

The power consumed by the modulator for each driving configuration is listed in Table 5.1. For the case without signal processing, the bitrate used is 40 Gbps and the driving voltage corresponds to error-free transmission with a received power of -5.17 dBm. The

DSP cases were also found to be error-free with FEC. Error counting could not be performed for the case with ASP due to the lack of a fast photo-receiver. The bitrate used in the energy estimation is the approximate bitrate of the payload.

Table 5.1 Estimated power consumed by the SPP TWMZM.

Modulation Format	Bitrate [Gbps]	V_{rms}	P_{Dynamic} [W]	Energy [pJ/bit]
OOK	40	4.07	0.33	8.27
OOK (ASP)	60	1.74	0.06	0.96
OOK (DSP)	62	0.98	0.019	0.31
PAM 4 (DSP)	100	0.94	0.018	0.18
PAM8 (DSP)	100	1.32	0.035	0.35

The calculations show that with signal processing the power consumed by the TWMZM is reduced. The energy per bit of the TWMZM is also lower with signal processing because of the higher bitrate. The lowest energy consumption is for PAM-4. PAM-8 had twice the power consumption. This table illustrates how the power consumption is transferred from the modulator to digital electronics, e.g., FPGA/ASICs for digital filtering and equalization, DACs, encoders/decoders for FEC, CDRs, and ADCs. In order to obtain a total power consumption of the entire transceiver, more insight is needed about the power consumption of electronics and its distribution for DSP.

Chapter 6

Conclusion

Silicon photonic offers a high index contrast system for developing dense very large scale integrated circuits. In addition, the reliable CMOS fabrication techniques make it possible to mass produce devices suitable with electronic integration at low costs. The Mach-Zehnder modulator is a key component of optical transmitters, which are becoming universal in datacenters for meeting the bandwidth and energy requirements. In this thesis, the design, analysis, and performance of a silicon photonic traveling wave Mach-Zehnder modulator was presented.

A series push-pull traveling wave MZM was designed to optimize the electro-optic bandwidth with respect to: (1) Microwave and optical velocity matching, (2) impedance matching, and (3) lower microwave loss. The electrodes used T-shaped extensions to increase the microwave index. The modulator design had a characteristic impedance near 50 Ω for matching with common microwave drivers and terminations.

A 37 GHz -3 dB electro-optic bandwidth was measured at 1 V reverse bias. At a reverse-bias voltage of 4 V, the -3 dB electro-optic bandwidth increased to 41 GHz. The DC V_π of the modulator is 7 V in the most efficient arm, which corresponds to a $V_\pi L_\pi$ of 2.8 V-cm. The RF V_π at 50 GHz is estimated to be 11.4 V, which is the lowest 50 GHz RF V_π reported to date in a MPW silicon photonic process.

For on-off keying modulation, open eye diagrams are visible up to 60 Gbps. Error-free operation with a BER $<1\times10^{-12}$ was measured up to 45 Gbps, the limit of our photoreceiver. Eye diagrams with a clear opening were observed up to 70 Gbps using a passive filter to compensate for the driving electronics. It is inferred that the modulator can operate to at least 90 Gbps with better data sources.

The device was further characterized using digital-signal-processing to compensate the analog bandwidth limitations of the entire communication link. Using PAM-4, 112 Gbps transmission on 2 km of SMF fiber is achieved below the pre-FEC BER threshold of 3.8×10^{-3}, corresponding to a post-FEC BER of $<1\times10^{-15}$. With the PAM-8 modulation format, this was achievable up to a distance of 5 km. The lowest energy consumed by the modulator was for the case with PAM-4 using DSP at 0.18 pJ/bit.

The work accomplished in this thesis shows that it is possible to create high bandwidth modulators in a silicon photonic process. It then demonstrates that 100 Gbps modulation in a single wavelength is possible using DSP with lower bandwidth components. Using multiplexing, 400 Gbps and 1 Tbps silicon photonic short reach interconnects can be developed.

Improvements to this design can be made by simply increasing the slow-wave factor using the trade-offs presented in this thesis. As discussed in Section 5.5, the E-E -6.4 dB point is higher in frequency than the E-O -3 dB point. With a better velocity matching, the E-O bandwidth at 4 V bias is expected to increase to at least 50 GHz.

A key factor in improving the performance of modulators in this platform is their modulation efficiency. A higher modulation efficiency will lower the drive voltage necessary to obtain the highest OMA and ER. This will also allow for shorter modulators consequently leading to higher bandwidths. Although much research has been done on this front, the most efficient junctions are still to be made available in MPW runs.

In this thesis, it was also shown that microwave loss is the dominant source of bandwidth limitation. The major source of loss is the capacitance of the p-n junction, followed by the resistance of the junction, the conductor loss, and then the substrate loss. The reduction in loss by optimizing these contributions involves a complicated trade-off space that should studied further. Since efficiency is typically related to capacitance, this is a compromise between microwave loss and modulation efficiency. The resistance of the junction can be lowered at the expense of higher optical loss. The conductor loss may be reduced by using electrodes made in a different material, for instance, copper instead of aluminum, and simulations should be performed to compare their benefit. Substrate loss can be further reduced by using a higher resistivity substrate. Moreover, large bandwidth may not be desired for certain tasks and can be traded off for lower drive voltages. Future work in this field can consist of quantifying these trade-offs and optimizing TWMZMs based on their desired application.

References

[1] Cisco, "Cisco global cloud index: Forecast and methodology, 2013-2018," 2014.

[2] A. Benner, "Optical interconnect opportunities in supercomputers and high end computing," in *Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference*, pp. 1–60, March 2012.

[3] M.-C. Chang, V. Roychowdhury, L. Zhang, H. Shin, and Y. Qian, "RF/wireless interconnect for inter- and intra-chip communications," *Proceedings of the IEEE*, vol. 89, pp. 456–466, Apr 2001.

[4] D. Miller, "Device requirements for optical interconnects to silicon chips," *Proc. IEEE*, vol. 97, pp. 1166–1185, July 2009.

[5] J. M. Shainline, J. S. Orcutt, M. T. Wade, K. Nammari, B. Moss, M. Georgas, C. Sun, R. J. Ram, V. Stojanović, and M. A. Popović, "Depletion-mode carrier-plasma optical modulator in zero-change advanced CMOS," *Opt. Lett.*, vol. 38, pp. 2657–2659, Aug 2013.

[6] "CMC Microsystems. http://www.cmc.ca."

[7] "NSERC CREATE Silicon Electronic-Photonic Integrated Circuits (Si-EPIC) program. http://siepic.ubc.ca."

[8] T. Baehr-Jones, T. Pinguet, P. L. Guo-Qiang, S. Danziger, D. Prather, and M. Hochberg, "Myths and rumours of silicon photonics," *Nature Photonics*, vol. 6, no. 4, pp. 206–208, 2012.

[9] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," *Nature*, vol. 435, no. 7040, pp. 325–327, 2005.

[10] X. Wang, M. Caverley, J. Flueckiger, Y. Wang, N. A. F. Jaeger, and L. Chrostowski, "Silicon photonic Bragg grating modulators," *IEEE Photonics Conference (IPC)*, 10/2014 2014.

[11] L. Gu, W. Jiang, Y. Jiang, X. Chen, and R. Chen, "Photonic-crystal-waveguide-based silicon Mach-Zehnder modulators," in *Integrated Photonics Research and Applications/Nanophotonics*, p. NWA3, Optical Society of America, 2006.

[12] A. Brimont, D. J. Thomson, P. Sanchis, J. Herrera, F. Gardes, J. M. Fedeli, G. T. Reed, and J. Martí, "High speed silicon electro-optical modulators enhanced via slow light propagation," *Opt. Express*, vol. 19, pp. 20876–20885, Oct 2011.

[13] D. Gill, S. Patel, M. Rasras, K.-Y. Tu, A. White, Y.-k. Chen, A. Pomerene, D. Carothers, R. Kamocsai, C. Hill, and J. Beattie, "CMOS-compatible Si-ring-assisted Mach-Zehnder interferometer with internal bandwidth equalization," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 16, pp. 45–52, Jan 2010.

[14] T. Baehr-Jones, R. Ding, Y. Liu, A. Ayazi, T. Pinguet, N. C. Harris, M. Streshinsky, P. Lee, Y. Zhang, A. E.-J. Lim, T.-Y. Liow, S. H.-G. Teo, G.-Q. Lo, and M. Hochberg, "Ultralow drive voltage silicon traveling-wave modulator," *Opt. Express*, vol. 20, pp. 12014–12020, May 2012.

[15] Z. Xuan, Y. Ma, Y. Liu, R. Ding, Y. Li, N. Ophir, A. E.-J. Lim, G.-Q. Lo, P. Magill, K. Bergman, T. Baehr-Jones, and M. Hochberg, "Silicon microring modulator for 40 Gb/s NRZ-OOK metro networks in O-band," *Opt. Express*, vol. 22, pp. 28284–28291, Nov 2014.

[16] W. D. Sacher, W. M. J. Green, S. Assefa, T. Barwicz, H. Pan, S. M. Shank, Y. A. Vlasov, and J. K. S. Poon, "Coupling modulation of microrings at rates beyond the linewidth limit," *Opt. Express*, vol. 21, pp. 9722–9733, Apr 2013.

[17] G. T. Reed, G. Z. Mashanovich, F. Y. Gardes, M. Nedeljkovic, Y. Hu, D. J. Thomson, K. Li, P. R. Wilson, S.-W. Chen, and S. S. Hsu, "Recent breakthroughs in carrier depletion based silicon optical modulators," *Nanophotonics*, vol. 3, no. 4-5, p. 229, 2014.

[18] P. Dong, X. Liu, S. Chandrasekhar, L. Buhl, R. Aroca, and Y.-K. Chen, "Monolithic silicon photonic integrated circuits for compact 100 $^+$ Gb/s coherent optical receivers and transmitters," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 20, pp. 150–157, July 2014.

[19] G. Denoyer, A. Chen, B. Park, Y. Zhou, A. Santipo, R. Russo, and F. Corporation, "Hybrid silicon photonic circuits and transceiver for 56 Gb/s NRZ 2.2 km transmission over single mode fiber," no. 1, pp. 4–6, 2014.

[20] C. Xiong, D. Gill, J. Rosenberg, M. Khater, T. Barwicz, S. Assefa, S. Shank, C. Reinholm, E. Kiewra, J. Ellis-Monaghan, S. Kamlapurkar, A. Stricker, W. Green, Y. A.

Vlasov, and W. Haensch, "A linear push-pull silicon optical modulator," in *Frontiers in Optics 2014*, p. FM3A.4, Optical Society of America, 2014.

[21] D. M. Gill, W. M. Green, C. Xiong, J. E. Proesel, A. V. Rylyakov, C. L. Schow, J. Rosenberg, T. Barwicz, M. Khater, S. Assefa, S. Shank, C. Reinholm, E. Kiewra, S. Kamlapurkar, and Y. Vlasov, "Distributed electrode Mach-Zehnder modulator with double-pass phase shifters and integrated inductors," in *CLEO: 2014*, p. STh4M.6, Optical Society of America, 2014.

[22] R. Ding, Y. Liu, Q. Li, Y. Yang, Y. Ma, K. Padmaraju, A. E.-J. Lim, G.-Q. Lo, K. Bergman, T. Baehr-Jones, and M. Hochberg, "Design and characterization of a 30-GHz bandwidth low-power silicon traveling-wave modulator," *Optics Communications*, vol. 321, no. 0, pp. 124 – 133, 2014.

[23] M. Chagnon, M. Osman, M. Poulin, C. Latrasse, J.-F. Gagné, Y. Painchaud, C. Paquet, S. Lessard, and D. Plant, "Experimental study of 112 Gb/s short reach transmission employing PAM formats and SiP intensity modulator at 1.3 μm," *Opt. Express*, vol. 22, pp. 21018–21036, Aug 2014.

[24] M. Poulin, C. Latrasse, J.-F. Gagne, Y. Painchaud, M. Cyr, C. Paquet, M. Osman, S. Lessard, and D. V. Plant, "107 Gb/s PAM-4 transmission over 10 km using a SiP series push-pull modulator at 1310 nm," in *European Conference on Optical Communications (ECOC)*, 2014.

[25] H. Xu, X. Li, X. Xiao, P. Zhou, Z. Li, J. Yu, and Y. Yu, "High-speed silicon modulator with band equalization," *Opt. Lett.*, vol. 39, pp. 4839–4842, Aug 2014.

[26] R. Ding, Y. Liu, Y. Ma, Y. Yang, Q. Li, A. E.-J. Lim, G.-Q. Lo, K. Bergman, T. Baehr-Jones, and M. Hochberg, "High-speed silicon modulator with slow-wave electrodes and fully independent differential drive," *J. Lightwave Technol.*, vol. 32, pp. 2240–2247, Jun 2014.

[27] X. Tu, K.-F. Chang, T.-Y. Liow, J. Song, X. Luo, L. Jia, Q. Fang, M. Yu, G.-Q. Lo, P. Dong, and Y.-K. Chen, "Silicon optical modulator with shield coplanar waveguide electrodes," *Opt. Express*, vol. 22, pp. 23724–23731, Sep 2014.

[28] D. Marris-Morini, L. Virot, C. Baudot, J.-M. Fédéli, G. Rasigade, D. Perez-Galacho, J.-M. Hartmann, S. Olivier, P. Brindel, P. Crozat, F. Bœuf, and L. Vivien, "A 40 Gbit/s optical link on a 300-mm silicon platform," *Opt. Express*, vol. 22, pp. 6674–6679, Mar 2014.

[29] Y. Yang, Q. Fang, M. Yu, X. Tu, R. Rusli, and G.-Q. Lo, "High-efficiency Si optical modulator using Cu travelling-wave electrode," *Opt. Express*, vol. 22, pp. 29978–29985, Dec 2014.

[30] X. Tu, T.-Y. Liow, J. Song, X. Luo, Q. Fang, M. Yu, and G.-Q. Lo, "50-Gb/s silicon optical modulator with traveling-wave electrodes," *Opt. Express*, vol. 21, pp. 12776–12782, May 2013.

[31] A. Novack, Y. Liu, R. Ding, M. Gould, T. Baehr-Jones, Q. Li, Y. Yang, Y. Ma, Y. Zhang, K. Padmaraju, *et al.*, "A 30 GHz silicon photonic platform," in *SPIE Optics+ Optoelectronics*, pp. 878107–878107, International Society for Optics and Photonics, 2013.

[32] B. Milivojevic, C. Raabe, A. Shastri, M. Webster, P. Metz, S. Sunder, B. Chattin, S. Wiese, B. Dama, and K. Shastri, "112Gb/s DP-QPSK transmission over 2427km SSMF using small-size silicon photonic IQ modulator and low-power CMOS driver," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, p. OTh1D.1, Optical Society of America, 2013.

[33] F. Merget, S. S. Azadeh, J. Mueller, B. Shen, M. P. Nezhad, J. Hauck, and J. Witzens, "Silicon photonics plasma-modulators with advanced transmission line design," *Opt. Express*, vol. 21, pp. 19593–19607, Aug 2013.

[34] M. Streshinsky, R. Ding, Y. Liu, A. Novack, Y. Yang, Y. Ma, X. Tu, E. K. S. Chee, A. E.-J. Lim, P. G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, "Low power 50 Gb/s silicon traveling wave Mach-Zehnder modulator near 1300 nm," *Opt. Express*, vol. 21, pp. 30350–30357, Dec 2013.

[35] D. Marris-Morini, C. Baudot, J.-M. Fédéli, G. Rasigade, N. Vulliet, A. Souhaité, M. Ziebell, P. Rivallin, S. Olivier, P. Crozat, X. L. Roux, D. Bouville, S. Menezo, F. Bœuf, and L. Vivien, "Low loss 40 Gbit/s silicon modulator based on interleaved junctions and fabricated on 300 mm SOI wafers," *Opt. Express*, vol. 21, pp. 22471–22475, Sep 2013.

[36] X. Xiao, H. Xu, X. Li, Z. Li, T. Chu, Y. Yu, and J. Yu, "High-speed, low-loss silicon Mach-Zehnder modulators with doping optimization," *Opt. Express*, vol. 21, pp. 4116–4125, Feb 2013.

[37] H. Yu, M. Pantouvaki, J. V. Campenhout, D. Korn, K. Komorowska, P. Dumon, Y. Li, P. Verheyen, P. Absil, L. Alloatti, D. Hillerkuss, J. Leuthold, R. Baets, and W. Bogaerts, "Performance tradeoff between lateral and interdigitated doping patterns for high speed carrier-depletion based silicon modulators," *Opt. Express*, vol. 20, pp. 12926–12938, Jun 2012.

[38] Y. Zhang, T. Baehr-Jones, R. Ding, T. Pinguet, Z. Xuan, and M. Hochberg, "Silicon multi-project wafer platforms for optoelectronic system integration," in *Group IV Photonics (GFP), 2012 IEEE 9th International Conference on*, pp. 63–65, Aug 2012.

[39] J. Ding, H. Chen, L. Yang, L. Zhang, R. Ji, Y. Tian, W. Zhu, Y. Lu, P. Zhou, and R. Min, "Low-voltage, high-extinction-ratio, Mach-Zehnder silicon optical modulator for CMOS-compatible integration," *Opt. Express*, vol. 20, pp. 3209–3218, Jan 2012.

[40] P. Dong, L. Chen, and Y. kai Chen, "High-speed low-voltage single-drive push-pull silicon Mach-Zehnder modulators," *Opt. Express*, vol. 20, pp. 6163–6169, Mar 2012.

[41] H. Xu, X. Xiao, X. Li, Y. Hu, Z. Li, T. Chu, Y. Yu, and J. Yu, "High speed silicon Mach-Zehnder modulator based on interleaved PN junctions," *Opt. Express*, vol. 20, pp. 15093–15099, Jul 2012.

[42] D. Thomson, F. Gardes, J.-M. Fedeli, S. Zlatanovic, Y. Hu, B. P. P. Kuo, E. Myslivets, N. Alic, S. Radic, G. Mashanovich, and G. Reed, "50-Gb/s silicon optical modulator," *Photonics Technology Letters, IEEE*, vol. 24, pp. 234–236, Feb 2012.

[43] C. DeRose, D. Trotter, W. Zortman, and M. Watts, "High speed travelling wave carrier depletion silicon Mach-Zehnder modulator," in *Optical Interconnects Conference, 2012 IEEE*, pp. 135–136, May 2012.

[44] M. Ziebell, D. Marris-Morini, G. Rasigade, J.-M. Fédéli, P. Crozat, E. Cassan, D. Bouville, and L. Vivien, "40 Gbit/s low-loss silicon optical modulator based on a pipin diode," *Opt. Express*, vol. 20, pp. 10591–10596, May 2012.

[45] F. Y. Gardes, D. J. Thomson, and G. T. Reed, "40 Gb/s high-speed silicon modulator for TE and TM polarisation," vol. 7943, pp. 794319–794319–7, 2011.

[46] D. J. Thomson, F. Y. Gardes, Y. Hu, G. Mashanovich, M. Fournier, P. Grosse, J.-M. Fedeli, and G. T. Reed, "High contrast 40Gbit/s optical modulation in silicon," *Opt. Express*, vol. 19, pp. 11507–11516, Jun 2011.

[47] L. Chen, C. R. Doerr, P. Dong, and Y. kai Chen, "Monolithic silicon chip with 10 modulator channels at 25 Gbps and 100-GHz spacing," *Opt. Express*, vol. 19, pp. B946–B951, Dec 2011.

[48] A. Liu, L. Liao, D. Rubin, J. Basak, Y. Chetrit, H. Nguyen, R. Cohen, N. Izhaky, and M. Paniccia, "Recent development in a high-speed silicon optical modulator based on reverse-biased pn diode in a silicon waveguide," *Semiconductor Science and Technology*, vol. 23, no. 6, p. 064001, 2008.

[49] Y. Cui, "GaAs-based traveling wave electrooptic modulators," Master's thesis, University of Ottawa, 2004.

[50] K. Holger, *Integrated InP Mach-Zehnder Modulators for 100 Gbit/s Ethernet Applications using QPSK Modulation.* PhD thesis, Technische Universität Berlin, 2010.

[51] H. Chen, *Development of an 80 Gbit/s InP-based Mach-Zehnder Modulator.* PhD thesis, Technische Universität Berlin, 2007.

[52] L. Chrostowski and M. Hochberg, *Silicon Photonics Design.*

[53] G. Reed, *Silicon Photonics: The State of the Art.* Wiley, 2008.

[54] E. Palik, *Handbook of Optical Constants of Solids.* No. v. 3 in Academic Press handbook series, Academic Press, 1998.

[55] L. Vivien, F. Grillot, E. Cassan, D. Pascal, S. Lardenois, A. Lupu, S. Laval, M. Heitzmann, and J.-M. Fédéli, "Comparison between strip and rib {SOI} microwaveguides for intra-chip light distribution," *Optical Materials*, vol. 27, no. 5, pp. 756 – 762, 2005. Si-based Photonics: Towards True Monolithic Integration Proceedings of the European Materials Research Society — Symposium {A1} European Materials Research Society 2004 Spring Meeting.

[56] M. Streshinsky, A. Novack, Y. Liu, R. Ding, M. Gould, T. Baehr-Jones, Q. Li, Y. Yang, Y. Ma, Y. Zhang, K. Padmaraju, K. Bergmen, A.-J. Lim, G.-Q. Lo, and M. Hochberg, "A 30 GHz silicon photonic platform: Multi-project wafer shuttles for next-generation optical systems," in *Photonics Society Summer Topical Meeting Series, 2013 IEEE*, pp. 225–226, July 2013.

[57] Y. Zhang, S. Yang, A. E.-J. Lim, G.-Q. Lo, C. Galland, T. Baehr-Jones, and M. Hochberg, "A compact and low loss Y-junction for submicron silicon waveguide," *Opt. Express*, vol. 21, pp. 1310–1316, Jan 2013.

[58] R. Halir, I. Molina-Fernández, A. O.-M. nux, J. G. Wangüemert-Pérez, D.-X. Xu, P. Cheben, and S. Janz, "A design procedure for high-performance, rib-waveguide-based multimode interference couplers in silicon-on-insulator," *J. Lightwave Technol.*, vol. 26, pp. 2928–2936, Aug 2008.

[59] J. C. Mikkelsen, W. D. Sacher, and J. K. S. Poon, "Dimensional variation tolerant silicon-on-insulator directional couplers," *Opt. Express*, vol. 22, pp. 3145–3150, Feb 2014.

[60] H. Yun, W. Shi, Y. Wang, L. Chrostowski, and N. A. Jaeger, "2× 2 adiabatic 3-db coupler on silicon-on-insulator rib waveguides," in *Photonics North 2013*, pp. 89150V–89150V, International Society for Optics and Photonics, 2013.

[61] D. Patel, V. Veerasubramanian, S. Ghosh, A. Samani, Q. Zhong, and D. V. Plant, "High-speed compact silicon photonic Michelson interferometric modulator," *Opt. Express*, vol. 22, pp. 26788–26802, Nov 2014.

[62] S. Selvaraja, W. Bogaerts, P. Dumon, D. Van Thourhout, and R. Baets, "Subnanometer linewidth uniformity in silicon nanophotonic waveguide devices using CMOS fabrication technology," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 16, pp. 316–324, Jan 2010.

[63] L. Chen, P. Dong, and Y.-K. Chen, "Chirp and dispersion tolerance of a single-drive push-pull silicon modulator at 28 Gb/s," *Photonics Technology Letters, IEEE*, vol. 24, pp. 936–938, June 2012.

[64] B. Chmielak, M. Waldow, C. Matheisen, C. Ripperda, J. Bolten, T. Wahlbrink, M. Nagel, F. Merget, and H. Kurz, "Pockels effect based fully integrated, strained silicon electro-optic modulator," *Opt. Express*, vol. 19, pp. 17212–17219, Aug 2011.

[65] R. A. Soref and B. Bennett, "Electrooptical effects in silicon," *IEEE J. Quantum Electron.*, vol. 23, pp. 123–129, Jan 1987.

[66] M. Nedeljkovic, R. Soref, and G. Mashanovich, "Free-carrier electrorefraction and electroabsorption modulation predictions for silicon over the 1-14- μm infrared wavelength range," *Photonics Journal, IEEE*, vol. 3, pp. 1171–1180, Dec 2011.

[67] G. T. Reed, G. Mashanovich, F. Gardes, and D. Thomson, "Silicon optical modulators," *Nature photonics*, vol. 4, no. 8, pp. 518–526, 2010.

[68] X. Wu, B. Dama, P. Gothoskar, P. Metz, K. Shastri, S. Sunder, J. Van der Spiegel, Y. Wang, M. Webster, and W. Wilson, "A 20Gb/s NRZ/PAM-4 1V transmitter in 40nm CMOS driving a Si-photonic modulator in 0.13 um CMOS," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International*, pp. 128–129, Feb 2013.

[69] P. Bhattacharya, *Semiconductor optoelectronic devices.* Prentice Hall, 1994.

[70] S. Akiyama, M. Imai, T. Baba, T. Akagawa, N. Hirayama, Y. Noguchi, M. Seki, K. Koshino, M. Toyama, T. Horikawa, and T. Usuki, "Compact PIN-diode-based silicon modulator using side-wall-grating waveguide," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 19, pp. 74–84, Nov 2013.

[71] H. Jayatilleka, W. Sacher, and J. Poon, "Analytical model and fringing-field parasitics of carrier-depletion silicon-on-insulator optical modulation diodes," *Photonics Journal, IEEE*, vol. 5, pp. 2200211–2200211, Feb 2013.

[72] H. Yu, M. Pantouvaki, J. Van Campenhout, K. Komorowska, P. Dumon, P. Verheyen, G. Lepage, P. Absil, D. Korn, D. Hillerkuss, J. Leuthold, R. Baets, and W. Bogaerts, "Silicon carrier-depletion-based Mach-Zehnder and ring modulators with different doping patterns for telecommunication and optical interconnect," in *Transparent Optical Networks (ICTON), 2012 14th International Conference on*, pp. 1–5, July 2012.

[73] T.-Y. Liow, J. Song, X. Tu, A.-J. Lim, Q. Fang, N. Duan, M. Yu, and G.-Q. Lo, "Silicon optical interconnect device technologies for 40 Gb/s and beyond," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 19, pp. 8200312–8200312, March 2013.

[74] T. Baehr-Jones, R. Ding, A. Ayazi, T. Pinguet, M. Streshinsky, N. Harris, J. Li, L. He, M. Gould, Y. Zhang, *et al.*, "A 25 gb/s silicon photonics platform," *arXiv preprint arXiv:1203.0767*, 2012.

[75] Y. Liu, S. Dunham, T. Baehr-Jones, A.-J. Lim, G.-Q. Lo, and M. Hochberg, "Ultra-responsive phase shifters for depletion mode silicon modulators," *Lightwave Technology, Journal of*, vol. 31, pp. 3787–3793, Dec 2013.

[76] K. Gupta, *Microstrip Lines and Slotlines.* ARTECH HOUSE ANTENNAS AND PROPAGATION LIBRARY, Artech House, 1996.

[77] F. Di Paolo, *Networks and Devices Using Planar Transmissions Lines.* Taylor & Francis, 2000.

[78] R. Simons, *Coplanar Waveguide Circuits, Components, and Systems.* Wiley Series in Microwave and Optical Engineering, Wiley, 2004.

[79] T. Edwards and B. Steer, *Foundations of Interconnect and Microstrip Design.* Wiley, 2000.

[80] C. Nguyen, *Analysis Methods for RF, Microwave, and Millimeter-Wave Planar Transmission Line Structures.* Wiley Series in Microwave and Optical Engineering, Wiley, 2003.

[81] N. Jaeger and Z. Lee, "Slow-wave electrode for use in compound semiconductor electrooptic modulators," *IEEE J. Quantum Electron.*, vol. 28, pp. 1778–1784, Aug 1992.

[82] S. Akiyama, H. Itoh, S. Sekiguchi, S. Hirose, T. Takeuchi, A. Kuramata, and T. Yamamoto, "InP-based Mach-Zehnder modulator with capacitively loaded traveling-wave electrodes," *J. Lightw. Technol.*, vol. 26, pp. 608–615, March 2008.

[83] J. Shin, S. Sakamoto, and N. Dagli, "Conductor loss of capacitively loaded slow wave electrodes for high-speed photonic devices," *Lightwave Technology, Journal of*, vol. 29, pp. 48–52, Jan 2011.

[84] D. Frickey, "Conversions between S, Z, Y, h, ABCD, and T parameters which are valid for complex source and load impedances," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 42, pp. 205–211, Feb 1994.

[85] J. Witzens, T. Baehr-Jones, and M. Hochberg, “Design of transmission line driven slot waveguide Mach-Zehnder interferometers and application to analog optical links,” *Opt. Express*, vol. 18, pp. 16902–16928, Aug 2010.

[86] G. Ghione, *Semiconductor Devices for High-Speed Optoelectronics.* Cambridge University Press, 2009.

[87] H. Yu and W. Bogaerts, “An equivalent circuit model of the traveling wave electrode for carrier-depletion-based silicon optical modulators,” *Lightwave Technology, Journal of*, vol. 30, pp. 1602–1609, June 2012.

[88] R. Spickermann, S. Sakamoto, and N. Dagli, “In traveling wave modulators which velocity to match?,” in *Lasers and Electro-Optics Society Annual Meeting, 1996. LEOS 96., IEEE*, vol. 2, pp. 97–98 vol.2, Nov 1996.

[89] G. Li, T. Mason, and P. Yu, “Analysis of segmented traveling-wave optical modulators,” *Lightwave Technology, Journal of*, vol. 22, pp. 1789–1796, July 2004.

[90] Microwaves101, “Microwaves101 — artificial transmission lines,” 2013. [Online; accessed 13-November-2013].