CHARACTERIZATION AND PERFORMANCE COMPARISON OF LOW-VOLTAGE, HIGH-SPEED, PUSH-PULL AND TRAVELING-WAVE SILICON MACH-ZEHNDER MODULATORS

THESIS

Tristan A. Latchu, Second Lieutenant, USAF

AFIT-ENG-14-M-48

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the United States Government.

AFIT-ENG-14-M-48

CHARACTERIZATION AND PERFORMANCE COMPARISON OF LOW-VOLTAGE, HIGH-SPEED, PUSH-PULL AND TRAVELING-WAVE SILICON MACH-ZEHNDER MODULATORS

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Electrical Engineering

Tristan A. Latchu, B.S.E.E.

Second Lieutenant, USAF

March 2014

AFIT-ENG-14-M-48

CHARACTERIZATION AND PERFORMANCE COMPARISON OF LOW-VOLTAGE, HIGH-SPEED, PUSH-PULL AND TRAVELING-WAVE SILICON MACH-ZEHNDER MODULATORS

Tristan A. Latchu, B.S.E.E.
Second Lieutenant, USAF

Approved:

//SIGNED//	28 Feb 2014
Maj Michael C. Pochet, PhD (Chairman)	Date
//SIGNED//	13 Mar 2014
Dr. Nicholas G. Usechak, PhD (Member)	Date
//SIGNED//	10 Mar 2014
Dr. Michael A. Marciniak, PhD (Member)	Date

AFIT-ENG-14-M-48

Abstract

The well-known power and memory walls are recognized as the current bottlenecks in computing performance, and with the increasing computational load of commonly run applications, it is necessary to find ways to alleviate the issues presented by the aforementioned bottlenecks. It is therefore necessary to not focus solely on extracting performance improvement by way of changes to the processing architecture, but rather by holistically improving the computing platform, namely the communications backbone. This work focuses on the characterization and performance comparison of two families of optical data modulators, both fundamentally Mach–Zehnder modulators (MZMs); namely, a modulator with a Push-Pull (PP) modulation scheme, and another modulator with a Traveling-Wave Electrode (TWE) design, prioritizing the effects on high-speed performance.

A number of operating conditions — temperature, wavelength, bias voltage, and bit rate — were varied to determine effects of modulator performance, measured in terms of DC performance characterization, bit error rate, electrical bandwidth, and power-penalty. Custom fixtures were designed and fabricated to achieve long-term experimental stability, and software was written to accomplish long term experimentation; the confluence of the two resulted in a wealth of data for use in performance comparison.

Despite the use of a push-pull modulation scheme, the devices using a traveling-wave electrode outperformed the push-pull modulators in almost all metrics, even at what was assumed to be a relatively low bit rate of 3 Gbps. This work then accentuates the importance of velocity and impedance matching, even at presumably low data rates, in spite of increased device fabrication complexity.

Acknowledgements

This thesis was by no means a solitary composition, but instead the collaborative product of a great number of minds, whose sole mission over this past year seems to have been to rescue me from my own lunacy and inaccuracies. The words 'thank you' seem insufficient, so instead I will do my best to dress them up with literary disguise; however, the message is and remains, thank you kindly for your support, encouragement, and assistance.

From the start, my faculty advisor Major Michael Pochet saved me from future peril by paring down my thesis, initially scoped as somewhere between saving the world and curing cancer, to silicon optical data modulators. The advisor-student relationship, I believe, can make or break a thesis. Maj Pochet was the type of advisor I needed, as he let me go off and explore the research areas into which I wished to venture, and reeled me in when I was straying too far for my own good. He was always there to answer my incessant, and on occasion irrational, questions, and entertain my folly by providing what resources were his to give. If this thesis strays in any way from its expected course, it is because I did not hear you correctly.

Dr. Nicholas Usechak is yet another person who deserves my wholehearted appreciation. There has not been a problem I have encountered during the course of this work for which he has not been able to provide a countably infinite set of solutions, or courses of action. Dr. Usechak was the fount of many a brilliant idea during the course of this work, namely the automation software which significantly eased the burden of my data-collection tasks. I must also express my gratitude for the use of your lab, your equipment, and some percentage of your brain's processing power every time I visited your office, and my refusal to leave the premise in less than an hour.

Rounding out my thesis committee is Dr. Michael Marciniak, a physicist by recent trade, but a necessary piece of the puzzle for an electrical engineering student struggling in the ways of the academic world. Dr. Marciniak was able to step back and look at the research in ways in which I had never bothered, and his insight and perspective raised many an intriguing question, saving me from the multitudinous pitfalls towards which I vectored myself. He also served as the inspiration for an entire chapter of this thesis, for which I am completely indebted — were it not for you, I would have been forever embarrassed at the work I put forth.

There was a certain group of people on base who saw me far more than they would most likely have cared, and put up with my ridiculous demands and absurd timelines — the top notch machinists at the AFIT Model Fabrication Shop: Christopher Harkless, Daniel Ryan, and Brian Crabtree. Quite honestly, without your efforts, this thesis would have been nothing short of dismal – both in scope and effort. I shudder to think what would have been, had you not fabricated the mounts used during the course of this work. You are wonderful folks, and I thank you sincerely for your work, for teaching me what you could during my stint at the machine shop, and for ensuring that I still have as many digits as when I was born, despite my best efforts.

Irrespective of whether you have been expressly named in the preceding paragraphs, all of you in my life during this previous year has had an indelible effect on me, and have contributed to the conclusion of this thesis; whether it be to calm my impending insanity, or to help get me off my quest for perfection, and just submit my thesis. To you all, I am much obliged. Any mistakes are of course mine, and the accuracies are theirs.

Tristan A. Latchu

Table of Contents

Page

List of Figures

Figure Page

List of Tables

CHARACTERIZATION AND PERFORMANCE COMPARISON OF LOW-VOLTAGE, HIGH-SPEED, PUSH-PULL AND TRAVELING-WAVE SILICON MACH-ZEHNDER MODULATORS

I. Introduction

Computational performance has made incredible gains since the advent of the computer — with computers having taken many form factors — from portable smartphones to the larger high-performance computing clusters, performance has increased significantly in every category. Benchmarks on computational performance taken over a number of years are shown in Figure 1.1. Initially, the large-scale growth in computing performance was due to hardware architectural advancements, such as reductions in feature size, increases in transistor count, and an increase in clock speeds. The trend to increase processor performance by way of increasing clock speeds has largely passed though, as manufacturers have been unable to effectively deal with issues of thermal management and heat dissipation brought on by the large number of transistors on a single die, and the speed at which they are switching. Between the years of 1986 and 2003, clock speeds increased by 40 % per year, and it was not until 2003 that clock speed increases slowed down to nearly 1 % per year, as seen in Figure 1.2 [1]. Rather than continuing to extract performance improvements by means of increasing clock speeds, manufacturers have instead turned to parallel processing and increasing the number of cores per die. Amdahl's Law, however, dictates that ultimately if communications networks are not improved, the limiting factor and bottleneck in computing performance will be the ability to quickly and accurately move data between processing cores and nodes [2]. Therefore, improving the processor alone is insufficient, and the communications backbone must be considered as vital a component

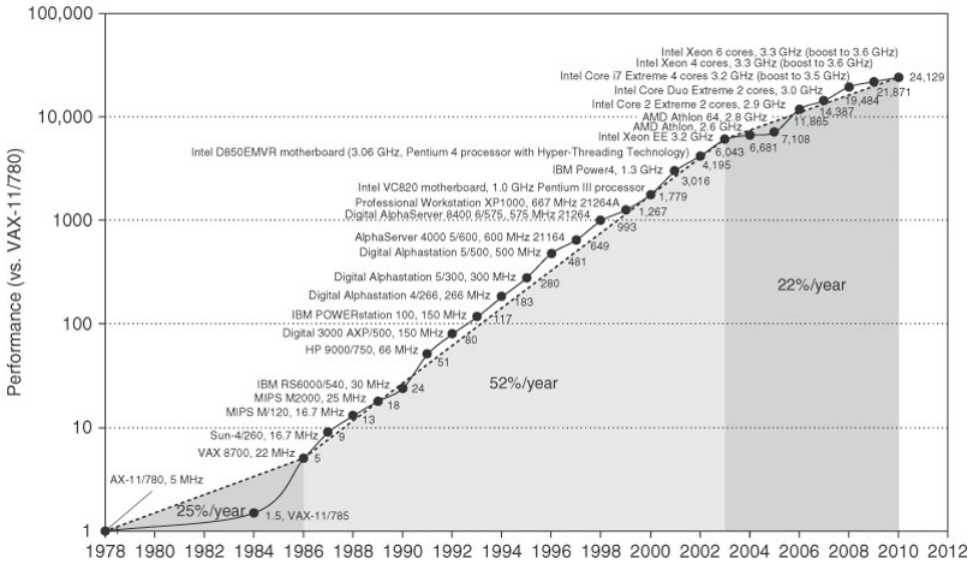


Figure 1.1: Computing performance, when benchmarked, has grown significantly since the inception of the microprocessor. [1]

as any other. The photonics field stands poised to deal with the bandwidth and power consumption issues presented by modern communications networks, both on and off chip.

1.1 Motivation

The photonics field has recently been the subject of a great deal of research, as the need for integrated small scale, efficient, electro-optics has made itself apparent. While the cost to install the fiber backbone is the dominant expense in long-distance communications networks, the primary cost in short-distance communications networks is the large number of integrated photonics components. Elementary semiconductors, compound semiconductors, silicates, and rare-earth doped glasses are the major players in the photonics market today, but silicon in particular is well-suited to the task, as devices constructed of such a material can be easily and monolithically fabricated using existing Complementary Metal-Oxide-Semiconductor (CMOS) and Very Large Scale Integration

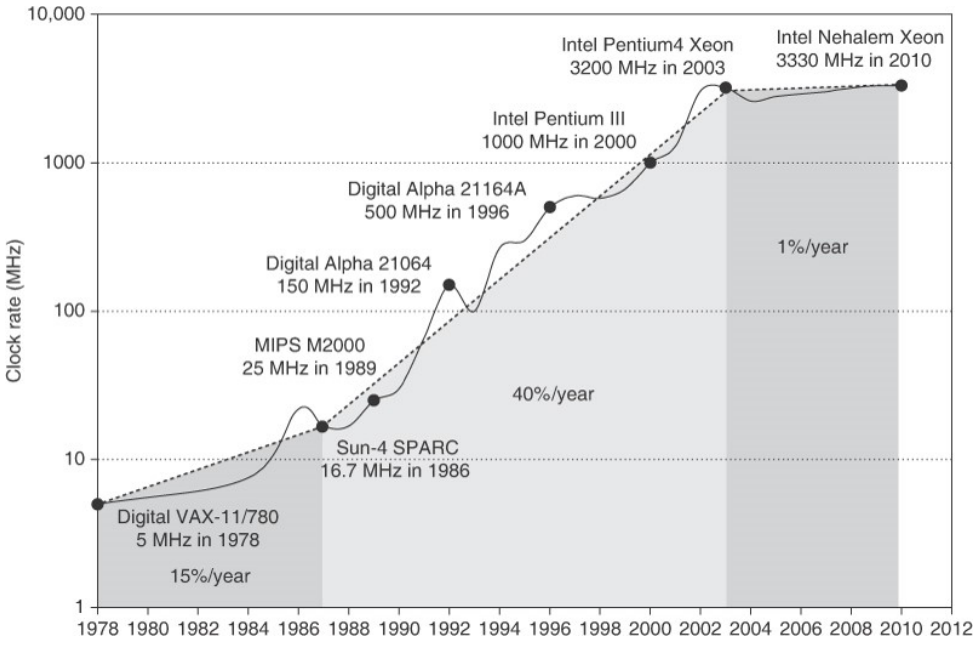


Figure 1.2: Year over year change in processor clock speeds, showing that clock speeds have now nearly leveled out. [1]

(VLSI) technologies and facilities. Additionally, silicon wafers have the lowest cost per unit area, making them less expensive to produce than other semiconductor photonic components. Add to that the minor degree to which defects appear in the crystalline structure of silicon, and the result is higher yields for devices manufactured on silicon [3].

Silicon's intrinsic bandgap of 1.1 eV makes it an excellent photodetector at wavelengths shorter than 1200 nm; however, the fact that typical telecommunications occur at wavelengths of 1310 and 1550 nm, coupled with the fact that silicon has a high index ($\sim$ 3.5), indicate that silicon could be used as a waveguide for telecommunications [4] [5]. With a high index contrast between silicon and its counterpart, silicon dioxide cladding ($\sim$ 1.5), bend radii can be reduced such that it is on the order of the wavelength of light propagating within the waveguide. Compared with the bend radii of more traditional silicates, this is a decrease of four orders of magnitude, with a corresponding increase

in component density by eight orders of magnitude [6]. Waveguide dimensions can correspondingly be shrunk while still maintaining strong optical confinement, leading to device dimensions that are on par with requirements for integrated circuit (IC) processing [3]. With the ever increasing need for higher bandwidths for use in on- and off-chip interconnects, the high-power requirements of communications links in high performance computing, the inherent limitations on communication lengths over copper, their associated losses (skin effect and dielectric losses from the PCB substrate), and susceptibility to cross-talk, the need for a full suite of integrated electro-optics continues to grow. Furthermore, the transition to smaller process sizes for transistors has shifted performance bottlenecks to electrical interconnects, which suffer from RC delay, signal distortion, and high power consumption. Replacing these electrical interconnects with optical interconnects can reduce signal distortion suffered at high frequencies, or over long distances, by reducing latency, skew, and jitter [7]. Ideally, such a suite would include silicon photon sources (lasers), photodetectors, modulators, and waveguides. Unfortunately, the lack of a direct bandgap in silicon has resulted in a dearth of silicon lasers, one of the most promising being optically-pumped silicon Raman laser demonstrated by Boyraz and Jalali [8]. A number of passive and active components have been implemented in silicon however, such as: couplers, polarization beam splitters, filters, waveguide crossings, and modulators. The aggregate effect of this wealth of silicon integrated photonic devices is such that these components can be tightly integrated in larger systems to achieve complex functionality [6]. Silicon also has a higher damage threshold and thermal conductivity when compared to conventional materials, roughly one order of magnitude higher in both cases, making it better suited to use in photonics [3].

The use of silicon-on-insulator (SOI) substrates also presents an advantage in that the CMOS industry has a wealth of experience using SOI substrates for fabrication, and large, nearly defect-free wafers can be readily produced [7]. That said, the standard optical

lithography process can be used to define features for full or partial etching of silicon films. This ultimately results in high yields per wafer, and uniform performance across wafers and production runs [9].

1.2 Research Focus

While a number of photonic components would be necessary for the creation of an integrated electro-optic chip, data modulators seem to present the largest challenge, as they consume the most power and have largely lagged behind the performance of III-V semiconductors or the more traditional lithium niobate modulators. This work focuses on the experimental characterization of two families of Mach–Zehnder Modulators (MZM), that vary in transmission electrode design and modulation scheme, with an emphasis on high-speed data performance. Specifically, two of the MZMs tested utilized a lumped-element electrode and a push-pull modulation scheme, and only varied in active region length — one MZM at 1000-μm, and the other at 2000-μm. The third and final MZM tested used a traveling-wave electrode and a single-arm modulation scheme at an effective active region length of 1500-μm.

In preparation for experimentation, a custom experimental fixture was designed and fabricated to hold the silicon chips that contain the devices under test. Additionally, a program called AutoLab was written in C# to automate the collection of large datasets over an extended period of time. Experimentation started with characterizing device performance at DC, the data for which was then used to design the experiments run at high-speeds. High-speed experimentation consisted of determining the Bit Error Rate (BER) at different data rates, using different test patterns, and at different levels of received optical power. Additionally, the device bandwidth was measured using Scattering Parameters, which are familiar to those in microwave electronics.

1.3 Thesis Organization

This thesis consists of six chapters, in addition to the present chapter, as well as appendices. Chapter Two discusses previous work in the field of optical communications, with a focus on optical data modulators implemented in silicon. Starting with a macroscopic perspective on silicon optical data modulators, the chapter begins by introducing a number of devices which have been used as optical modulators in silicon photonics, and then narrows to specific technologies used in MZMs. Each component necessary to assemble an MZM in silicon is discussed, from power splitters and couplers to the active region, which implements the necessary phase delay between arms. Performance metrics are also discussed, as well as other factors that may affect the performance of the MZM.

Chapter Three then moves on to discuss the principles by which the experiments were designed to accurately characterize device performance. It includes an explanation of the theory and intent behind each type of experiment, as well as the importance of each type of experiment. The chapter also includes an in-depth discussion of the preparation for the experiments, which is to say the design of the experimental fixture and the program AutoLab. The chapter then concludes with the specific equipment and setup for each experimental type.

Chapter Four includes the results and analysis of the experimentation conducted during the course of this research. The discussion of the push-pull MZMs and the traveling-wave electrode MZM is segregated — the two push-pull MZMs are discussed initially, and the traveling-wave electrode device is discussed afterwards. The discourse on the traveling-wave electrode MZM also includes a quantitative performance comparison to the push-pull family of MZMs. The aforementioned DC and high-speed experiments are all discussed independently for the push-pull and traveling-wave electrode MZMs, figures and statistics are then provided for each experiment. The sheer number of factors involved

in device operation precluded the possibility of producing rigorous a predictive nonlinear multi-variate model within the time frame given for this thesis.

Chapter Five utilizes modeling and simulation techniques to gain insight into the performance and operation of the MZMs, with a strong focus on the TWE devices, and operation at high-speeds. Fundamental device characteristics such as impedance, optical index and microwave index are calculated, and attempts are then made to simulate device performance in the presence of an applied microwave signal.

Chapter Six concludes with a brief summary of the findings, and then goes on to suggest potential areas for future research on these devices, with a large focus given to temperature-related effects and variations in device performance at high-speeds. Modeling and extracting device parameters from the collected data would also be beneficial, as suggestions could be made for future iterations of these MZMs, which would then result in enhanced performance. It would also be interesting to characterize and analyze identical device designs taken from different fabrication batches, in an effort to see the degree to which variation in fabrication affects device performance, i.e. to determine device fabrication tolerances and the effects of process variation.

Lastly, the appendix includes the derivation of equations presented in Chapter Two, as well as additional data not included in Chapter Four.

II. Literature and Theory Review

2.1 Theory

The function of a modulator is to modify a carrier waveform by encoding it with a high-speed electrical data signal. An optical modulator provides an altered optical signal output based upon an input electrical signal and a continuous input optical beam. An ideal optical modulator would have the properties of low insertion loss and large extinction ratios under high-frequency operation [7]. In optical communications, it is often the simplest form of amplitude shift-keying that is used, On-Off Keying (OOK), which takes the form of intensity variations on the output optical waveform. At the most fundamental level, the process of modulation is nothing more than altering the number of photons in a waveform.

Two general classifications of modulation methods are direct and indirect modulation. Direct modulation modulates the intensity of the laser diode itself, typically by altering the current to the laser diode, thereby sending the laser above and below its threshold value. Indirect modulation schemes use a continuous wave (CW) laser, and then inserts a device in line with the laser to modulate the optical waveform. Certainly this can be done by a variety of means, each of which may possess characteristics that make it better suited to a particular application: switching, absorption, resonance, and interferometry. Demonstrated by Silberberg, Perlmutter and Baran in the late 1980s, the digital optical switch can preferentially switch an input signal to one of two output arms with an almost step-like nature, based on the application of a voltage across a particular arm [10]. Constructing a silicon electro-absorption modulator (EAM) is not a straightforward process, as silicon is an indirect bandgap material, and as such the absorption cross-section is low when compared to III-V materials. Modulation via absorption can be achieved by either growing germanium on silicon, or bonding III-V materials to silicon. The EAM then makes use of the Franz-Keldysh effect, which describes a change in the

absorption coefficients of bulk semiconductor as the voltage through that semiconductor is changed, or the Quantum Confined Stark Effect (QCSE), which makes use of a bias voltage to modify the absorption spectra of a quantum well [11]. The optical waveform to be modulated can also be placed inside a resonant cavity, and the resonant modes of the cavity modified appropriately by the electrical data signal to produce OOK. The fact that resonant cavities can exhibit incredibly large quality factors means that small deviations in optical path length in the cavity can entirely extinguish a mode's output coupling. Resonant cavity devices have been demonstrated with power efficiencies of less than $10\,\mathrm{fJ\,bit^{-1}}$, voltage levels comparable to CMOS logic levels, and data rates in excess of 50 Gbps [6] [12]. Unfortunately, this exceptional performance due to high Q also has the effect of making a resonant cavity sensitive to temperature variations, fabrication imperfections, and introduces a strong polarization dependence [13]. By way of example, the thermo-optic effect of $2 \times 10^{-4}\,\mathrm{K^{-1}}$ in silicon at near-IR wavelengths induces a modulation of roughly $5\,\%\,\mathrm{K^{-1}}$ for a resonant cavity with a quality factor of $\sim 10^3$ Furthermore, resonant cavities are inherently narrow-bandwidth devices, and although thermal heaters can be used to alter the resonant frequencies, their addition increases device complexity and power usage.

As it applies here, the technique of interest is modulation via interference, or more specifically, achieving modulation with a Mach–Zehnder Modulator (hereafter referred to as an MZM). MZMs stand as a promising technology in modulating a data stream onto an optical source, as they convert a phase modulation into an amplitude modulation. Additionally, MZMs are inherently broadband optical devices, limited primarily in optical bandwidth by the performance of the optical power splitter/coupler. Moreover, MZMs have a low temperature sensitivity, as any external thermal effects are applied equally to both arms so that there are minimal deleterious effects on modulation performance [14] [15] [7]. The typical rib-guide MZMs realized in silicon have a device footprint of several millimeters, making them far larger than desirable for use in large-scale integration [16].

Additionally, the relatively low values for mobility, coupled with the long electron and hole lifetimes in silicon, lead to slow switching speeds when compared to faster III-V compound semiconductor devices. These two points serve as the backstop for further research into silicon optical modulators. In order to achieve high-speed modulation, electrical and optical characteristics such as resistance, capacitance, extinction ratio, and insertion loss need to be optimized to maximize the modulator bandwidth [7].

2.1.1 Methods of Modulation in a Mach–Zehnder Modulator.

An MZM is typically built using a 3-dB splitter to send the input beam through two arms, one which will not be modified, and the other which will have some phase shift induced. The two arms are then recombined, where the two beams interfere appropriately, which results in amplitude modulation. Alternatively, both arms can have a phase shift induced in them, with the phase advanced in one arm, and the phase retarded in the other. In the ideal case, when there is no phase change between the two arms, the full input intensity is recovered at the output after the light is recombined – this state is known as the "on" state. In the "off" state, there is a π phase shift between the two arms. When the light is recombined, there is complete destructive interference, which results in zero intensity at the output. The transmission intensity at the output of the MZM can be expressed as

$$T = \frac{c\varepsilon_r\varepsilon_0}{n}\left[E_0^2\, e^{-2(\alpha_0 L+\Delta\alpha L)} + E_0^2\, e^{-2(\alpha_0 L+\Delta\alpha L)}\cos\left(\frac{2\pi n_{\text{eff}}(V)L}{\lambda_0}\left(\Delta n_{\text{eff1}}(V) - \Delta n_{\text{eff2}}(V)\right)\right)\right] \tag{2.1}$$

where E_0 is the electric field's amplitude at the input of the MZM, n_{eff} and Δn_{eff} are the effective refractive indices of the waveguide and the change in effective index due to modulation (with the subscripts 1 and 2 indicating the two different arms), and λ_0 is the free-space wavelength of the electromagnetic radiation. In this equation, α_0 and $\Delta\alpha$ are the optical loss per unit length of the waveguide and the change in optical loss due to modulation and L is the length of the modulation section. See Appendix B for a full derivation of the above equation.

It is important to note that an ideally zero output intensity will only be achieved if the optical power in both arms is equal, and the power splitters/combiners operate ideally. In such a case, the maximum transmitted intensity of the MZM is given by the following equation

$$T = \frac{2c\varepsilon_r\varepsilon_0}{n} E_0^2 \, e^{-2(\alpha_0 L + \Delta\alpha L)} \tag{2.2}$$

Under this ideal case, the output intensity varies from its maximum in Equation 2.2 to a minimum of 0. It then stands to reason that for a given voltage applied to the MZM, there is a necessary minimum length for the arms of the MZM to achieve a full π phase shift – that length is known as L_π, and its associated voltage as V_π. The two are related by the following equation:

$$L_\pi = \frac{\lambda}{2 \cdot \Delta n_{\text{eff}}(V)} = \frac{\lambda_0}{2n \cdot \Delta n_{\text{eff}}(V)} \tag{2.3}$$

where λ is the wavelength in the medium.

The phase modulation in the signal arms is generally achieved using one of two methods: the plasma-dispersion effect or the electro-optic effect; the first of which describes the change in refractive index with a change in free-carrier concentration, and the second a change in refractive index upon the application of an electric field. Plasma (free-carrier) dispersion is readily achieved in silicon; however, unstrained pure silicon crystal displays no appreciable Pockels effect, and the Kerr and Franz-Keldysh effects are weak, as is to be expected with higher order, non-linear optical processes. The lack of a Pockels effect can be attributed to crystalline symmetry in silicon, and more specifically that silicon is a centrosymmetric crystal, and therefore $\chi_{ijk}^{(\text{even})} = 0$ (where χ_{ijk} are elements of the nonlinear optical susceptibility tensor, and the superscript 'even' designates the even order processes) [17]. The electro-optic effect can be utilized if the MZM is constructed in a photonic crystal, where the defects of the photonic crystal are infiltrated with a material that does display an appreciable electro-optic coefficient, and that material is then used to alter the effective index of the waveguide. Although the thermal effect can also be used to

alter the free-carrier concentration, it acts on the order of milliseconds, and as such cannot be used in high-speed modulators. It can and will appear as a parasitic effect during normal operations though, and will be proportional to the amount of electrical power dissipated by the modulator [7]. The primary focus here though, is the use of the plasma-dispersion effect, the governing equation of which was given by Soref and Bennett at 1550 nm as

$$\Delta n = -(8.8 \cdot 10^{-22}\Delta N + 8.5 \cdot 10^{-18}(\Delta P)^{0.8}) \tag{2.4}$$

where ΔN and ΔP represent the change in carrier concentrations for electrons and holes, respectively [18]. The use of the plasma-dispersion effect also has the unfortunate drawback of inducing optical free-carrier absorption, which behaves according to

$$\Delta\alpha = 8.5 \times 10^{-18}\Delta N + 6 \times 10^{-18}\Delta P \tag{2.5}$$

where ΔN and ΔP represent the change in carrier concentrations for electrons and holes, respectively, and $\Delta\alpha$ is the change in the absorption coefficient [19]. The above equations came from the Drude-Lorenz model in silicon, the formulas for which are

$$\Delta n = \frac{-q^2\lambda}{8\pi^2c^2\epsilon_0 n}\left(\frac{\Delta N}{m_e^*} + \frac{\Delta P}{m_h^*}\right) \tag{2.6}$$

$$\Delta\alpha = \frac{q^3\lambda}{4\pi^2c^3\epsilon_0 n}\left(\frac{\Delta N}{m_e^{*2}\mu_e} + \frac{\Delta P}{m_h^{*2}\mu_h}\right) \tag{2.7}$$

where q is the electronic charge, ϵ_0 is the free-space permittivity, n is the refractive index of Si, m_e^* is the effective mass of an electron in the conduction band (specified as $0.26m_0$), m_h^* is the effective mass of a hole in the valence band (specified as $0.39m_0$), μ_e is the electron mobility, and μ_h is the hole mobility. Experimental data was in good agreement with the Drude-Lorenz model for electrons, but not holes, and Soref and Bennet fit the empirical data with the function given in (2.4) and (2.5). When a refractive index change, Δn, is induced in the waveguide, the result is a phase shift in the propagating optical mode, $\Delta\phi$, manifest as a shift in the wavenumber, Δk, when operating at a fixed frequency. An increase

in carrier density results in a decrease in refractive index, whereas a decrease in carrier density results in an increase in refractive index [18]. In a p–n structure though, the index change is not discrete, but rather a gradient across the dimensions of the waveguide. As such, it is far more convenient to model the change in index as an effective index change of the entire waveguide, n_{eff}, rather than as a change of the index at each point in the waveguide. The phase shift resulting from a change to n_{eff} can then be modeled as

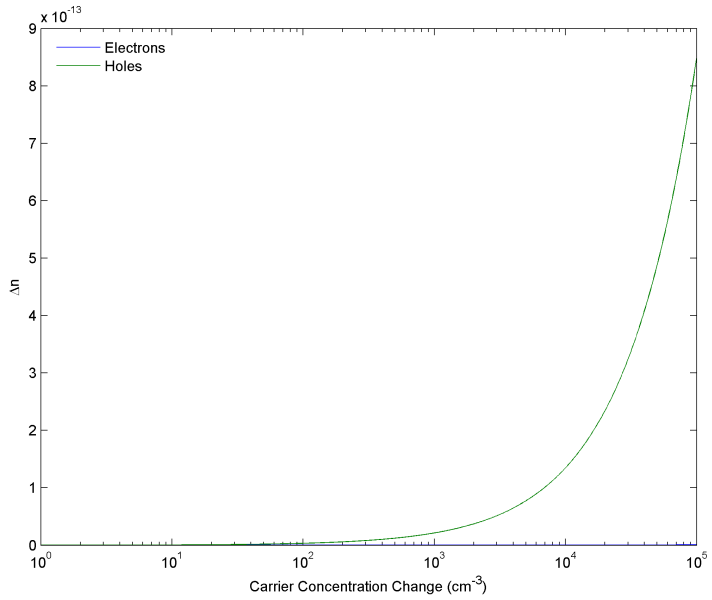
$$\Delta\phi = \Delta n_{\text{eff}} kL = \Delta n_{\text{eff}} \frac{2\pi}{\lambda} L = \Delta n_{\text{eff}} \frac{2\pi n_{\text{eff}}}{\lambda_0} L \tag{2.8}$$

where k is the wavenumber, λ_0 is the free-space wavelength, and L is the effective length of the modulator. The fact that the required length of the modulator can be on the order of a millimeter can be understood through the context of the relatively low change in refractive index achieved through the use of the plasma-dispersion effect [20]. As can be seen from both the above equation and from (2.6), the phase shift achieved is dependent upon the wavelength, and so the linewidth of the laser source affects performance. The applied phase shift should always be chosen such that it provides extinction at the primary spectral component of the laser; however, lasers with larger linewidths will achieve incomplete extinction at the remaining wavelengths, and so the maximum achievable extinction ratio is lower than it would be for a narrow-linewidth laser. In order to achieve an ideal output with zero energy in the output optical mode (the off state), the necessary phase shift between the two arms of the MZM is 180°, or equivalently, π-radians. In practice though, there are a variety of factors that preclude us from reaching an ideal off state.

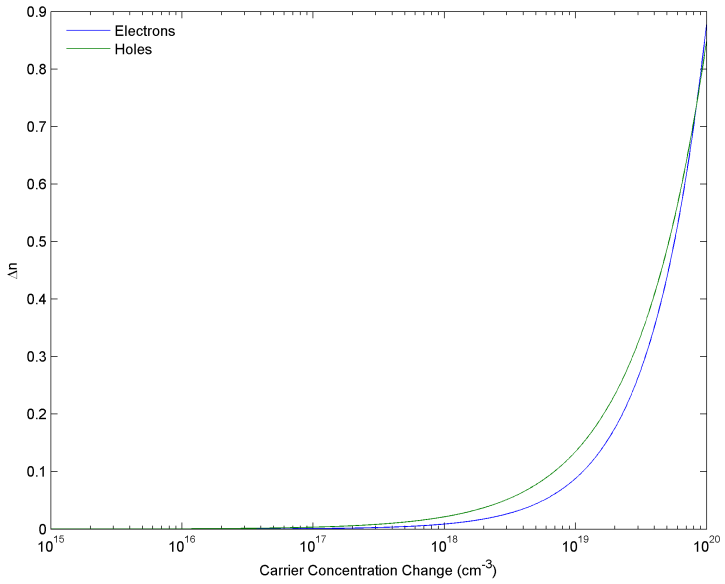
While it is possible to broadly classify modulators based on their mode of operation, it is also possible to classify MZMs in terms of their active regions: injection mode, depletion mode, and Metal-Oxide-Semiconductor-capacitor (MOS-capacitor). Injection and depletion mode devices are p–n diodes operating in forward and reverse bias, respectively. It is well known that when a p-type region and an n-type region are brought into contact with one another, a built-in voltage exists between the electronic bands of the

two doped regions. The area which the built-in voltage spans is known as the depletion region, named because the region is depleted of mobile charge carriers. When a voltage is applied across the p–n diode, the voltage either acts to decrease or increase the voltage drop across the junction, in turn decreasing or increasing the width of the depletion region, respectively. When the applied voltage acts to decrease the width of the depletion region, majority carriers are injected into the depletion region, and the device is said to be forward-biased. When the applied voltage acts to increase the width of the depletion region, majority carriers are extracted from the area surrounding the depletion region, and the device is said to be reverse-biased. It is for these reasons that forward-biased devices are said to act in injection mode, and reverse-biased devices are said to act in depletion mode. In an injection-mode MZM, the refractive index change is achieved primarily by means of injecting additional carriers, and in a depletion mode device, the opposite is true, and the refractive index change is achieved primarily by means of extracting carriers. Injection-mode devices can readily achieve the necessary phase change at a low voltage; however, the long free-carrier recombination lifetime in silicon-on-insulator (SOI) devices limits the bandwidths of injection mode devices [21]. In fact, the voltage across the junction is clamped to the knee voltage of the diode, and it is the current through the junction that becomes important. As the I-V curve in a forward-biased diode is exponential, small changes in V result in large changes in $V_\pi L$, in contrast to the relatively small effect of a change in L [22]. It is therefore possible to shorten the device significantly, and adjust V_π accordingly to compensate, until such time as the diode resistance begins to dominate power consumption [22]. Regrettably, the typical 3 dB bandwidth of such a device is less than 1 GHz. Signal pre-emphasis, or the increase in amplitude of the driving signal, can help to mitigate the effects of a long free-carrier lifetime and results in operation over a larger bandwidth; however, this is done at the expense of increased circuit complexity, and ultimately higher energy consumption per bit. Depletion mode devices, in contrast, have

successfully been operated at 10 GHz, and less commonly at speeds of 40 GHz, as well as one being operated at 50 GHz, however, a tightly confining structure is necessary to achieve a large refractive index change [23]. Without an optimized waveguide structure, depletion mode devices require substantially larger driving voltages than injection mode devices. The low reverse current present in depletion-mode devices typically means that there is low power dissipation [7]. Rather than using a p–n junction diode, the familiar three layer MOS-capacitor structure can be exploited, and the carrier concentration changes are in localized areas of the device — primarily the area immediately below the gate oxide — rather than injecting or depleting a junction. A MOS-capacitor is composed of a poly-silicon layer, an oxide layer, and a buried SOI layer. The poly-silicon layer functions as the gate terminal, and by the application of a voltage at the gate terminal, the silicon layer below the oxide will alternatively be inverted or depleted of free-carriers, thereby inducing the index change necessary to achieve a phase shift. MOS-capacitors are something of a compromise between injection- and depletion-mode devices, clearly not by way of operation or structure, but in terms of the switching speeds and power consumption. MOS-capacitors have been demonstrated to be capable of switching speeds in excess of 20 GHz, but with an energy consumption larger than a depletion mode device, yet smaller than an injection mode device. A MOS-capacitor achieves such figures on the basis of high electric field strength and large charge accumulation [21] [24]. Unfortunately, the structure of the MZM is such that the poly-silicon gate introduces excess optical losses in the form of free-carrier absorption, which although seemingly insignificant, can and will notably affect performance.



(a) Index change for reasonable values of carrier concentration change. The plot for change due to electrons lies nearly along the x axis.



(b) Index change for large values of carrier concentration change.

Figure 2.1: Effect on the real index as a function of carrier concentration change, broken out by carrier. As can be seen, for reasonable values of carrier concentration change, holes primarily contribute to the change in real index, and it is only at a value of 8.41×10^{19} that the electrons and holes produce an equal change in real index, after which electrons contribute the majority of change in real index.

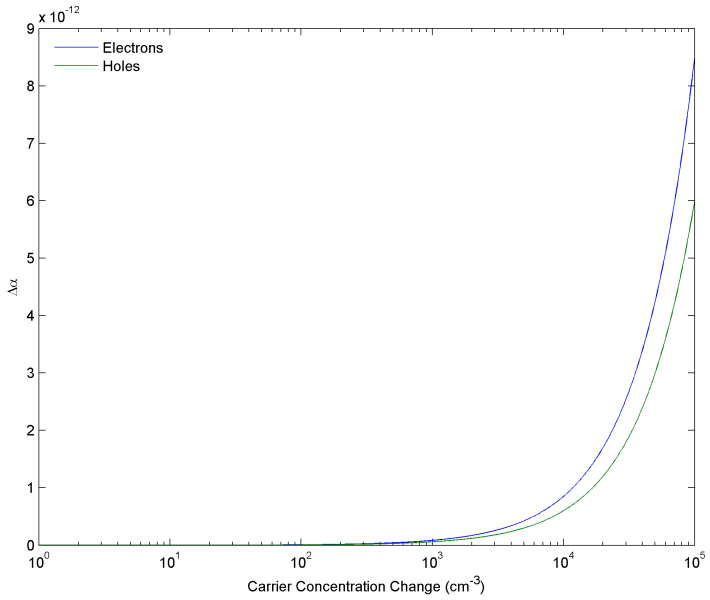


Figure 2.2: Effect on the imaginary index — absorption — as a function of carrier concentration change, broken out by carrier. As can be seen, electrons primarily contribute to the change in imaginary index.

There also exists some very basic knowledge with which a rudimentary design for the active region of a p–n structure can be devised. Returning to (2.4), the coefficient for index change governing the hole population is larger than that of the coefficient for electron population, and thus it is advantageous to maximize the overlap of the optical mode with the highest change in hole concentration. To this end, optimization of the waveguide parameters such as height and width stand to decrease $V_\pi L$. This intuition also translates into a desire to decrease the p-type dopant level relative to the n-type dopant level, in order to ensure that the depletion extends primarily into the p-type region of the p–n junction [23]. Moreover, an equal change in electron and hole concentrations produces a larger change in real index for holes than it does for electrons, and a smaller change in imaginary index for holes than it does for electrons, as can be seen in equations (2.4) and (2.5) [21]. Figure 2.1 also shows the contribution of each carrier type to the change in real index. Holes dominate

the process until a carrier concentration change of $8.41 \times 10^{19}\,\mathrm{cm}^{-3}$, which is a value that cannot be reached in practice. Looking at the depletion approximation, given by [25]

$$x_d = x_n + x_p = \sqrt{\frac{2\varepsilon_s}{q}\left(\frac{1}{N_a} + \frac{1}{N_d}\right)(\phi_i - V_a)} \tag{2.9}$$

where x_d is the width of the entire depletion region, x_p and x_n are the widths of the depletion region located in the p- and n-type sections, ε_s is the permittivity of silicon, N_a is the concentration of acceptors, N_d is the concentration of donors, ϕ_i is the built in voltage, and V_a is the applied voltage, it is possible to determine whether it is best to have a tall or wide waveguide. The ratio of width to height of the waveguide can be succinctly described by its aspect ratio (AR), defined as AR = width/height, where an aspect ratio of greater than one indicates a waveguide that is wider than it is tall, and one less than one, a waveguide that is taller than it is wide. Because the goal is to achieve the largest possible fractional change in effective index of refraction for the waveguide, and the width of the depletion region in the waveguide is inversely proportional to the square root of the carrier concentration, and it is directly proportional to the square root of the applied voltage, it is best to use a wide waveguide rather than a tall one, and thus a waveguide with $AR < 1$ is desirable [21]. It is therefore beneficial to form a vertical rather than horizontal p–n junction, which thereby reduces the voltage necessary to achieve a given change in carrier concentration. The location of the junction in the waveguide is also critical, as the modulation and phase change efficiency is maximized when the active region is centered within the waveguide, i.e. the location where the optical mode is largest in intensity [26].

2.1.2 Performance Metrics.

At this point it is best to introduce a number of performance metrics which will be considered throughout the course of the document.

$\mathbf{V}_\pi$ The voltage necessary to achieve a π-radian phase shift between the two arms of the MZM

L The length of the active region of the MZM

$V_\pi L$ The product of the two values above, and the primary figure of merit concerning MZMs, which can be used as a basis for comparison between alternatives

Insertion Loss (IL) The difference between the intensities of the input and output beam in the absence of any modulation, typically given in dB

$$\mathrm{IL} = 10 \cdot \log_{10}\left(\frac{P_{\mathrm{out}}}{P_{\mathrm{in}}}\right)$$

Extinction Ratio (ER) The difference between the intensities of the on and off states of the output, typically given in dB

$$\mathrm{ER} = 10 \cdot \log_{10}\left(\frac{P_{\mathrm{high}}}{P_{\mathrm{low}}}\right)$$

Bias Voltage The DC voltage operating point of the MOS-capacitor MZM

Driving Voltage The AC voltage applied to the MOS-capacitor MZM which has the effect of modulating the data onto the output waveform

Rise Time The time that it takes the output waveform to rise from 10% of the maximum intensity to 90% of the maximum intensity

Fall Time The time that it takes the output waveform to fall from 90% of the maximum intensity to 10% of the maximum intensity

Maximum Bit Rate The bit rate at which any further increase in bit rate would introduce an unacceptable level of bit errors at the receiver

Bit Error Rate (BER) A ratio of number of bit errors at the receiver to the number of bits transferred

Performance goals for an MZM are to decrease $V_\pi L$ while maintaining maximum bit rate and bit error rate. For the sake of comparison, injection-mode devices typically have the lowest $V_\pi L$, whereas depletion-mode devices have the highest, and MOS-capacitors occupy the middle ground. Ideally, a silicon MZM would be compatible with current on-chip CMOS drive voltages, and consume a roughly equivalent amount of power per bit as traditional MZMs, translating to a goal for V_{pp} of less than 1 V, and an energy consumption of between 2 and 3 pJ per bit [27]. A bit error rate of 10^{-3} is the typical upper limit for a receiver employing Forward Error Correction (FEC) techniques; otherwise, a bit error rate of less than 10^{-12} is generally targeted, as it is considered to be error-free [9].

By way of comparison, commercially available $LiNbO_3$ modulators have a 20-dB extinction ratio, with a driving voltage of 5 V_{pp}, and exhibit 4.5 dB of coupling loss, with a bandwidth of 60 GHz [28]. However, $LiNbO_3$ modulators are too expensive for backplane integration, and are too large for use in intrachip interconnects.

Yet another performance metric is the energy required per bit of data transmitted. The energy required to charge or discharge a capacitor, which is the circuit equivalent load that a modulation arm presents, is calculated as CV_{pp}^2. In the limit as time approaches infinity, in a random data signal where the probability of transitioning between a one and zero logic levels is 50 %, the number of transitions average to $B/4$, where B is the bit rate of the data signal [29]. The energy per bit is doubled to account for the fact that a series push-pull structure needs to drive double the capacitance of a modulator that only drives one arm. The formula to calculate the energy per bit is then computed as

$$P = \frac{CV_{pp}^2}{4} \tag{2.10}$$

for a single-arm modulation scheme, and

$$P = \frac{CV_{pp}^2}{2} \tag{2.11}$$

for a series push-pull modulation scheme. Note that this formula ignores any static power consumption that occurs as a result of DC bias. Assuming a forward voltage of 0.7 V for a

forward-biased diode, current can be calculated using the Shockley ideal diode equation:

$$I = I_S \left(e^{V_D/V_T} - 1 \right) \tag{2.12}$$

where I_S is the reverse saturation current (defined in 2.13), V_D is the diode's forward voltage, and V_T is the thermal voltage. In the case of a reverse-biased diode that has not broken down, the current is simply the reverse saturation current, given by

$$I = I_S = q\,A \left(\sqrt{\frac{D_p}{\tau_p}} \frac{n_i^2}{N_D} + \sqrt{\frac{D_n}{\tau_n}} \frac{n_i^2}{N_A} \right) \tag{2.13}$$

where q is the charge of an electron, D is the diffusion coefficient, τ is the carrier lifetime, n_i is the intrinsic carrier concentration, N_D and N_A are the donor and acceptor concentrations in the n- and p-type materials, respectively, and the subscripts p and n are meant to denote holes and electrons, respectively. Static power consumption can then be calculated using the well-known formula for electrical power, $P = IV$.

2.1.3 Loss Mechanisms.

The primary challenge with using silicon as a material in photonic circuits is the high propagation loss associated with the medium, which is atypical of waveguides in optical communications. An ideal waveguide, one which is perfectly straight and uniform, would have zero propagation loss [30]. Fabrication imperfections lead to a variety of propagation loss mechanisms, caused primarily by scattering off rough sidewalls, coupling to radiation modes, and to a negligible degree by intrinsic absorption in the bulk crystal (free-carrier absorption or FCA) [15]. The propagation losses experienced are dependent upon the configuration of the actual waveguide, common configurations being channel, ridge, photonic crystal, and slot waveguides, depicted in Figure 2.3. In channel and ridge waveguide configurations, light is confined to the core via index guiding and its associated total internal reflection. In a 2D photonic crystal, light is confined in one dimension by index-guiding, and confined in the other by way of the photonic crystals bandgap. In a slot waveguide, light is confined in a low-index channel that is embedded between two

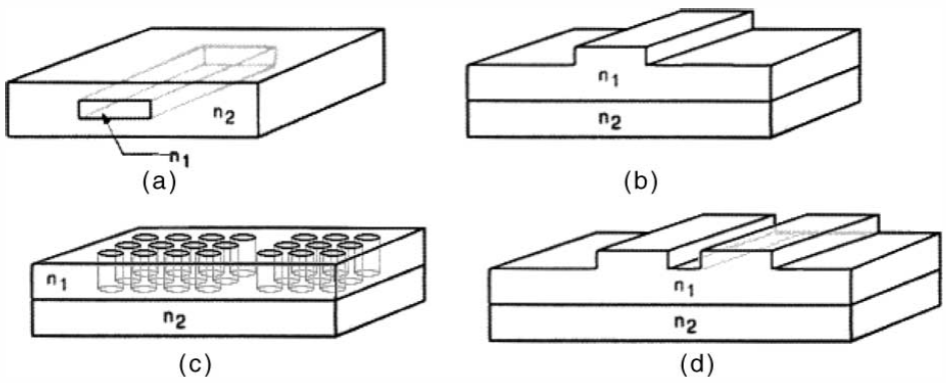


Figure 2.3: Different waveguide geometries: (a) channel waveguide, (b) ridge waveguide, (c) 2D photonic crystal waveguide, and (d) slot waveguide. [15]

silicon waveguides. Photonic crystal and slot waveguides display higher confinement factors than channel or ridge waveguides, and because slot waveguides propagate in a true eigenmode, they display no coupling to radiation modes, the downside being fabrication complexity [15]. Demonstrated losses for channel and rib waveguides have achieved figures of $80\,\mathrm{dB\,m^{-1}}$ and $20\,\mathrm{dB\,m^{-1}}$, respectively [31].

Coupling to a radiation mode in a waveguide occurs primarily when the geometry of the waveguide is altered significantly over a short distance, such as in a bend of the waveguide. This process can be understood in the context of the phase front of the optical mode. While the spatial mode of the field is not appreciably changed in the course of traversing a bend in a waveguide, the phase front must rotate around the center of curvature of the bend. In order to maintain a spatially coherent phase front, the components of the phase front farther away from the center of curvature, and in the cladding, must travel with a phase velocity that is higher than the phase velocity of the front closer to the center of curvature [15]. Given that the phase velocity can be no greater than the local speed of light, c/n, there will be some minimum radius of curvature, beyond which the phase front must

exceed the local speed of light in order to remain spatially coherent. As a result of the limiting local speed of light, the phase front breaks off and couples to a radiation mode.

Waveguide geometry also strongly influences the minimum allowable radius of curvature, as waveguides with a small cross-sectional area display a higher confinement factor, and as such localize power within the core. Thus, a single-mode waveguide will display a higher confinement factor than a multimode waveguide, thereby reducing coupling to radiation modes. Finally, the confinement factor is also directly proportional to the contrast between effective indices of refraction of the core and cladding, and so a channel waveguide, which has a high contrast, has a higher confinement factor than would a ridge waveguide, which typically has a low contrast [15].

The introduction of a bend in the waveguide, and the corresponding coupling to radiation modes, can also be understood as the introduction of a non-uniformity in the waveguide. So long as the length scale, L_n, is significantly larger than the characteristic length scale for coupling to radiation modes, z_c, then radiation losses will be minimal. Taking β to be the propagation constant of the fundamental mode in the waveguide, n_{cl} to be the index of the waveguide cladding, and k to be the wavenumber, then the characteristic length is the beat length between the fundamental and radiation modes, and is expressed as

$$z_c = \frac{2\pi}{\beta - kn_{cl}} \tag{2.14}$$

From [30], coupling to radiation modes via bend losses is minimal provided the following condition is satisfied

$$R_b \gtrsim \frac{W}{\rho}\frac{z_c^2}{2}, z_c = \frac{4\pi k n_{co}}{\beta^2 - kn_{cl}^2} = \frac{4\pi k \rho^2 n_{co}}{W^2} = \frac{4\pi\rho}{(2\Delta)^{1/2}}\frac{V}{W^2}, \beta \cong kn_{co} \tag{2.15}$$

where R_b is the radius of the bend, ρ is core radius, Δ is the difference in index between core and cladding, V and W are the model parameters for the fiber and cladding (defined in [32]), n_{co} is the index of the core, and n_{cl} is the index of the cladding.

Scattering losses in a silicon waveguide occur when the optical mode interacts with rough sidewalls, the roughness being the result of fabrication imperfections which are large

in size relative to the dimensions of the waveguide. From a macroscopic perspective, scattering theory tell us that the electromagnetic field will recognize non-uniformities if the length scale of the non-uniformities is significantly larger than the wavelength of the electromagnetic radiation, and so scattering becomes significant when $\lambda_0/n > L_n$, where L_n is the length scale of the non-uniformity in the waveguide, and n is the refractive index of the waveguide. The fact that the buried oxide and cladding are fabricated by means of deposition results in a uniform and specular surface, as opposed to the etching process used to define the horizontal dimensions of the waveguide, which results in a smoother, more Lambertian surface. Surface roughness can also be traced to mask-pattern imperfections during lithography. These small and random perturbations in the geometry of the waveguide influence the propagation losses, α, according to the equation

$$\alpha = \frac{\sigma^2 k_0^2 n_{\text{core}}}{n_{\text{eff}}} \frac{E_S^2}{\int E^2 \, \mathrm{d}x} \Delta n^2 \tag{2.16}$$

where k_0 is the free-space wavenumber, E_S is the field intensity at the core/cladding interface, E is the field intensity at a particular position along the cross-section of the waveguide, and n_{eff} is the effective index of the waveguide. If the sidewall roughness is periodic, then the scattering losses are also affected by the autocorrelation length. When the autocorrelation length approaches the wavelength of the electromagnetic radiation, the scattering points begin to resemble a grating, and scattering losses are further increased [15]. The fact that the confinement factor of a waveguide is wavelength dependent, and the fact that the periodicity of the roughness to wavelength ratio are both wavelength dependent, creates an overall wavelength dependence for the waveguide. In the former case, wavelengths larger than the one for which the waveguide was designed have lower confinement factors, and thus scattering losses are increased. In the latter case, as the wavelength decreases it approaches the autocorrelation length of the sidewall roughness, thus scattering losses are increased. The use of a tightly confined ridge waveguide minimizes losses by reducing the interaction between the fundamental propagating TE

and TM modes, thereby lessening the effects of sidewall roughness, and its associated scattering [31][15]. It is sensible to define the geometry of the waveguide such that the modal overlap with the sidewalls is minimized in comparison to the modal overlap with the buried oxide and cladding of the waveguide, i.e. utilize a waveguide with a low aspect ratio as opposed to a high aspect ratio, as high aspect ratio waveguides have higher scattering losses [15]. It is also possible to reduce sidewall roughness by means of thermal oxidization [33]. Optimal oxidation temperatures and times are necessary to ensure that the cross-sectional profile of the silicon wire waveguide is not deformed due to oxidation-induced stress. Experimentation shows that oxidation at 1100 °C allows the waveguide to retain its original cross-sectional profile, due to the stress release by the viscous flow of SiO_2 at such high temperatures [33]. It follows then that surface roughness rms scale decreases with oxidation time. Figure 2.4 shows the effects of oxidation at different temperatures and over different periods on the surface morphology of an Si/SiO_2 sidewall interface, and the associated effect on surface roughness. High-temperature hydrogen annealing has also been demonstrated to reduce sidewall roughness, but has yet to produce the lowest losses [3].

Considering both scattering and coupling to radiation modes as the primary mechanisms by which propagation loss occurs, both will be significant factors if the length scale of the non-uniformities lies within the scattering and coupling length scales, or mathematically, if $\lambda_0/n \lesssim L_n \lesssim z_c$. Otherwise, propagation losses due to scattering or radiation are minimal [30].

As silicon cannot be electrically pumped to act as a photon source, the design of an MZM must also account for insertion (coupling) loss, which in many scenarios is between 5 and 10 dB [5]. Insertion loss is determined not only by the contrast of the indices of refraction of the waveguide and the transmitting medium external to the waveguide, but also the geometry on the waveguide and the mode mismatch between the two waveguides.

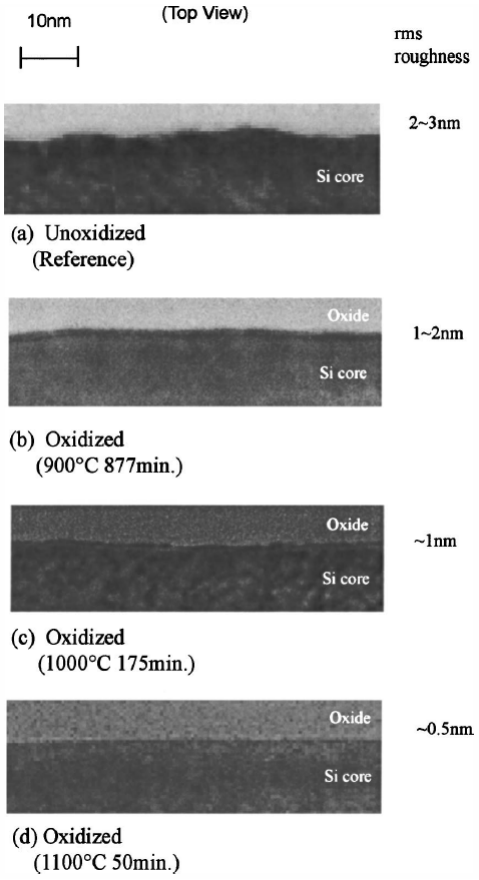


Figure 2.4: Effects of oxidation time and temperature on the sidewall morphology of an Si/SiO_2 interface. Temperature and oxidation time affect surface roughness, and an unoxidized reference sample is shown in (a), whereas (b), (c), and (d) were oxidized at different times and temperatures, as shown in the figure itself. [33]

Given that light is typically coupled to and from a silicon photonic chip by way of optical fiber, it would behoove us to match the core diameter of a fiber and the cross-sectional area of the silicon waveguide. Knowing that the core of a fiber is typically around 10 μm in diameter, using a large cross-sectional area silicon waveguide would make it easier to couple light into and out of the silicon chip. However, increasing the cross-sectional area decreases the confinement factor, which can have a negative impact on modulation efficiency via the overlap of the optical mode and the carrier concentration change, as well as require large bend radii in the waveguide. Beyond those difficulties, fabrication

complexity is also an issue. Therefore, silicon waveguides are typically engineered with smaller cross-sectional areas to increase the confinement factor, and thereby decrease minimum radius of curvature, and as such the modal diameter mismatch is high (A_{eff} of $0.1\,\mu\text{m}^2$ in a silicon wire waveguide, and A_{eff} of $50\,\mu\text{m}^2$ in SSMF-28) [3]. In order to ameliorate this, it is possible to use a lensed fiber to reduce the spot size coupling to the silicon waveguide, or one can use a variety of coupling techniques such as prism coupling, grating-assisted coupling, inverse taper-based coupling, or cross-sectional taper coupling, all of which serve to reduce coupling losses [5]. A cross-sectional taper aims to reduce the initial mode mismatch mentioned earlier, by increasing the cross-sectional area of the silicon waveguide at the entrance facet, and then gradually tapering down to the silicon waveguide (Figure 2.5). Inverse tapers rely upon a low-index material with a mode size roughly the same as that of the fiber, and then introduce a high-index waveguide inside of the low index material, adiabatically widening the high-index waveguide until such time as a majority of the mode is internally confined [34]. Cross-sectional tapers have achieved coupling losses of 0.5 dB, inverse tapers have been used to achieve a coupling loss of 0.2 dB, and surface gratings have achieved a 1 dB loss. Importantly, grating-assisted couplers can be used at any point on a photonic chip to couple light in and out, not simply at the facet of the chip. Regardless of the specific geometry of the silicon waveguide, the coupling efficiency of the two waveguides can be described by the equation

$$\eta = \frac{4\beta_1\beta_2}{(\beta+\beta_1)^2}\frac{\left(\int E_1^*(r,\phi)E_2(r,\phi)r\,dr d\phi\right)^2}{\int E_1(r,\phi)E_1^*(r,\phi)r\,dr d\phi \int E_2(r,\phi)E_2^*(r,\phi)r\,dr d\phi} \tag{2.17}$$

where β_1 and β_2 are the waveguide propagation constants, and the electric field amplitudes are given by E_1 and E_2, respectively. A fiber with a 10-μm core diameter edge coupled to a ridge waveguide of dimensions given in [15] displays a coupling efficiency of approximately 0.01.

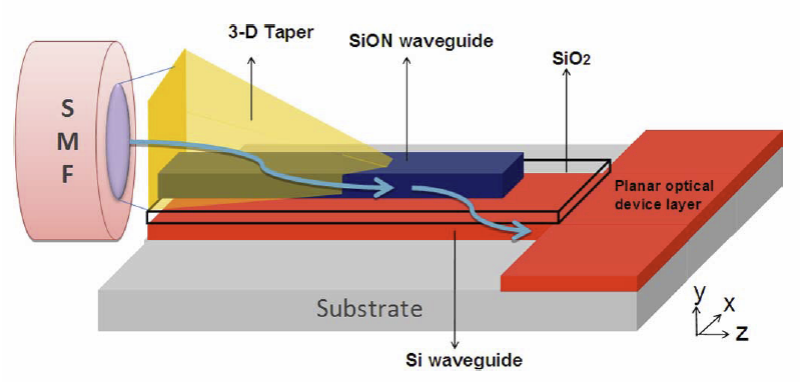


Figure 2.5: Schematic of a 3-D cross-sectional taper design to reduce insertion loss. [35]

In addition to simply considering the modal mismatch and overlap of the optical mode with the waveguide, one must also consider reflection due to the contrast in incident and transmitting medium indices of refraction. The reflectance and transmittance between two interfaces are defined by the well-known Fresnel Equations, and are shown below in (2.18 - 2.23). The reflectance and transmittance themselves are graphed in Figure 2.6. The parallel and perpendicular amplitude coefficients are equal when the incident wave is normal to the interface, because the plane of incidence becomes undefined, and so parallel and perpendicular become indistinguishable from one another. This is the manner in which the wave is launched into the silicon waveguide, and so the expected reflectance off the silicon interface is 31 %. It is for this reason that some research efforts have focused on index matching the coupling interface, or have chosen to use an index matching epoxy between the lensed fiber and the silicon waveguide. No such epoxy was used in this work, in order to maintain the ability to translate the input and output fibers between devices.

$$r_{\perp} = \frac{n_i \cos\theta_i - n_t \cos\theta_t}{n_i \cos\theta_i + n_t \cos\theta_t} \tag{2.18}$$

$$r_{\parallel} = \frac{n_t \cos\theta_i - n_i \cos\theta_t}{n_i \cos\theta_t + n_t \cos\theta_i} \tag{2.19}$$

$$R = r^2 \tag{2.20}$$

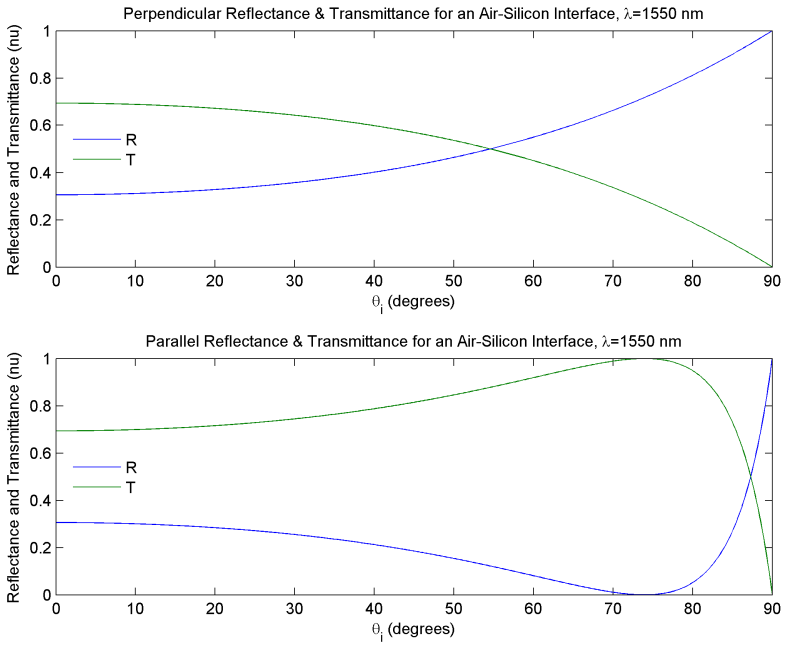


Figure 2.6: Reflectance and Transmittance between an Air-Silicon interface at $\lambda = 1550\,\text{nm}$. Note that at $\theta_i = 0$, the expected reflectance is 31 %.

$$t_{\perp} = \frac{2n_i \cos\theta_i}{n_i \cos\theta_i + n_t \cos\theta_t} \tag{2.21}$$

$$t_{\parallel} = \frac{2n_i \cos\theta_i}{n_i \cos\theta_t + n_t \cos\theta_i} \tag{2.22}$$

$$T = \left(\frac{n_t \cos\theta_t}{n_i \cos\theta_i}\right) t^2 \tag{2.23}$$

2.1.4 Power Division and Couplers.

The type of 3-dB coupler used to initially split, and then couple the light in the two arms of the MZM is also a factor. While it is common to use y junctions, multi-mode interference couplers, or directional couplers, such designs are highly intolerant to fabrication imperfections, and incredibly dispersive. In practice, this means that achieving the optimal 3 dB coupling ratio is quite a challenge when using a y-junction, directional

coupler, or multimode interference coupler. An adiabatic coupler, on the other hand, is minimally dispersive, and highly tolerant of fabrication imperfections — as long as the adiabatic coupler meets the minimum requirement for length, it shall adequately couple the light [36]. The tradeoff for tolerance to fabrication imperfections is the fact that adiabatic couplers are generally longer than the previously mentioned types of couplers. A coupler, or rather any process, can be termed adiabatic when the process evolves sufficiently slowly such that there are no transitions between energy levels, and therefore no changes in population levels, during the course of said process [37]. In the field of optics, an adiabatic process corresponds to one where energy remains confined to its original mode. A four-port adiabatic coupler is an asymmetric waveguide, which has two input and two output ports, with no modal coupling between the two ports. Physically, the coupler has three regions, the first of which has two asymmetric waveguides with a large gap between them, the second with a smaller gap where the geometry changes linearly to ensure symmetry at the end of the region, and the third region where the waveguides again diverge adiabatically. As the waveguide geometry becomes similar throughout region two, the power transfer takes place until such time as the two waveguides are entirely similar in shape and the energy is split evenly between the two modes. At this point, it is necessary to adiabatically separate the two waveguides to ensure that the energy remains in either the fundamental even or odd mode of each individual waveguide[37]. The coupler has a single system mode (the fundamental even and first-order odd modes) and waveguide mode. The waveguide mode evolves adiabatically to the system mode as it progresses longitudinally through the waveguide, the evolution being caused by a gradual change in the cross-sectional geometry of the waveguide [38]. Provided the geometry of the waveguides is sufficiently different, and the geometry of the waveguides evolves sufficiently slowly, the two waveguides can be brought together without an exchange of power. In order to effect a transfer of power between waveguides, the geometry of the waveguides must be made symmetric at its

junction. The wideband operation of an adiabatic coupler is caused by the fact that it exhibits no coupling to radiation modes, and because it has a wideband operation, it is also largely tolerant to changes in operating temperature and fabrication imperfections. The length over which it is necessary to evolve the geometry of the adiabatic coupler's waveguides is known as the coupling length, the optimal value for which is affected by the coupled local normal mode equations given by [39] as

$$\frac{\partial A_i}{\partial z} = -C_{ij}\frac{\partial \rho}{\partial z}A_j + j\beta_i A_i \tag{2.24}$$

$$\frac{\partial A_j}{\partial z} = C_{ij}\frac{\partial \rho}{\partial z}A_i + j\beta_j A_j \tag{2.25}$$

where A_i, A_j, β_i, β_j, and C_i, C_j are the amplitudes, propagation constants, and coupling coefficients of the fundamental and next-higher order system modes, respectively, and the derivative $\frac{\partial \rho}{\partial z}$ is the rate of change of a general parameter that is defined by the specific geometry of the waveguide (see [39]). The general solution to these equations can be computed for a given structure, yielding the optimal coupling length for a given error tolerance, ϵ, asynchronicity parameter between the waveguides at the start of the coupling region, X_{IIo}, coupling coefficient κ_{II}, and optimum length ratio between input and output coupling region lengths, α_{opt}, of

$$L_{3\text{dB}} = \frac{1}{\epsilon}\frac{X_{IIo}}{\kappa_{II}}(1+\alpha_{\text{opt}})^{3/2} \tag{2.26}$$

Simulations in [38] reveal that under a worst-case process variation, the splitting ratio is 45–55%, and under nominal conditions is 48–52%. The coupler does, however, display a polarization sensitivity, with the TM mode displaying a constant 3 dB coupling, and the TE mode varying between 0.25–0.60 and 0.45–0.65 power transmission as a function of gap size in region 1, and wavelength, respectively, with gap sizes ranging from 100 nm to 240 nm, and wavelength varying from 1530 nm to 1640 nm [36]. It is also worth noting that the use of a four-port adiabatic coupler in the MZM will ultimately lead to two complementary outputs, which is a fact that can be useful in analog applications, where

a differential signal can be used to compensate for external fluctuations, such as a change in the input laser power [40].

2.1.5 Other Factors Affecting Performance.

The waveguide itself also displays a polarization sensitivity, due to the built-in stresses of each epitaxial layer [41]. It is a well known fact that stresses and strains in a crystal can and will induce a birefringence. The stresses and strains in the waveguide manifest as a rotation of the principal axes of the crystal through the optomechanical effect. As a result of these effects, unless the optical mode is polarized so that it lies parallel to one of the principal axes of the crystal, the optical mode will see two distinctly different indices of refraction, and modal dispersion will occur over finite propagation distances. Additionally, even if the polarization of the optical mode were well-aligned with the principle axes of the crystal at the input, there is no guarantee that the stresses and strains in the crystal which govern the rotation of the principal axes will remain constant, and indeed it is unlikely that they will. As such, there will be minimal but random fluctuations in the index that the optical mode sees along the entire length of the waveguide, resulting in an additional and undesirable phase shift between the two arms of the MZM. The rotation matrix for the optical mode is nothing more than the standard rotation matrix in Euclidean space,

$$R(\rho) = \begin{bmatrix} \cos(\rho) & -\sin(\rho) \\ \sin(\rho) & \cos(\rho) \end{bmatrix} \tag{2.27}$$

where ρ is an empirical polarization axis rotation angle [41]. The optical mode is retarded according to the retardation matrix

$$D(P) = \begin{bmatrix} e^{ikn_0L} & 0 \\ 0 & e^{ikn_eL} \end{bmatrix} \tag{2.28}$$

where L is the length of the arm of the MZM, n_o and n_e are the indices of refraction for the ordinary and extraordinary rays, and k_0 is the free-space propagation constant, which is equal to $2\pi/\lambda$. The state of polarization can then be expressed as a Jones matrix, using the

product of the rotation and retardation matrices.

$$J = D(P)R(\rho) \tag{2.29}$$

Moreover, if the waveguide experiences a different level of confinement for TE and TM modes, then the mode with higher confinement would experience less loss and a higher degree of phase shift, which results in stronger modulation at a certain polarization. The plasma-dispersion effect itself is isotropic, and so any phase shift applied in the active region is applied to all polarization states equally [42].

Chirp is yet another characteristic that will play a role in determining the performance of the modulator. Broadly speaking, chirp is a mechanism which increases the spectrum of a signal, and as it applies to a modulator, is a variance in the time it takes to transition between on and off states. In a directly modulated scheme, chirp is a far larger factor than it is in an external modulator, due to the fact that turning the laser on and off introduces several frequencies into the signal. Even among other forms of external modulators, MZMs are considered to display low-chirp performance. Nonetheless, chirp when combined with the chromatic dispersion characteristics of a fiber, is often the limiting factor over long-distance communications links [43]. The chirp parameter α_{chirp} identifies the relation between phase and intensity modulation. In fiber-optic communications, the chromatic dispersion of the fiber is typically positive, and thus it is desirable to have a negative chirp parameter to achieve a zero-dispersion communications system [44]. Unfortunately, the chirp parameter of the common $LiNbO_3$ family of modulators is positive, which means that it is not able to counterbalance the dispersion found in fibers. The chirp parameter of an MZM in silicon though, is negative, and can be found by

$$\alpha_{\text{chirp}} = 2\frac{d\Phi/dt}{(1/I)dI/dt} \tag{2.30}$$

where I is the intensity of the wave, and Φ is the instantaneous output phase of the wave [43]. Chirp is introduced in a silicon MZM by way of (2.4) and (2.5), which details a

parasitic phase modulation accompanying an amplitude modulation, respectively. As chirp is a function of the derivative of output phase and intensity, the α_{chirp} is not a constant, but rather a parameter that varies with time, as long as the applied signal is also a function of time, as it often is. For this reason, the chirp parameter is often thought of in terms of its time average, $\alpha_{chirp-avg}$. The chirp parameter is then bounded by its maximum and minimum values, $\alpha_{chirp-max}$ and $\alpha_{chirp-min}$ respectively. The value $\alpha_{chirp-max}$ is reached when $\Delta\varphi$ reaches its maximum value (the phase shift in a single arm), and $\alpha_{chirp-min}$ is reached when I and $\Delta\Phi$ simultaneously reach their maximum and minimum values, respectively. In addition, as free-carrier absorption increases with signal amplitude, the absolute value of the chirp parameter will increase as the applied signal amplitude increases, as will the difference between $\alpha_{chirp-max}$ and $\alpha_{chirp-min}$ [43]. As per [43], the average chirp parameter does not change much with an increase in the clock or transmission frequency, although it has been demonstrated that the magnitude of $\alpha_{chirp-avg}$ is inversely proportional to the clock transmission frequency.

In an MZM where only one arm is modulated, the electrical field of the output optical mode can be expressed as

$$E \cdot e^{j\varphi} \cdot e^{j\omega_0 t} = \frac{1}{2} E_i \left(e^{-(\Delta\alpha+\alpha_0)L/2} \cdot e^{j(\omega_0 t+\Delta\varphi+\varphi_{DC})} + e^{-\alpha_0 L/2} \cdot e^{j(\omega_0 t+\varphi_{DC}+\varphi_{bias})} \right) \tag{2.31}$$

where E_i is the amplitude of the input electrical field of the optical mode, L is the length of the phase shifter, φ is the parasitic phase applied by the modulator, ω_0 is the angular frequency of the applied microwave signal, φ_{DC} and α_0 are the phase shift and absorption introduced by the applied DC bias, respectively, $\Delta\varphi$ and $\Delta\alpha$ are the phase shift and absorption introduced by the applied microwave signal, respectively, and φ_{bias} is any thermo-optic phase shift present [43]. From there, the instantaneous intensity and phase (in radians) can be derived as

$$I = \frac{1}{4} E_i^2 e^{-\alpha_0 L} \left[1 + e^{-\Delta\alpha L} + 2e^{-\Delta\alpha L/2} \cdot \cos\left(\Delta\varphi - \varphi_{bias}\right) \right] \tag{2.32}$$

$$\Phi = \tan^{-1} \frac{e^{-\Delta\alpha L/2} \cdot \sin(\Delta\varphi) + \sin\varphi_{bias}}{e^{-\Delta\alpha L/2} \cdot \cos(\Delta\varphi) + \cos\varphi_{bias}} \tag{2.33}$$

The chirp parameter, α_{chirp}, can then be computed from (2.30), and is given below, with the superscript $'$ denoting the time derivative [43].

$$\alpha_{chirp} = \frac{(L/2) \cdot (\Delta\alpha)' \cdot \sin(\Delta\varphi - \varphi_{bias}) - (\Delta\varphi)' \cdot e^{-\Delta\alpha L/2} - (\Delta\varphi)' \cdot \cos(\Delta\varphi - \varphi_{bias})}{(L/2) \cdot (\Delta\alpha)' \cdot \cos(\Delta\varphi - \varphi_{bias}) + (L/2) \cdot (\Delta\alpha)' \cdot e^{-\Delta\alpha L/2} + (\Delta\varphi)' \cdot \sin(\Delta\varphi - \varphi_{bias})} \tag{2.34}$$

Push-pull structures reduce chirp as they produce less parasitic phase shift during modulation, but the general trends regarding chirp and how it is affected by signal amplitude and frequency remain similar [43]. The equations themselves change to reflect the fact that there is now absorption and phase changes in both arms of the MZM. Ideally, the same phase shift (with opposite signs), $\Delta\phi_1$ and $\Delta\phi_2$, is achieved in both arms of the MZM, therefore, the push-pull configuration halves the amplitude and phase shift of an MZM that is only modulating one arm, therefore reducing $\alpha_{chirp-avg}$. Under the aforementioned ideal conditions, the carrier concentration changes in each arm would also be equal, however, carrier response times differ in injection and depletion mode – specifically, the carrier response time in depletion mode is shorter than in injection mode, which is the same reason that depletion-mode MZMs generally achieve faster speeds than injection-mode MZMs. This results in different carrier concentration changes being achieved in each arm, and so the phase shifts $\Delta\phi_1$ and $\Delta\phi_2$ are not in fact equal, and neither are the changes in absorption coefficients, $\Delta\alpha_1$ and $\Delta\alpha_2$; thus the chirp parameter is non-zero. As $\Delta\phi_1$ and $\Delta\phi_2$ have opposite signs, as do $\Delta\alpha_1$ and $\Delta\alpha_2$, the chirp parameter is reduced from that of an MZM in which only one arm is modulated. The pertinent equations for a push-pull modulation scheme are shown below.

$$E \cdot e^{j\varphi} \cdot e^{j\omega_0 t} = \frac{1}{2} E_i \left(e^{-(\Delta\alpha_1 + \Delta\alpha_2 + \alpha_0)L/2} \cdot e^{j(\omega_0 t + \Delta\varphi_1 + \Delta\varphi_2 + \varphi_{DC})} + e^{-\alpha_0 L/2} \cdot e^{j(\omega_0 t + \varphi_{DC} + \varphi_{bias})} \right) \tag{2.35}$$

$$I = \frac{1}{4} E_i^2 e^{-\alpha_0 L} \left[e^{-\Delta\alpha_1 L} + e^{-\Delta\alpha_2 L} + 2e^{-(\Delta\alpha_1 + \Delta\alpha_2)L/2} \cdot \cos(\Delta\varphi_1 - \Delta\varphi_2 - \varphi_{bias}) \right] \tag{2.36}$$

$$\Phi = \tan^{-1} \frac{e^{-\Delta\alpha_1 L/2} \cdot \sin(\Delta\varphi_1) + e^{-\Delta\alpha_2 L/2} \cdot \sin(\Delta\varphi_2 + \varphi_{bias})}{e^{-\Delta\alpha_1 L/2} \cdot \cos(\Delta\varphi_1) + e^{-\Delta\alpha_2 L/2} \cdot \cos(\Delta\varphi_2 + \varphi_{bias})} \tag{2.37}$$

$$\$\ S\ = e^{(\Delta\alpha_1-\Delta\alpha_2)L/2}$$

$$\alpha_{chirp} = \frac{(\Delta\varphi_1)' + S^2\,(\Delta\varphi_2)' + S\left[(\Delta\varphi_1)' + (\Delta\varphi_2)'\right]\cdot\cos\left(\Delta\varphi_1 - \Delta\varphi_2 - \varphi_{bias}\right) + S'\cdot\sin\left(-\Delta\varphi_1 + \Delta\varphi_2 + \varphi_{bias}\right)}{S\cdot S' - S\left[(\Delta\varphi_1)' - (\Delta\varphi_2)'\right]\cdot\sin\left(\Delta\varphi_1 - \Delta\varphi_2 - \varphi_{bias}\right) + S'\cdot\cos\left(\Delta\varphi_1 - \Delta\varphi_2 - \varphi_{bias}\right)} \tag{2.38}$$

It is possible to measure the small-signal chirp parameter by measuring the S_{21} response of the modulator with a dispersive fiber in line; the interplay between the fiber dispersion and the chirp of the modulator will result in resonance points in the response, which can then be modeled as

$$f_n^2\ L = \frac{c}{2D\lambda^2}\left(1 + 2u - \frac{2}{\pi}\arctan\ \alpha_{chirp}\right) \tag{2.39}$$

where f_n is the nth order of resonance, c is the speed of light, D is the fiber dispersion, λ is the wavelength of light in the medium, and α_{chirp} is the chirp parameter of the modulator [45].

The extinction ratio is contingent not only upon the phase change induced in the arms of the MZM, but also upon the intensity in each arm. Should the two arms recombine with a phase difference of π-radians between them, but be of unequal intensity, the interference will not be completely destructive. As can be seen from (2.5), there will certainly be a power imbalance between the two arms, which has the effect of decreasing the extinction ratio. This power imbalance creates a fundamental limit to the effectiveness of an MZM. One means of reducing the impact of free-carrier absorption on crosstalk is using a power splitter/combiner with a variable ratio (as opposed to 3 dB). However, as explained by [46], the on-chip losses increase in a manner such that any performance gains that might be achieved by recombining with matched powers are masked by the additional propagation losses, and therefore, performance in fact decreases when using something other than an ideal 3-dB power splitter/combiner. The overall effect of a power imbalance can be described by the following equations, where the amplitude of the two beams prior to

recombining after the modulation section can be expressed

$$E_1 = \sqrt{\frac{r_s}{2}} \tag{2.40}$$

$$E_1 = \sqrt{\frac{r_s}{2}} e^{-(\Delta\alpha \cdot L_\pi + j \cdot \Delta\varphi)} \tag{2.41}$$

where the power loss in each arm is described by r_s. The normalized output power after interference is then expressed as

$$P_{out} = \frac{r_s^2}{4}\left[1 + 2e^{-\Delta\alpha \cdot L_\pi}\cos(\Delta\varphi) + e^{-2\Delta\alpha \cdot L_\pi}\right] \tag{2.42}$$

and the output power in decibels can be expressed as $P_{dB} = 10\log_{10} P_{out}$ [19]. Moreover, there is more absorption at smaller wavelengths, leading to a decrease in ER at said wavelengths [47]. After propagating through a length, L, of the modulation section, the normalized amplitude of the optical mode can be calculated as given in [46] by

$$P_m = e^{(-\Delta\alpha \cdot L)}$$

$$\begin{aligned} L = L_\pi = \pm\lambda/2\Delta n \rightarrow P_m &= e^{(-\Delta\alpha \cdot L_\pi)} \\ &= e^{(-\Delta\alpha \cdot \lambda/2\Delta n)} \\ &= e^{(-\lambda/2)\cdot(\Delta\alpha/\Delta n)} \end{aligned}$$

The optical loss (in dB) can then be calculated as

$$\begin{aligned} A_m &= -10 * \log_{10} P_m \\ &= -10 * \log_{10} e^{(-\lambda/2)\cdot(\Delta\alpha/\Delta n)} \\ &\approx \frac{-10}{\ln 10}\frac{-\lambda}{2} \cdot \frac{\Delta\alpha}{\Delta n} \\ &\approx \frac{-10}{2\ln 10} \cdot -\frac{\lambda \cdot \Delta\alpha}{\Delta n} \\ &\approx \frac{2.171\ \lambda \cdot \Delta\alpha}{\Delta n} \end{aligned}$$

As can be seen, there exists a balance between the induced FCA and the interaction length: larger interaction lengths lead to greater optical loss (considering both FCA and propagation loss), but reduce drive voltages, whereas shorter interaction lengths decrease optical loss but increase the required drive voltages. Moreover, larger drive voltages may in and of themselves lead to more FCA, and so the interplay between the three parameters is not particularly straightforward. However, the change in carrier concentration impacts free-carrier absorption at a quasi-linear rate, but reduces the necessary interaction length at a higher-order rate [46]. The overall effect of the optical loss is, as previously mentioned, an imbalance of the optical power in each arm, leading to reduced extinction ratios.

The total optical loss can be expressed as a summation of the propagation losses and those due to FCA, where the A_t is the total loss, A_0 is the propagation loss, and A_m is the loss due to free-carrier absorption. A_0 has an expression similar in form to A_m

$$\begin{aligned} A_0 &= -10 * \log_{10} P_0 \\ &= -10 * \log_{10} e^{-\alpha_0 \cdot L} \\ &\approx \frac{-10}{\ln 10} \cdot \alpha_0 \cdot L \\ &\approx 4.343 \alpha_0 \cdot L \end{aligned}$$

$$L = L_\pi = \frac{\pm\lambda}{2\Delta n} \rightarrow A_0 \approx 4.343 \alpha_0 \cdot \frac{\lambda}{2\Delta n} \approx \frac{2.171\ \alpha_0 \cdot \lambda}{\Delta n}$$

$$\begin{aligned} A_t &= A_0 + A_m \\ &\approx \frac{2.171\ \alpha_0 \cdot \lambda}{\Delta n} + \frac{2.171\ \lambda \cdot \Delta\alpha}{\Delta n} \\ &\approx \frac{2.171\ \lambda\ (\alpha_0 + \Delta\alpha)}{\Delta n} \end{aligned}$$

There are two configurations for push-pull modulation schemes, series push-pull and parallel push-pull, each of which provides a different benefit. In both configurations, the two modulation arms undergo phase changes in anti-phase, and so the length requirement

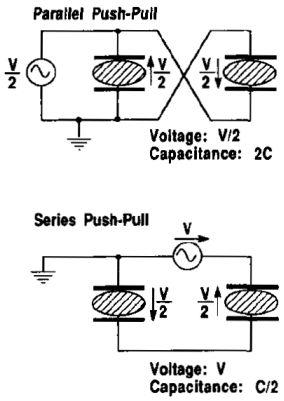


Figure 2.7: Difference between series and parallel push-pull modulation configurations. The series configuration doubles the device bandwidth by halving the capacitance, whereas the parallel configuration halves the necessary drive voltage. [48]

for the active region of the modulator is theoretically reduced to half its original length, by virtue of the fact that an interferometer measures phase differences. Figure 2.7 shows the difference between configurations. The series push-pull configuration places the two junctions in series – both junctions appear as a capacitive load – and thus the total capacitance of the modulation sections is halved. The reduction in capacitance by a factor of two, in turn, translates to a doubling of the device bandwidth. The drive voltage, however, is not necessarily halved, and may only experience a minor reduction. In a parallel push-pull configuration, the two junction capacitances appear in parallel, and thus the capacitance of the device is doubled with a corresponding reduction in device bandwidth by a factor of two. Advantageously, the parallel push-pull configuration requires only half of the original driving voltage to achieve the same phase shift. The tradeoff is bandwidth for efficiency of modulation. The push-pull configuration can also be designed in either single- or the more common dual-drive mode, where a single-drive mode requires a single microwave source, and dual-drive mode requires two differential microwave sources [9].

Because the length of the electrodes is not small compared to the wavelength of the microwave modulating signal, the maximum modulation frequency can be limited by the group velocity mismatch of the optical and electrical waves. The traveling-wave condition is necessitated by the fact that at high frequencies, the optical wave and modulating microwave must propagate in the same direction and at the same speed, to ensure that the modulation caused by one wave is not canceled by the antiphase modulation of a subsequent wave. Any walkoff in group velocities between the optical wave and microwave results in lower phase modulation depths at a given frequency, thereby reducing the device bandwidth [48]. Another factor affecting phase modulation depth is impedance matching. The result of an impedance mismatch is a reflection of the microwave and a resulting counterpropagation of the microwave. If the impedance of the entire device is mismatched, then power reflects off the transmission line during high-speed operation and never truly makes it to the device, leading to a reduction in dynamic ER [49]. If the end of the transmission is incorrectly terminated, then the microwave is reflected off the end of the transmission line, and results in a cancellation of the phase modulation by subsequent waves. In either case, impedance mismatches lead to poor efficiency.

2.1.6 Modeling and Device Performance Estimation.

Optimization of all relevant metrics such as bandwidth, power consumption, extinction ratio, bias voltage, and chirp, emphasize the importance of accurate predictive models, as modeling is less expensive than iterative fabrication runs.

During carrier injection, the rise/fall time is proportional to the effective carrier lifetime, which is defined as

$$\frac{1}{\tau_{\text{eff}}} = \frac{1}{\tau_{\text{B}}} + \frac{1}{\tau_{\text{s}}} \tag{2.43}$$

where τ_{B} is the recombination lifetime in the bulk semiconductor, and τ_{s} is the recombination lifetime at the surface. Because most carrier recombination takes place

away from the surface states in devices with large geometric dimensions, the value of τ_{eff} is dominated by the bulk recombination lifetime [20].

In the case of the injection-mode MZM, it has already been noted that the electrical bandwidth of the device is limited primarily by the long recombination lifetime of majority carriers. However, because carrier drift is far faster at varying the number of carriers within the active region of the MZM, depletion-mode devices have an electrical bandwidth imposed by the RC constant of the device electrodes. Parameters affecting the resistance and capacitance of the electrodes are doping levels, widths of the electrodes, and the spacing between them [7]. Increasing the semiconductor slab height also has the effect of decreasing access resistance for the metal electrodes, and as such increases the electrical bandwidth of the device; however, increasing the slab height might also drive the optical waveguide into multimode operation, which would negatively impact performance [50]. The electrodes should be separated from the doped semiconductor modulation section such that they are far enough removed to not suffer appreciable microwave loss due to interaction of the electrical mode with the doped semiconductor, yet close enough that the impact on access resistance is not significant.

The device can be converted to its small-signal equivalent using a lumped-element model, which is the capacitance of the reverse-biased diode in series with the access resistance to the junction from the metal electrodes due to the doped regions between the two [7]. The configuration is identical to an RC low-pass filter, which has an electrical bandwidth of

$$f_{3\,\text{dB}} = \frac{1}{2\pi RC} \tag{2.44}$$

where R is the access resistance, and C is the capacitance of the junction. The capacitance of the junction is directly proportional to the active region's surface area. In the absence of the ability to decrease the active region's length, the width of the active region should be reduced, which is inversely proportional to the cutoff frequency. The lower bounds

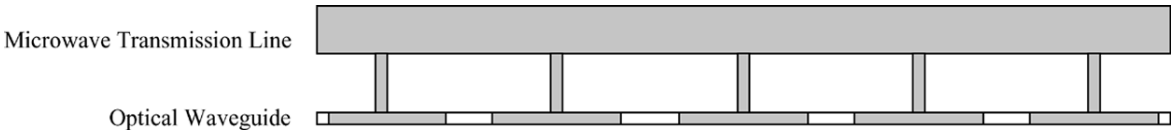


Figure 2.8: Segmented traveling-wave electrode design, showing the microwave transmission line along the top, and the optical waveguide and modulation sections on the bottom. [52]

on the active region width is the optical loss induced by an increase in interaction with the sidewalls of the waveguide. Additionally, the impedance presented by the junction capacitance should be significantly larger than the access resistance to ensure that the majority of voltage is applied to the junction capacitance to achieve a phase shift, rather than dropped across the access resistance [51]. Mathematically, this is the equivalent of satisfying the following equality

$$R \ll \left| \frac{1}{j\omega C} \right| \tag{2.45}$$

As one may recall though, simply increasing the capacitance to satisfy the equality will lower the device bandwidth, and so an increase in junction capacitance should be matched by a decrease in access resistance. This translates to an increase in dopant levels between the metal contacts and the junction, which in turn increases FCA — absorption of both the optical wave and the electrical microwave.

Traveling-wave electrodes can further increase device bandwidth by reducing velocity mismatch between the electrical and optical modes, and analysis shows that a segmented traveling-wave electrode offers better performance than either a lumped-element electrode or continuous-traveling-wave electrode [52]. In order to achieve optimal performance, traveling-wave electrodes require impedance matching, velocity matching, and low microwave losses. A segmented traveling-wave electrode has a separate transmission line running parallel to the silicon wire waveguide; see Figure 2.8 for a depiction of

the design. The microwave velocity of the electrode is faster than the group velocity of the optical mode, and the microwave impedance is higher than 50 Ω. Rather than modulating the entire length of the active optical region, distinct and periodic segments are connected to the transmission line, appearing as a capacitive load, thereby lowering the microwave velocity and impedance of the transmission line. Such a configuration is referred to as a capacitively-loaded transmission line, the loaded sections being considered capacitively loaded if the Bragg frequency of the periodic sections is much greater than the frequency range of the traveling-wave electrode itself [52] [53]. Advantageously, because a capacitively-loaded transmission line is located farther away from the doped semiconductor than the microwave electrode would be in a coplanar waveguide configuration, the electrical mode does not have a high overlap with the doped semiconductor, and microwave propagation losses are thereby reduced. The microwave impedance and velocity for loaded and unloaded transmission lines can be calculated using

$$Z_U = \sqrt{\frac{L_U}{C_U}} \tag{2.46}$$

$$n_U = c\sqrt{L_U C_U} \tag{2.47}$$

$$Z_L = \sqrt{\frac{L_U}{C_U + C_L}} \tag{2.48}$$

$$n_L = c\sqrt{L_U (C_U + C_L)} \tag{2.49}$$

where Z_U and Z_L represent the transmission line impedance in the unloaded and loaded case, respectively, n_U and n_L represent the transmission line velocity index in the unloaded and loaded case, respectively, L_U and C_U are the unloaded transmission line's inductance and capacitance per unit length, C_L is the loaded transmission line's capacitance per unit length, and c is the speed of light [52]. Translating the earlier requirements for impedance and velocity matching means Z_L must equal 50 Ω, and n_L must equal the index of the optical waveguide, in this case Si at 3.42. From the above equations, the following relations can

be derived:

$$L_U = \frac{Z_L n_L}{c} \tag{2.50}$$

$$Z_U n_U = Z_L n_L \tag{2.51}$$

$$\frac{n_L}{cZ_L} = C_U + C_L \tag{2.52}$$

$$C_L = \frac{n_L^2 - n_U^2}{cZ_L n_L} \tag{2.53}$$

According to Li *et al.*, these relations emphasize the following points: 1) to attain equal values for total capacitance, a large value of C_L is beneficial as it can reduce device length, accordingly lessening optical scattering losses and increasing device bandwidth, and 2) the values of Z_L and n_L are both fixed by the requirement that the optical and microwave velocity and index be equal, so minimizing the value of n_U can improve the value of C_L, which is the same as saying that the microwave velocity of the unloaded transmission line should be fast [52]. It is possible to determine the frequency response of the modulator using a microwave equivalent circuit model, as opposed to a lumped-element circuit model, which takes into consideration effects such as velocity mismatch, impedance mismatch, microwave loss, microwave dispersion, and filtering effects due to the periodic capacitative loading. The microwave equivalent circuit model considers N active modulation segments above the optical waveguide, shunting to the transmission line, with each active segment being modeled by a lumped $L_a R_a C_a$ circuit, the equivalent circuit model appearing in Figure 2.9. The inductance arises from the connection bridge, the resistance arises from the metal bridge resistance and the series resistance of the modulation sections, and the capacitance is the junction capacitance of the modulation waveguide [52]. Optical modulation is performed by varying the voltage across the junction capacitance, C_a. A detailed transmission matrix analysis is presented in [52], which results in an equation for the normalized frequency response of the modulator

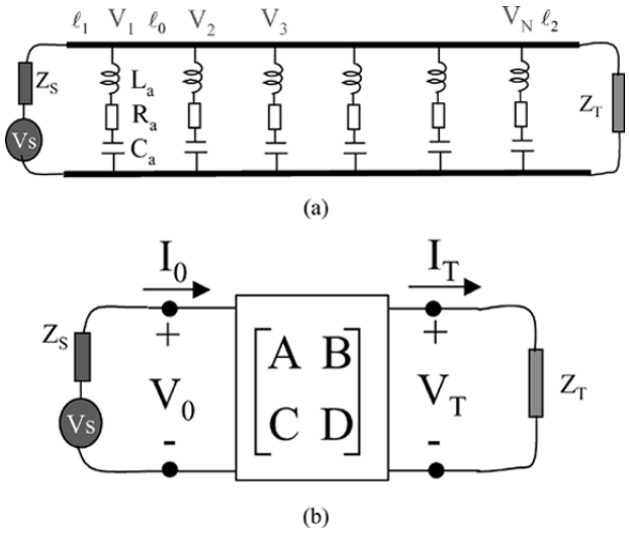


Figure 2.9: (a) A microwave equivalent circuit model of a segmented traveling-wave optical modulator. (b) The representation of the transmission matrix of a segmented traveling-wave optical modulator. [52]

$$M(f) = \left| \frac{2}{NV_s} \sum_{n=1}^{N} V_n e^{j\omega[\ell_1 + \ell_0(n-1)]/v_g} \frac{1}{1 - \omega^2 L_a C_a + j\omega R_a C_a} \right|^2 \tag{2.54}$$

where the input microwave voltage is V_s, N is the normalization constant for the summation, ℓ_0 is the microwave transmission line length between adjacent bridges, ℓ_1 is the microwave transmission line length between the applied voltage and the first bridge, v_g is the optical group velocity, and V_n is the microwave voltage across the nth loaded segment. The loaded transmission line is dispersive with regard to microwave impedance and phase velocity.

Varying the number of segments in the active region can lead to vastly different frequency responses for the device. In some cases, an increase in the number of segments can lead to a higher 3 dB bandwidth. In other cases, a reduction in the number of segments results in a hump in the frequency response, which then serves to increase the 3 dB bandwidth. Regardless of the effect on the 3 dB bandwidth, increasing the number of segments reduces the effect of dispersion and increases the cutoff frequency

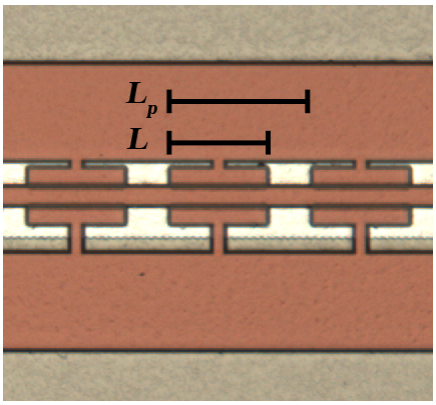


Figure 2.10: A segmented traveling-wave electrode, marked to denote L and L_p.

of the loaded transmission line, which has the effect of improving modulation responses at higher frequencies. As the loaded transmission line is dispersive, velocity and impedance matching can only be achieved at a select frequency, and therefore, it may be beneficial to exactly meet the impedance and velocity matching requirement at high frequencies rather than low frequencies [52].

It may also be beneficial to introduce the term fill factor, which is defined as $F = L/L_p$, where L is the length of each loaded section, and L_p is the length of the period itself. Both L and L_p are denoted in Figure 2.10. The effective length of the modulation section is then described as $L_{\text{eff}} = L\,N$, where N is the number of loaded sections [45].

2.2 Device Specifics

A variety of MZMs were designed and fabricated by Zortman *et al.* at Sandia National Laboratory and tested during the course of this thesis work. One family of devices utilized a push-pull (PP) modulation scheme and the electrodes were designed using a lumped-element model (see Figure 2.11), and another family of devices was designed using a traveling-wave electrode (TWE) but only used a single-arm modulation scheme (see Figure 2.12). Both families of devices were produced with active regions of different lengths;

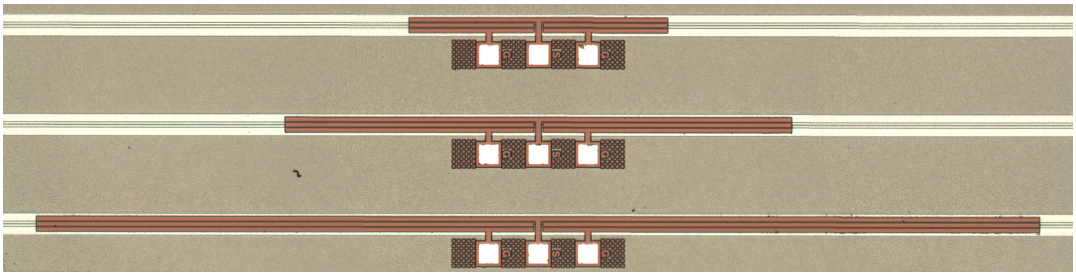


Figure 2.11: Push-pull Mach-Zehnder modulators, with active region lengths of 500-µm, 1000-µm, and 2000-µm from top to bottom.

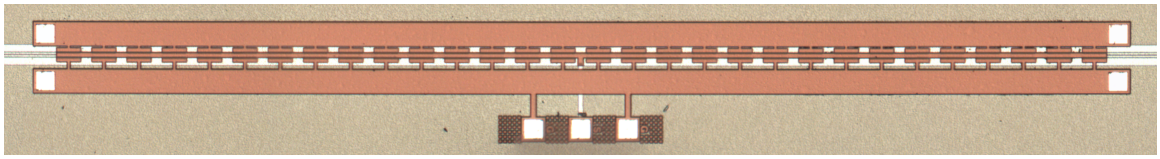


Figure 2.12: Traveling-wave electrode Mach–Zehnder modulator with an effective active region length of 1500-µm.

however, testing on the TWE family was only performed on the 1500-µm length device due to resource limitations.

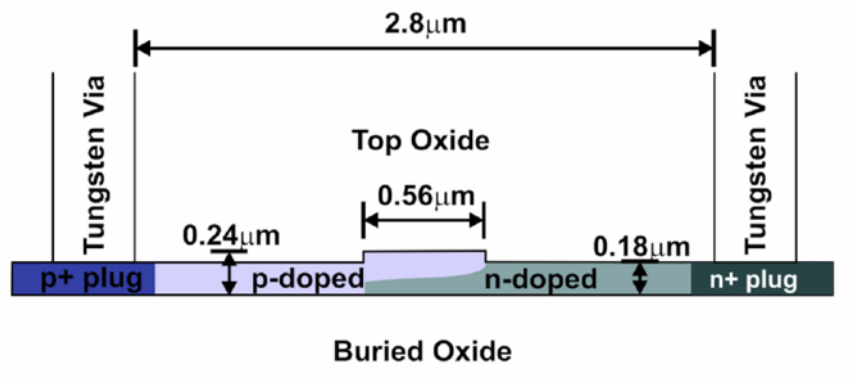
Both devices are vertical p–n junctions and intended to be operated in depletion mode, utilizing a 2x2 adiabatic coupler with a ridge waveguide. The use of a 2x2 port adiabatic coupler provides complementary outputs.

Both the PP and TWE MZMs use the same active region design and doping profile, as shown in Figure 2.13, and cross-sectional geometry shown in Figure 2.14. The two MZMs only differ in electrode design, and the manner in which the DC bias is applied to the device. The specific manner in which the DC bias is applied to the MZMs will be covered later, however what is important is the fact that the manner in which the bias is applied causes the PP MZM to be operated using a single-drive serial push-pull modulation scheme, and

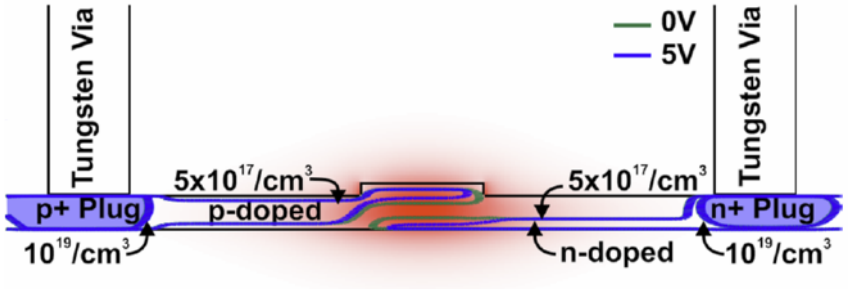
the TWE MZM to be operated using a single-drive single-arm modulation scheme. The waveguide consists of a 560-nm by 240-nm ridge which sits atop a 3-μm buried oxide slab. The silicon layer on top of the buried oxide is implanted with p- and n-type dopants to form the vertical junction only in the area underneath the ridge, elsewhere being solely p- or n-type. Bounding the p- and n-type regions are p+ and n+ plugs, contacted with tungsten vias, to enable modulation. The p- and n-type regions are doped with phosphorous and arsenic to levels of $5 \times 10^{17}\,\mathrm{cm}^{-3}$, and the p+ and n+ plugs are doped to levels of $10^{19}\,\mathrm{cm}^{-3}$. The two tungsten vias are gapped from the center of the ridge by a distance of 1.4 μm to minimize the device resistance and maintain low microwave loss operation, with a junction capacitance of 0.41 fF at 0 V bias.

From a macroscopic point of view, the push-pull design uses two electrodes, both sitting close to the ridge waveguide they are designed to modulate. The DC and microwave signals are applied using the same GSG pads below the device, with the aid of a bias tee. The bottom electrode is held at ground, and the top electrode is the signal electrode – the design of the junctions means that the push-pull modulation scheme is implemented as a single-drive push-pull modulator. The structure, as defined, then has an RC time constant of 5.6 ps, with a resulting intrinsic RC bandwidth of 29 GHz, though the transmission line used to drive the modulator limits the expected 3-dB bandwidth to 6.3 GHz [21].

In the case of the TWE design, there are three electrodes, two metal strips which are clearly visible bounding the active region, and a central electrode which runs the length of the active region. The top electrode is the signal electrode, the bottom is the ground electrode, and a DC bias is applied to the central electrode. The DC bias is applied by means of the GSG pads below the device, whereas the microwave signal is applied using the contact pads on the left or the right of the MZM, the other set of pads being used to terminate the signal with a 50 Ω termination. As can be seen from Figure 2.15, devices were also fabricated with a built-in termination, where the microwave signal would be applied



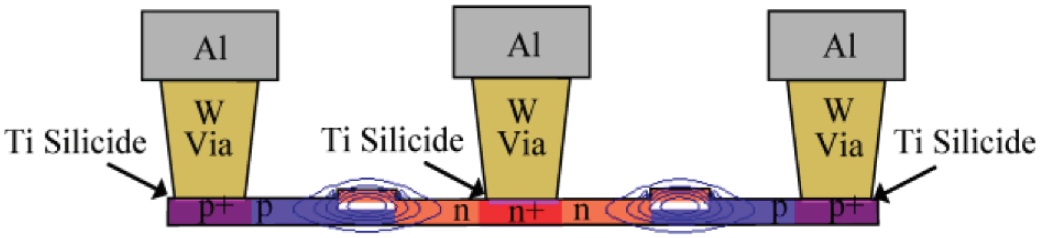
(a) Geometry of a single ridge waveguide with no applied signal.



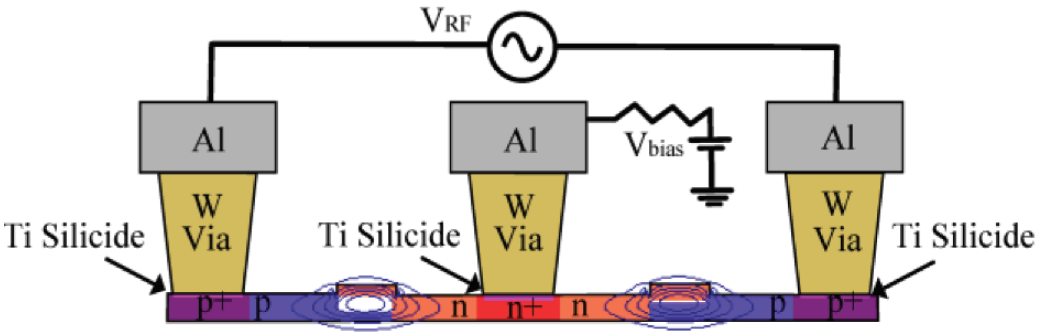
(b) With an applied bias noted in the figure, the carrier concentrations change with the contour lines shown. The optical mode is shown in red.

Figure 2.13: The cross-sectional geometry and doping profile of the PP and TWE MZMs with no applied signal (a), and an applied signal (b). [21]

on the left set of pads using a GS or SG probe. The unloaded slot line has an impedance of 95 Ω, which was then impedance matched using 50-μm segments with a fill factor of $F = 0.6$. The measured 3 dB bandwidth of the 1500-μm TWE device was 14 GHz, limited by the series resistance in the p–n junctions [54].



(a) No applied signal, and so the depletion regions are identical in width.



(b) With an applied bias and a differential microwave signal, the size of both depletion regions change, and the device is operated as a push-pull MZM. When using a single-ended microwave signal, only one arm is modulated, and so the depletion region of only one arm is changed. This image depicts the application of a differential microwave signal to the transmission line.

Figure 2.14: The cross-sectional geometry and the PP and TWE MZMs with no applied signal (a), and an applied signal (b). Note that applied signal configuration is only relevant for the TWE MZM, as the PP MZM leaves the central electrode floating. The depletion region is represented by the white coloring in the ridge waveguide, and a contour plot is shown for the E_x field of the optical mode. [54]

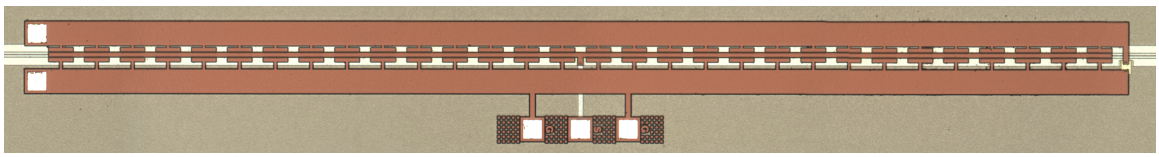


Figure 2.15: Traveling-wave electrode Mach–Zehnder modulator with an effective active region length of 1500-μm, and a built-in termination on the right.

III. Experimental Methodology

3.1 Experimental Design

A variety of experiment types were performed: DC Characterization, Bit Error Rate (BER) Characterization, and Bandwidth Characterization. While each experiment type sweeps a different set of factors, all rely upon the same general experimental design. All factors thought to be of relevance were controlled, however, factors such as ambient lighting, temperature, and humidity, were not controlled and would appear as nuisance factors in the results. The DC characterization was used as a screening experiment, as it typically took the least amount of time to run, while the BER and bandwidth experiments were designed with consideration, given the results of the DC characterization.

Most experiments were run using a full factorial design, where the experiment was specified in terms of the factor to be swept, and conditions were held constant. With the exception of optical polarization, each factor to be swept was specified in terms of a discrete starting point, a discrete ending point, and the step interval between points. The full experiment was then compiled as the Cartesian product of all factors and conditions. By way of example, in a DC characterization experiment, the parameters to vary are temperature of the silicon chip, bias voltage, wavelength, and the input polarization of the optical mode. In a single experiment, any of the above factors can be swept or held constant. Unfortunately, due to equipment limitations, input polarization was unable to be specified or controlled quantitatively. If choosing to sweep only wavelength and bias voltage, with the wavelength being swept from 1540 nm to 1550 nm in 5-nm steps, and bias voltage being swept from -2 V to -1 V in 1-volt steps, with temperature held constant at 20 °C, then the full experiment would consist of the conditions given in Table 3.1.

Experiments used a full factorial design as opposed to a one factor at a time (OFAT) design, because an OFAT design fails to capture the effects of interaction between

Table 3.1: Example experimental conditions using a normal ordering.

Temperature (°C)	Wavelength (nm)	Bias Voltage (V)
20	1540	-2
20	1540	-1
20	1545	-2
20	1545	-1
20	1550	-2
20	1550	-1

factors. For example, in an OFAT design, the previous experiment (specified by the experimental design in Table 3.1) would not have been able to be run because both wavelength and bias voltage are varied at the same time. In an OFAT design, one would have to choose to vary solely temperature, wavelength, or bias voltage. A full factorial design was chosen over a fractional factorial design because the length of time to run any particular experiment was typically not an issue. Additionally, multiple replicates of the same experiment were typically performed to minimize the effects of nuisance factors and random noise. Statistical methods require that experimental error (noise) be randomly distributed throughout the sample set, which equates to randomizing the order of experimental conditions; however, doing so proved to require a significantly larger amount of time than a normal experiment [55]. This is due to the fact that a certain amount of time is required to change each factor. For example, while it may take only 100 ms to change to any given bias voltage, it may take 3 s to change to any given wavelength. If the order of experimental trials is set such that the number of changes in wavelength is minimized, then the amount of time to complete a certain experiment is also minimized. If, however, the experiments are randomly distributed, it is likely that the number of wavelength changes is significant, as compared to a normally ordered experiment. By way of example, the

same experiment given in Table 3.1 is randomized as it is in Table 3.2, then the number of wavelength changes increases from two to four, and the number of wavelength changes dominates the time it takes to complete the experiment.

Table 3.2: Example experimental conditions using a randomized ordering.

Temperature (°C)	Wavelength (nm)	Bias Voltage (V)
20	1545	-1
20	1540	-2
20	1545	-2
20	1540	-1
20	1550	-1
20	1550	-2

Rather than perform experiments manually, all experimentation was automated through remote control interfaces exposed on lab instruments via GPIB or ethernet. Automation facilitates the collection of a vast amount of data at a rate faster than would have been possible if a human were to collect the data manually. This provided the capability of running experiments using very small step sizes, which took days, or in some cases weeks, to complete, while eliminating the need for human intervention. The program, called AutoLab, was written in C#, using the .NET 4.5 framework and Windows Presentation Foundation, exclusively for the purpose of this thesis work.

The use of automation allowed for the control of fixed values such as the input optical power level. To ensure that valid comparisons could be made between data sets, it is vital that the same amount of input power be supplied across trials and experiments. Because the gain spectrum of the laser and amplifier used is not flat, the current to the laser diode would have to be tuned as the wavelength is swept. AutoLab was configured to use a

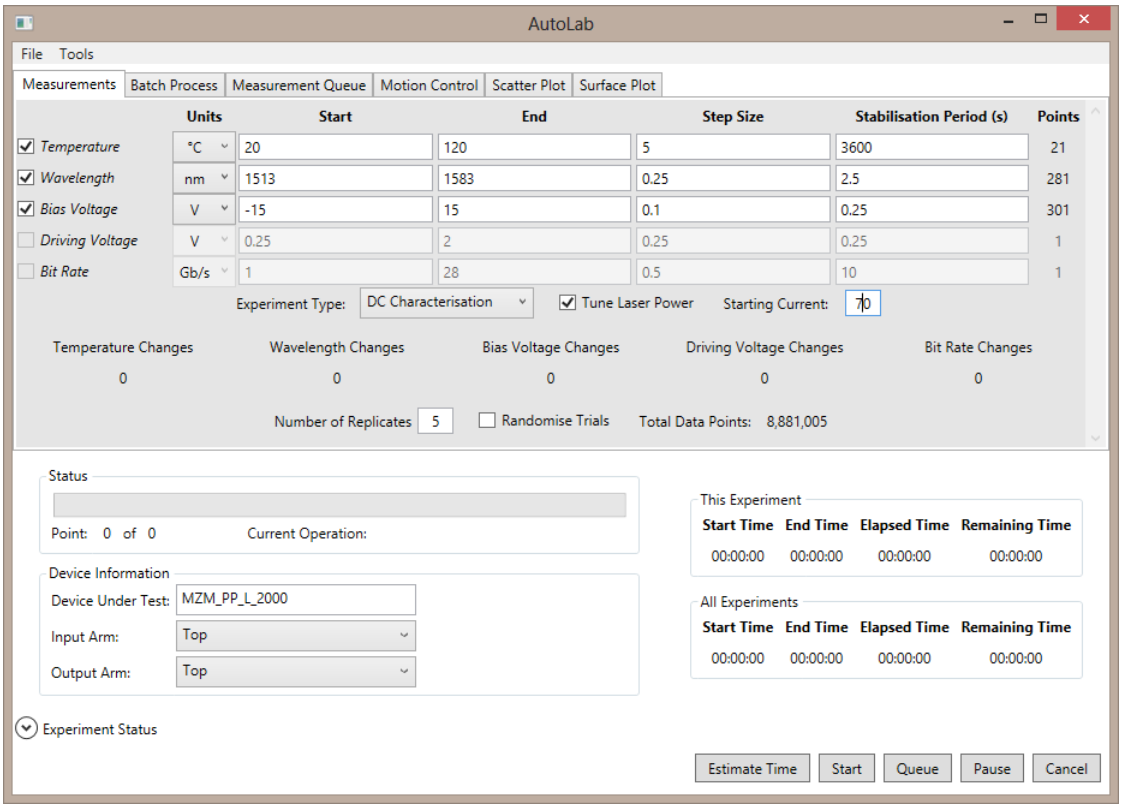


Figure 3.1: A snapshot from AutoLab, the automation program used during experimentation.

feedback loop to maintain a constant input power. It is important to note that although the laser source itself could tune from 1512.3 nm to 1583.4 nm, experiments were typically limited to the range 1540 nm to 1583 nm, as the dip in the gain spectrum from 1512.3 nm to 1540 nm could not be effectively compensated. A snapshot of the program interface is provided in Figure 3.1.

3.1.1 DC Measurements.

A wealth of characterization was performed on the various device designs under DC bias conditions. DC measurements enable the characterization of device performance across a variety of bias voltages, wavelengths, temperatures, and input optical polarization. Altering the bias voltage allows us to view the transmission as a function of bias voltage,

and altering the wavelength and temperature of the device allows us to see if either has an effect on the performance of the device. Altering the bias voltage obviously introduces a change in carrier concentration, and so the voltage across the junction was always expected to alter transmission.

Varying the input wavelength was not expected to be a large factor as the plasma-dispersion effect is wavelength dependent, but should vary minimally over the wavelengths of interest, and adiabatic couplers are designed to operate over a broad wavelength range. Temperature was expected to play a role in the performance of the device, as silicon displays a large thermo-optic coefficient; however, temperature should only affect device performance by shifting the transmission at each bias voltage, so that the application of an identical driving voltage at multiple temperatures would achieve the same device performance, as long as the bias voltage was adjusted to compensate for the change in temperature. Unless stated otherwise, all experiments were conducted at an actively controlled temperature of 20 °C, or equivalently, 68°F – the de facto standard for room temperature.

As explained earlier, some change in performance was expected based on a change in the input polarization state, as the wavelength displays a different confinement for different polarizations, the waveguide is birefringent, and the adiabatic coupler has a variable coupling ratio for the TM mode when sweeping wavelength, but holds constant for the TE mode.

3.1.2 Bit Error Rate Characterization.

A bit error rate test is performed by comparing a known data signal to a received modulated waveform – comparing each bit in sent and received waveforms to ensure accuracy. In comparison with eye diagrams, BER tests provide a high-bandwidth measurement of a device's ability to transmit a signal. While an eye-diagram is a good tool to use to explore a modulator's performance, an open eye diagram may give the impression

of a modulator that is operating error-free, and under the same test conditions, a BER test might say otherwise.

In a BER test, the microwave signal is sent by a pattern generator to the modulator, which is typically operated at a DC bias point providing an ideal extinction ratio, and the applied microwave signal provides a driving voltage that in turn modulates the electrical signal onto the optical waveform. That optical waveform is then reconverted into the electrical domain by a high-speed photodetector, and fed into an error detector. The error detector also receives the pattern directly from the pattern generator (reference pattern), time delays the reference pattern appropriately, and compares each bit in the sequence between the reference and signal patterns for accuracy. Any flipped bits are reported as errors, and the results are typically reported as a ratio of the number of errors detected to total bits sent, otherwise known as the bit error rate.

During a BER test, the temperature, wavelength, bias point, driving voltage, bit rate, and test pattern can all be tested as factors. Temperature and wavelength were not anticipated to have an effect on performance, and are therefore held constant during BER tests. The remaining four parameters — bias point, driving voltage, bit rate, and test pattern — were predicted to have a large effect on the device's BER. In practice, BER tests take a considerable amount of time to run, and the driving voltage was limited by the equipment to <2-V_{pp}, which was in all cases less than the driving voltage that maximized the static ER. As such, the driving voltage was fixed at 2-V_{pp}, and bias point, bit rate, and test pattern were varied. Testing was primarily conducted using Pseudorandom Bit Sequences (PRBS), which are specified in terms of the number n, where a single PRBS test pattern contains every possible combination of n bits, with the exception of a word consisting entirely of zeros, and therefore contains $2^n - 1$ different words. Convention names PRBS patterns as PRBS-n, so a PRBS pattern with 7 bit words is named a PRBS-7 pattern. Table 3.3 shows the different values for n, and the corresponding word length and number of unique

words in the PRBS pattern. Unless stated otherwise, a PRBS-7 pattern was used during experimentation. In the cases when another pattern was used, a PRBS-7 pattern was first used to explore the device performance versus bias voltage. Once an ideal bias voltage was determined, all possible PRBS test patterns were used while sweeping the bit rate. The use of a PRBS pattern stresses the modulator in ways that a 10 pattern simply does not. Whereas the power spectrum of a 10 pattern appears at a limited number of frequencies, the use of a PRBS pattern spans a broader range of frequencies, and accordingly tests the modulator's ability to transmit a signal with a large spectral content. The spacing between components of the power spectrum decreases as pattern length increases, and the envelope of the power spectrum changes with the frequency at which data is transmitted. Unless the S_{21} response is flat with respect to frequency, then the use of a PRBS pattern is necessary for testing the modulator as it would be used in real-world conditions.

Table 3.3: PRBS pattern specifications.

n	Name	Word Length (bits)	Unique Words
7	PRBS-7	7	127
11	PRBS-11	11	2,041
15	PRBS-15	15	32,767
20	PRBS-20	20	1,048,575
23	PRBS-23	23	8,388,807
31	PRBS-31	31	2,147,483,647

Another method by which device performance is benchmarked at microwave speeds is the eye diagram. Eye diagrams are a tool typically used in the communications field to intuitively inspect the ability of a modulator to transmit a signal. Eye diagrams repetitively sample the optical power level over time, and therefore trace the optical signal over the

course of several bit patterns. Eye diagrams can easily show the effects of jitter, overshoot, undershoot, and intersymbol interference. While eye diagrams record the same signal as a BER test, eye diagrams will only show events that occur with a higher probability, while a BER test will show any events that occur. For example, in a test consisting of 10^{12} bit with only 1 error, the BER test would show a BER of 1×10^{-12}, but the eye diagram would have a large opening and look to be error-free. This is because an eye diagram has an acquisition or sample rate well below the bit rate, while the BERT samples every transmitted bit. Because of this, eye diagrams typically do not show events that occur with a probability of less than 10×10^{-5}, despite the fact that such errors hamper transmission performance. Therefore, eye diagrams alone cannot fully characterize device performance, although they intuitively display the transmitted signal.

Another diagram, called the BER Contour, draws the eye diagram with BER contour lines inside the eye opening. If latching in the data value with a threshold voltage and delay time inside a given BER contour, the MZM will operate with a BER of less than or equal to that of the BER contour. Although a number of BER contours were not collected, one will be presented in order to illustrate the limitations of an eye diagram. That BER contour is given in Figure 3.2, and although the eye appears to be open, and as such one would expect error-free operation, there is only a small region where the BER is $<10^{-12}$. The algorithm used to optimize threshold voltage and delay time in fact chose to operate in a region where the BER was 7.74×10^{-7}.

3.1.3 Bandwidth Characterization.

The device's electrical bandwidth can be tested by measuring its scattering parameters, also known as the S-Parameters. S-Parameters constitute a method of vector network analysis, which measures a device's effect on the amplitude and phase of signals over a range of frequencies, which can broadly be thought of as distorting the original signal. Both linear and non-linear devices can distort signals, with linear distortion altering the

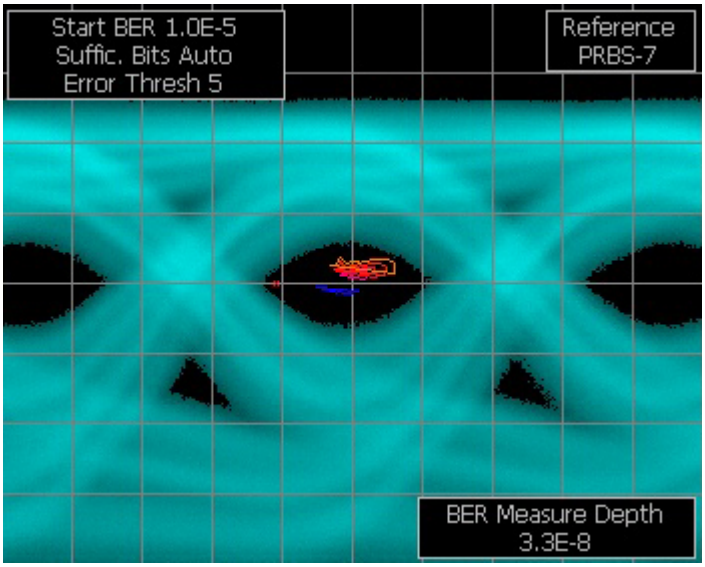


Figure 3.2: BER Contour for the 1500-μm TWE MZM, taken at 18 Gbps and 1550 nm. The diagram is at 33 mV per division. Looking at the eye alone, one would expect the MZM to be operating error-free because the eye is very much open. However, as can be seen from the contour lines, there is only a small region where the MZM is operating with a BER $<10^{-12}$ – and that would be inside the orange contour lines.

amplitude and phase of the original frequency content, and nonlinear distortion adding additional frequencies to the signal content.

Linear devices are said to impart signal distortion when the amplitude response is not flat, and the phase response is nonlinear. It is important to note that under normal circumstances a linear device does not alter the frequencies present in a signal, the exception being when linear devices are driven with large signals which may then cause saturation or compression, distinctly non-linear behavior.

Fundamentally, a vector network analysis is nothing more than a measurement of the energy sent by the source, reflected towards the source, and transmitted by the device. Any mismatch in the real or imaginary parts of the impedance of the source and the device generally results in a reflection of power. If the impedance is purely real, then the condition for zero reflection is that the source resistance exactly equal the device resistance,

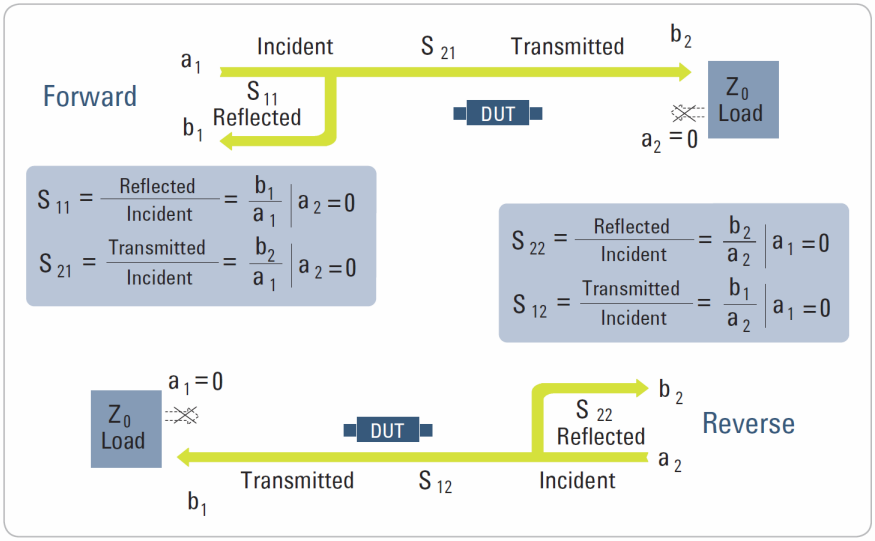


Figure 3.3: Relationship between two port S-parameters and energy. [56]

otherwise the condition is that the source impedance be the complex conjugate of the device impedance. At low frequencies, the resistance of a transmission line is relatively small and the measured current and voltage on said transmission line are roughly equal everywhere. When transitioning to high frequencies, the wavelength of the microwave signal becomes comparable to, or smaller than, the length of the transmission line, and it becomes beneficial to think of the microwave signal as a traveling-wave. Two port S-parameter measurements consist of four distinct measurements: S_{11}, S_{12}, S_{21}, and S_{22}. By convention, the first subscript indicates the port at which the energy is emerging from the device, and the second subscript indicates the port at which energy is entering the device. Therefore, measurements where both subscripts are the same indicate a reflection measurement (e.g. S_{11} and S_{22}, input and output reflection, respectively), and measurements where the subscripts differ indicate gain/loss measurements (e.g. S_{12} and S_{21}, reverse and forward gain/loss, respectively). Figure 3.3 shows how these measurements relate to transmitted and reflected energy.

The S-Parameters can also be converted to an ABCD matrix and used to extract the device parameters according to

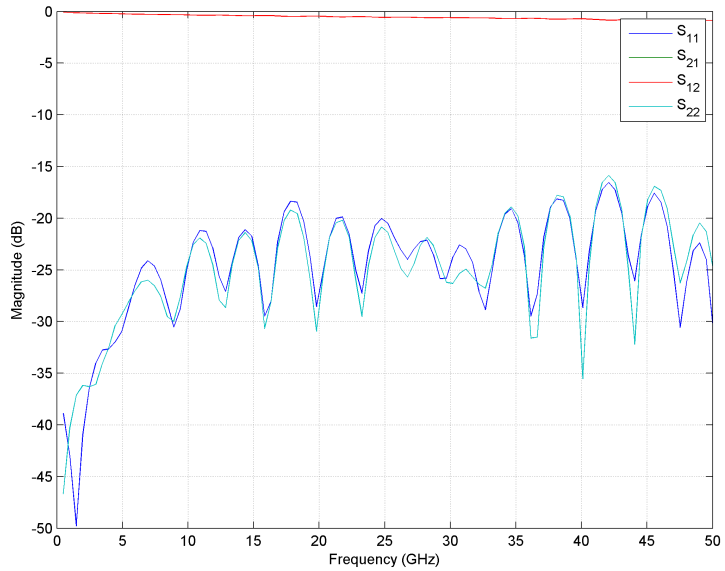
$$Z_d = \frac{2B}{D - A \pm \sqrt{(A+D)^2 - 4}} \tag{3.1}$$

$$\gamma = \alpha + j\beta = \frac{1}{L}\left[\ln\left(\frac{A+D}{2} \pm \sqrt{\left(\frac{A+D}{2}\right)^2 - 1}\right)\right] \tag{3.2}$$

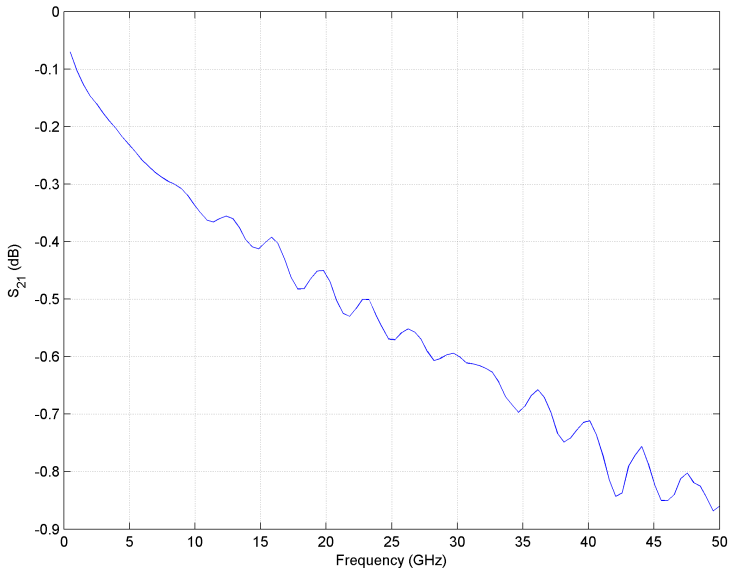
where Z_d is the device impedance, α is the microwave propagation loss, and β is the phase velocity [45].

In a scattering-parameter test many more options are available. Not only can the temperature, wavelength, and bias voltage be changed, but the frequency range and step size, along with the RF power can also be adjusted. RF power was set to the maximum allowable value, that is the highest value prior to the network analyzer producing a "Source Unleveled" error. That was done so as to maximize the signal-to-noise ratio. Because cabling, adapters, and probes have an effect on the scattering parameters, every cable, adapter, and probe used in line with the device has to be calibrated out, to ensure that the device under test (DUT) is only the MZM, and not the MZM in series with something else, e.g. cabling and electrical couplers. If something is unable to be calibrated out of the test setup, then it is possible to mathematically de-embed the response of that item should one have the S-parameters measurements on that item.

In the case of this test setup, the response of the microwave probe was not calibrated out due to the lack of a proper calibration substrate, so the S-parameters provided by the manufacturer for the specific serial number probe were used to de-embed its response from that of the device. In this manner, the S-parameters reflect, to the best of our ability, only the response of the MZM, and not of the test setup. The S-parameters for a GSG probe are shown in Figure 3.4. Importantly, the value for S_{11} is below -15 dB over the entire frequency range, while the value for S_{21} is greater than -1 dB over the same range.



(a) The complete two-port S-parameter response of the Cascade Infinity GSG 50-GHz probe, S/N 3V23G.



(b) S_{21} response

Figure 3.4: Note that the S_{21} and S_{12} parameters are overlaid exactly, and so the S_{21} response is broken out for increased visibility and clarity.

Unlike other experiments, temperature, wavelength, or bias voltage were not expected to have a large effect on the S-parameters. Within reason, the bias voltage should not alter the device's bandwidth, as an S-parameter test is fundamentally a small-signal test, and so changes in bias voltage which are meant to optimize ER, are largely unnecessary.

3.1.4 Power-Penalty Measurements.

The term power-penalty refers to the additional optical power necessary at the receiver to maintain a given BER with a change in either bit rate or transmission distance [57]. As such, a power-penalty measurement varies the received optical power using an in-line optical attenuator located after the MZM, and measures the BER for a given bit rate or a given transmission distance. Increasing the optical attenuation and decreasing the received optical power, the Optical Signal to Noise Ratio (OSNR) is decreased and the BER is increased. Measuring BER versus received power at different data rates primarily tests the ER of the MZM, and how well it can maintain an open eye, free from amplitude noise and timing jitter. Measuring BER versus received power at different transmission distances primarily tests the effects of dispersion induced by the optical fiber on the signal. The two variables, BER and received power are typically plotted using a logarithmic scale for the BER and then plotting the power in dBm. In some cases, the Q-factor of the BER is plotted rather than the BER itself. The amplitude noise associated with high and low logic levels follows a Gaussian distribution, and the Q-factor was initially used as an estimate of the BER, computed as a function of the mean and standard deviation in high and low logic levels. The fact that the Q-factor can be used to estimate BER by using only a random sampling of power levels in time, is important because that means BERs can be determined even while communications links are actively in use. This contrasts with a standard BER test, which requires a known data pattern to be transmitted, and the test itself can consume significant amounts of time [58]. The relationship between BER and Q-factor is given by (3.3) and (3.4). The frequency with which Q-factor is used to estimate BER means that in

some instances, even when the true BER is known, the Q-factor is instead plotted versus received power. Performing this mapping between BER and Q-factor results in a linear relationship between Q-factor and received power. When the power-penalty curves for two given data rates lie nearly atop one another, it indicates that the device is not bandwidth limited at those particular transmission speeds.

$$\mathrm{BER} = \frac{1}{2}\mathrm{erfc}\left(\frac{Q}{\sqrt{2}}\right) \tag{3.3}$$

$$Q = \frac{\mu_1 - \mu_0}{\sigma_1 - \sigma_0} \tag{3.4}$$

An issue encountered with the power-penalty measurements is the inclusion of ASE in the optical signal. Ideally, a tunable bandpass optical filter would have been inserted after the Erbium-doped fiber amplifier (EDFA) on the output side of the MZM, and prior to the Variable Optical Attenuator (VOA). However, the FC/APC facets on the bandpass filter were damaged, and introduced a large degree of instability in the power-penalty measurements, which became an uncontrolled random variable during experimentation. The variability in coupling loss between the bandpass filter and the VOA skewed measurements to the point that the bandpass filter had to be removed. As a result, the SNR of the optical signal decreased, and so a portion of the optical power was nothing more than noise.

3.2 Experimental Fixture

The overall setup consists of the silicon chip mounted between two linear stages which translate in X, Y, and Z — the stages holding the fiber which couples light into and out of the chip. From the very beginning, it was apparent that long-term optical coupling stability could become an issue if not addressed. 561D-XYZ stages from Newport Corporation were used in the setup; however, the standard fiber-holding block available from Newport would have had the fiber optic cable cantilevered out from the stage by a distance of 35 mm. Due to the lack of rigidity of fiber optic cables, random air currents would have easily jostled the fiber from its optimum coupling position, thereby causing power fluctuations during testing. To avoid such instabilities, a custom fiber-holding solution was designed and fabricated, which reduced the distance by which the fiber was cantilevered. A custom thermal mount to securely hold the chip in position and provide accurate temperature control was also fabricated. The overarching solution was first designed and simulated in a CAD environment to ensure compatibility, and then manufactured by the AFIT Machine Shop. With the custom fiber-holding mount, the distance over which the fiber is cantilevered can be reduced to as little as 500 μm, a reduction of 1/70th of its original length.

The fiber-holding mount itself is made of ferritic stainless steel, its ferritic property enabled magnetic clamps with elastomer bases were used to securely hold the fiber in place. As can be seen in Figure 3.5, the fiber-holding mount has a channel cut down the center in which the fiber rests. The fiber-holding mount is secured to the Newport 561D-XYZ stage by way of two cone point set screws on the side and a dovetail mounting rail on top of the 561D stage. A side view is also provided in Figure 3.6, which emphasizes the fact that the profile of the mount tapers from right to left to allow sufficient clearance over the chip mount. Note that two versions of the fiber-holding mount were designed and fabricated. Version one uses a chamfer taper at the end of the fiber-holding mount, and version two

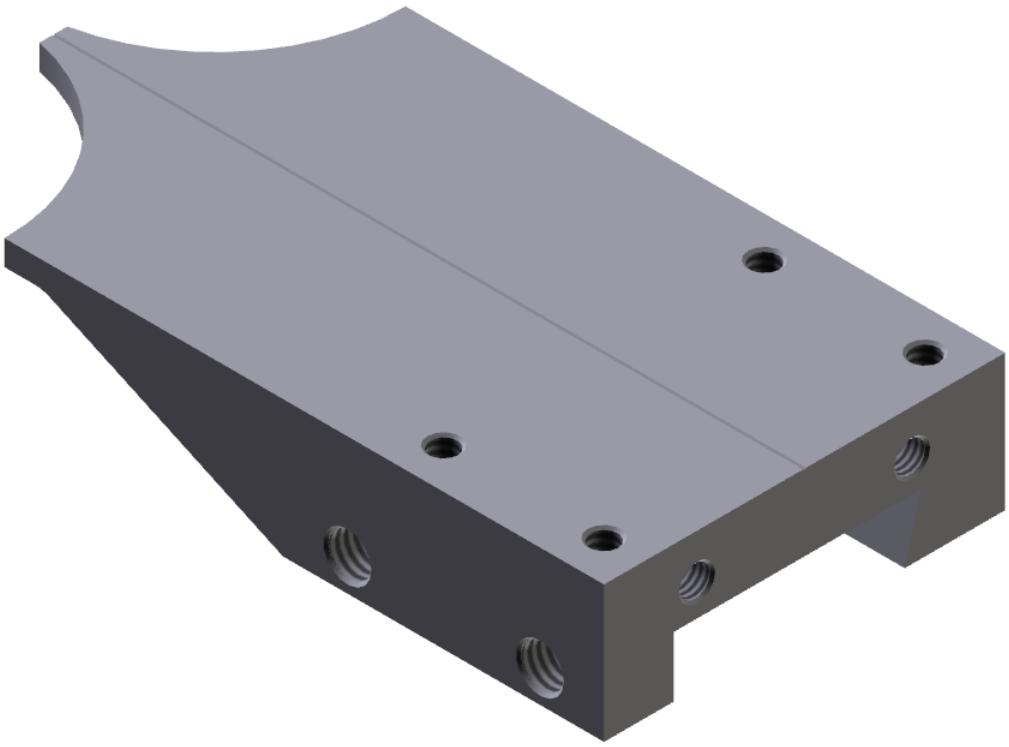


Figure 3.5: A perspective view of the custom fiber-holding mount.

Figure 3.6: A side view of the custom fiber-holding mount.

uses a fillet taper at the end of the fiber-holding mount. Version two is pictured in Figures 3.5 and 3.6, while version one is pictured in Figures 3.10 and 3.11.

Between the 561D stage and the fiber-holding mount is a tip/tilt platform, which allows us to correct for any deviations in tip or tilt by ±5°.

The chip mount is designed to hold a thermoelectric cooler (TEC) between two copper plates, and a temperature sensor in the top plate, above which the silicon chip sits. The chip mount is made of oxygen-free high-conductivity (OFHC) copper, which provides excellent thermal conductivity, as needed by the TEC. The chip mount is then affixed to a stainless steel stand which provides adequate spacing around the copper blocks for natural convection cooling. The entire chip mount is shown in Figure 3.7. The silicon chip rests in a small depression milled into the raised ridge on the top copper block, and is affixed with thermal compound to ensure good thermal conductivity between the silicon chip and the top copper block. For reference, the silicon chip contains a number of devices, and measures 7.5 mm by 26 mm. The ridge in the top copper clock holds the temperature sensor, so that the sensor sits directly between the TEC and the silicon chip. The leads of the TEC can be seen extending from the back of the chip mount in Figure 3.7. Thermal mass on the top copper block was minimized to ensure the temperature was uniform throughout the entire copper block.

The exploded view in Figure 3.8 provides a clear view of the clearance hole in which the temperature sensor is located. The temperature sensor is secured in place using thermal epoxy, once again to ensure good thermal conductivity between the top copper block and the temperature sensor. Thermal compound is also used between the TEC and top/bottom copper block interfaces rather than thermal epoxy, the TEC being the square device sitting between the top and bottom copper blocks in Figure 3.8. This is so that the TEC is replaceable in the event of a catastrophic failure. The top and bottom copper blocks are then held together using #0-80 socket cap alloy steel screws, which unfortunately had the effect of reducing thermal isolation between them. In applications where a large degree

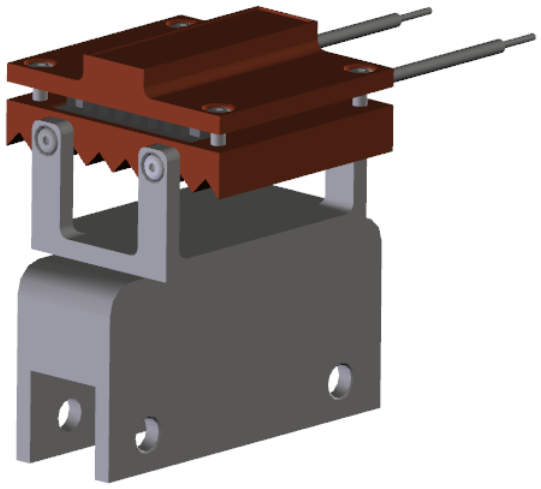


Figure 3.7: A perspective view of the custom thermal chip mount.

of cooling is needed, this might become a problem, as a TEC fundamentally operates by drawing heat from one interface (the cold side) to the other interface (the hot side).

In the experimental setup, the cold side is the side contacting the top copper block. This means that the alloy steel screws would allow heat to leak from the hot side to the cold side, potentially heating the top copper block to the point where the TEC would no longer be able to properly control the temperature of the thermal stage. The experiments did not need such a level of cooling though, and as such were not affected by the aforementioned problem. The lack of thermal isolation could be solved by replacing the alloy steel screws with nylon screws (which would require fabricating another mount, as nylon screws are not available in #0-80 size), or by using thermal epoxy, rather than thermal compound, between the TEC/copper interfaces.

The bottom copper block is milled with triangular grooves to increase the surface area, as the bottom copper block is a heat sink for the hot side of the TEC, and must radiate heat away as fast as possible. Increasing the surface area of the copper block increases

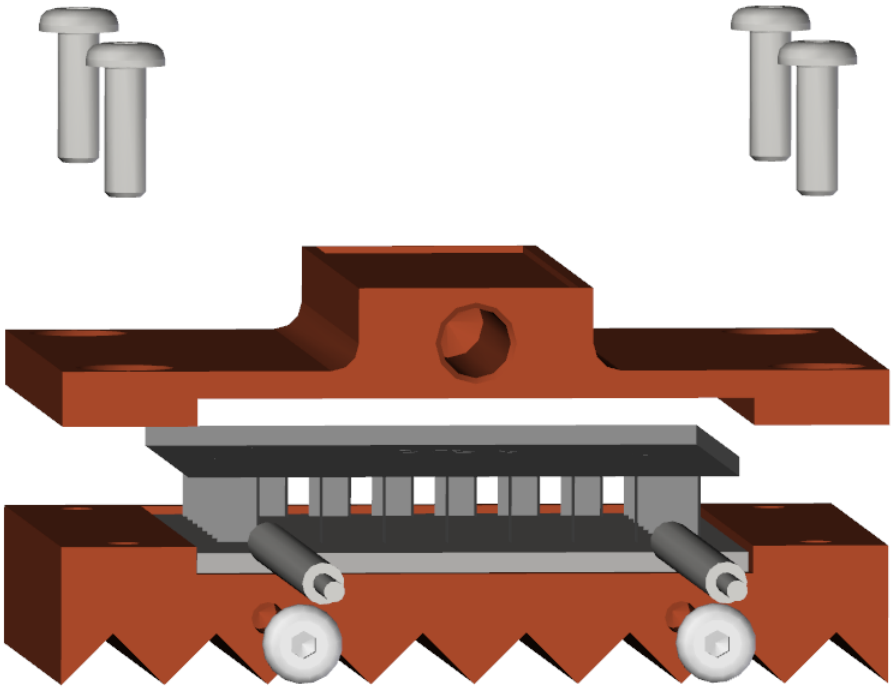


Figure 3.8: An exploded view of the custom thermal chip mount.

the amount of heat that can be radiated away from the copper block per unit time. Another design option would be to use straight fins as opposed to triangular grooves; however, using straight fins increases fabrication complexity and may increase material waste if the copper piece were milled from a single copper block instead of welding the fins onto a copper block. Welding on the fins would require a considerable amount of time, and would be less precise than milling the piece from a single copper block. Another option would be to use forced convection as opposed to natural convection (i.e. use a fan to blow air over the copper block); however, the use of a fan so close to the fiber optic cables would likely disturb coupling.

Together, the entire chip/fiber-holding simulated solution can be seen in Figures 3.10 and 3.11, and in practice in Figures 3.12, 3.13, and 3.14. The steel mount that holds the

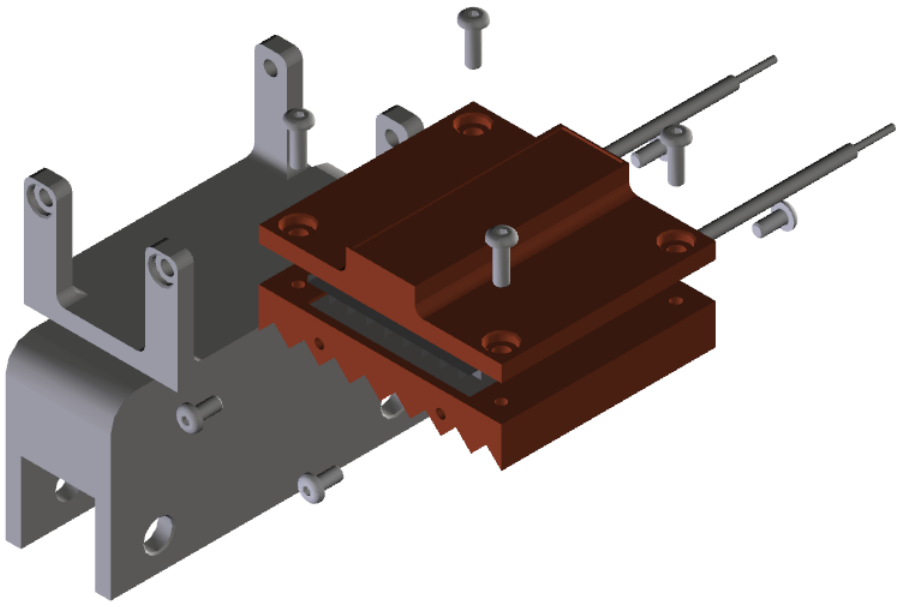


Figure 3.9: An exploded view of the custom thermal chip mount.

chip mount is resting atop a Newport 561-TS, which is a steel T-stand. The mounting plate atop the T-stand was removed, and the custom steel mount was bolted directly onto the T-stand.

The 561D-XYZ has crossed-roller bearings and three positions for micrometers, which allows the stage to translate in all three dimensions of Euclidean space. Rather than using hand-adjusted micrometers, precision picomotors from New Focus were selected, and provided a step resolution of <30 nm. The logic behind the decision to use picomotors is primarily based on the finer step resolution, as the misalignment tolerance with the waveguide is small. Additionally, the use of picomotors enabled the implementation of a feedback loop within AutoLab that automatically aligned the lensed fibers to the

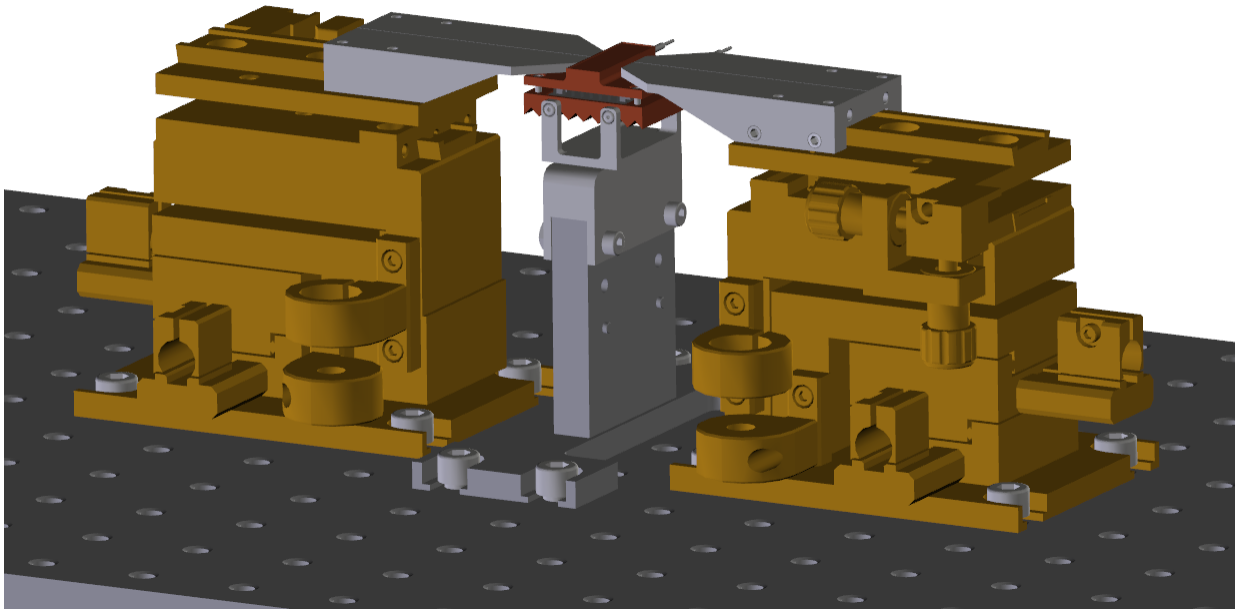


Figure 3.10: The overall chip/fiber-holding solution simulated in a CAD environment.

chip's optical waveguide. To illustrate the magnitude of the issue, it is beneficial to say that the waveguide measures 560 nm by 240 nm, and the smallest spot size lensed fiber commercially available has a spot size, ω_0, of 2.5 μm – dimensions which are very large compared to that of the waveguide. Figure 3.15 shows the lensed fiber optimally coupled to the top waveguide of an MZM, and illustrates the sheer difference in size between waveguide and fiber. Note though, that the entire fiber measures 125 μm in diameter, whereas the core is only 8 μm in diameter, so the image may overstate the size difference.

The general procedure for fiber alignment is to position the fiber in line with the waveguide using the microscope view, and then adjust the height to ensure that both the tip of the fiber and the waveguide are simultaneously in focus. The coupling is then optimized by peaking the output power received by a photodetector in each axis. Coupling

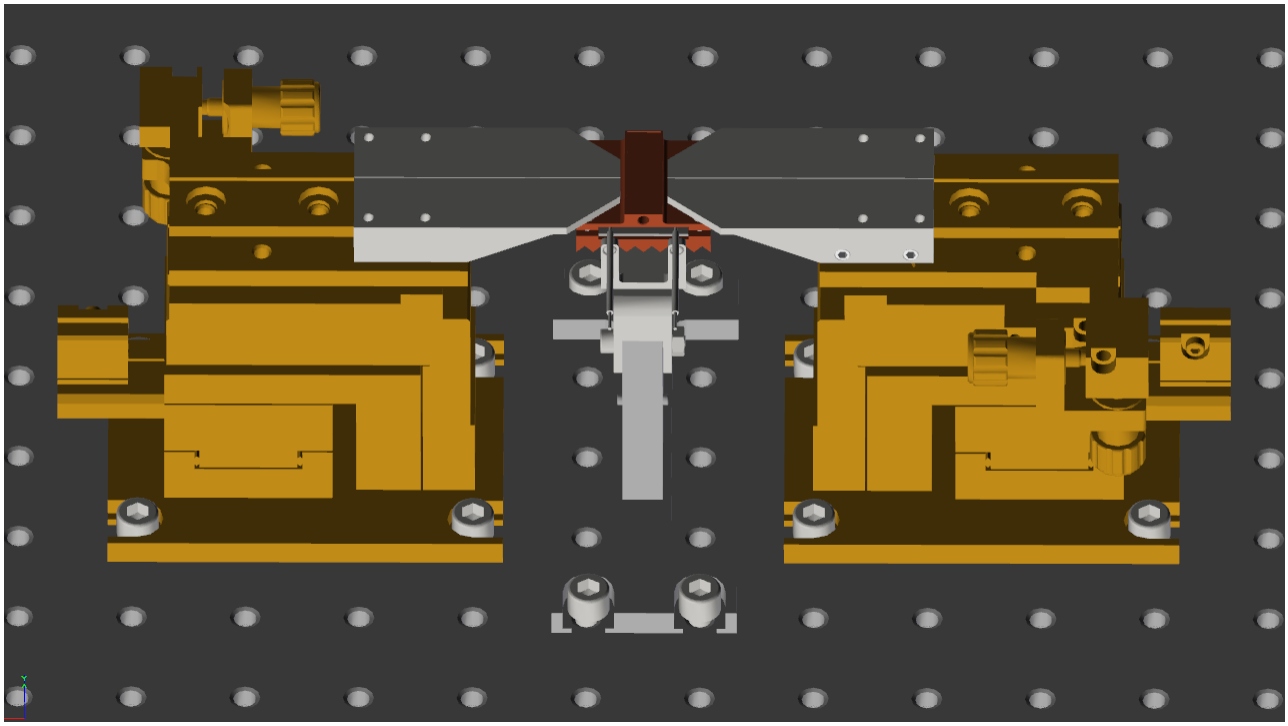


Figure 3.11: The overall chip/fiber-holding solution simulated in a CAD environment.

optimization has to be performed for both input and output fibers, and is an iterative process. Because the picomotors can be controlled remotely via Ethernet, a module was added to AutoLab to automate power optimization. Unfortunately, the picomotors are open-loop and not closed-loop motors, so a recursive search algorithm was developed to ensure that the power optimization algorithm accurately and precisely finds the optimum coupling each time. The search algorithm starts by translating a number of steps, $x/2$, past the waveguide, and then moves x steps in total, one at a time, recording the output power reading at each step. After completing the full movement, the algorithm then steps back to the location where the highest power reading was found.

Because of the fact that the motor is open-loop, the position to which the algorithm returns is not guaranteed to be the same as that where the maximum power reading was

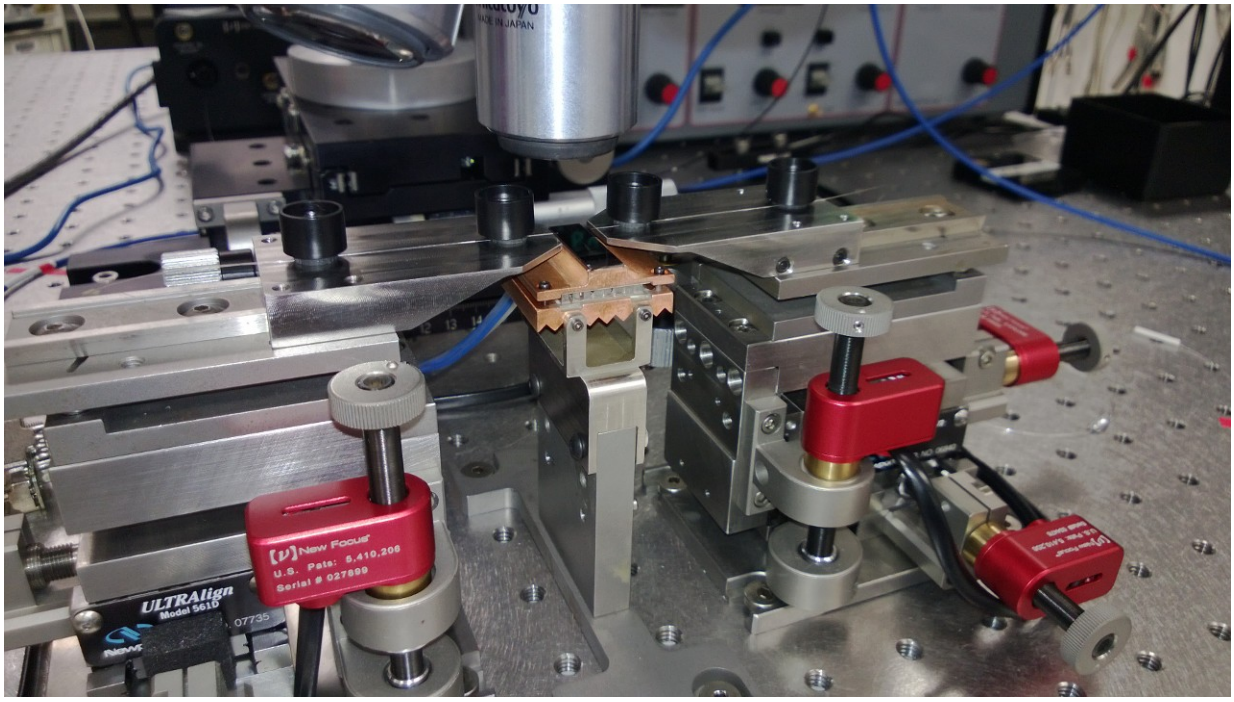


Figure 3.12: The overall chip/fiber-holding solution. Note the use of the first design for the fiber-holding mount, and the use of elastomer magnetic clamps to secure the fiber to the fiber-holding mount.

taken. Therefore, the algorithm steps $x/4$ steps beyond that point, and then moves $x/2$ steps in total, one at a time, recording the output power reading at each step, and at the end returning to the location where output power was maximized. This process is repeated until such time as $x/n \leq 1$, at which point the fiber is optimally aligned to the waveguide in one axis. The search algorithm is run on each axis independently, to ensure optimal coupling in each axis, and run iteratively on each axis. After initially optimizing power in X, Y, and Z, the search algorithm is run again on each axis with smaller starting search area sizes (i.e. the value of x decreases during each round of power optimization). A variable number of rounds of power optimization can be run, though the number was typically five, and negligible improvement in coupling was witnessed beyond round three.

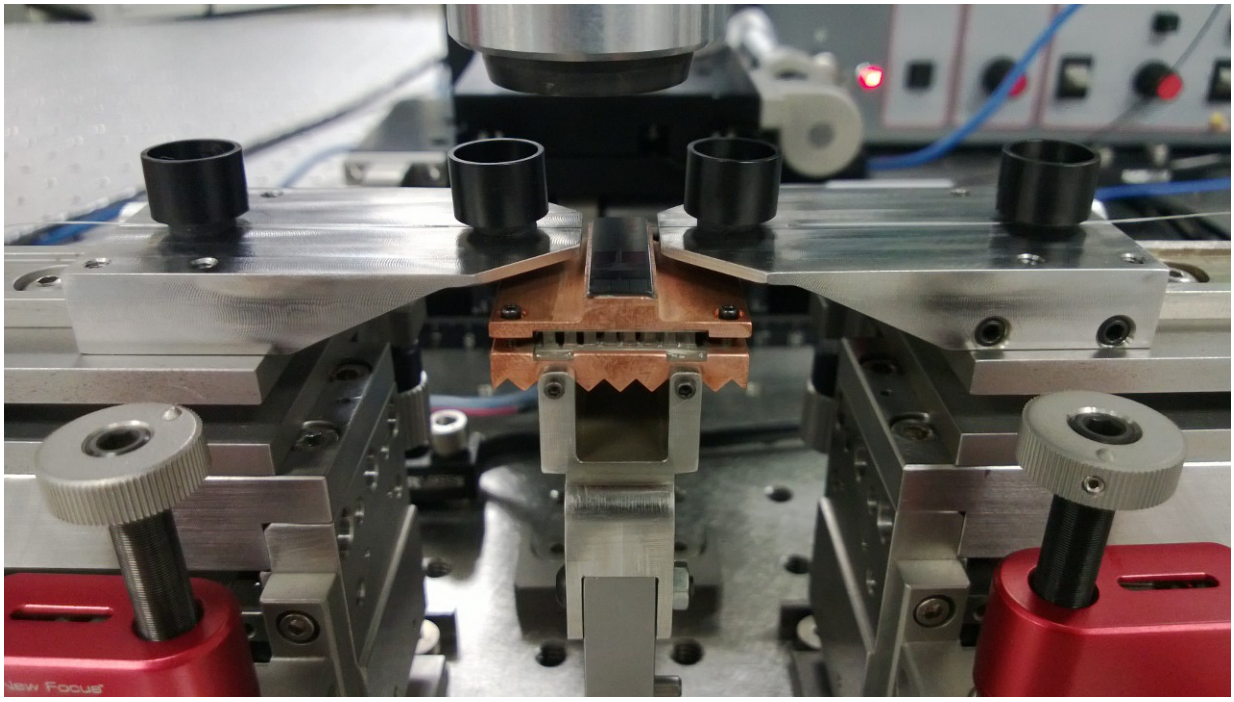


Figure 3.13: The overall chip/fiber-holding solution.

The use of picomotors also allowed us to automate movement between top and bottom waveguides, which takes place using translation in the plane parallel with the surface of the silicon chip (in the plane of the paper when looking at 3.15) followed by a power optimization routine. With this setup, it is possible to completely test all combinations of input and output waveguides (see Table 3.4 for all combinations) automatically without the need for human intervention. Due to the complementary nature of the 2x2 adiabatic coupler, the Input-Top and Output-Top state is identical in behavior to the Input-Bottom and Output-Bottom state, just as the Input-Top and Output-Bottom state is identical in behavior to the Input-Bottom and Output-Top state. Because the specific input arm is irrelevant to the operation and performance of the device, and it is only the combination of input and output arms that matter, the terms cross and straight ports will also be used. The straight port refers to a situation where the input is on the top arm and the output is on the top arm,

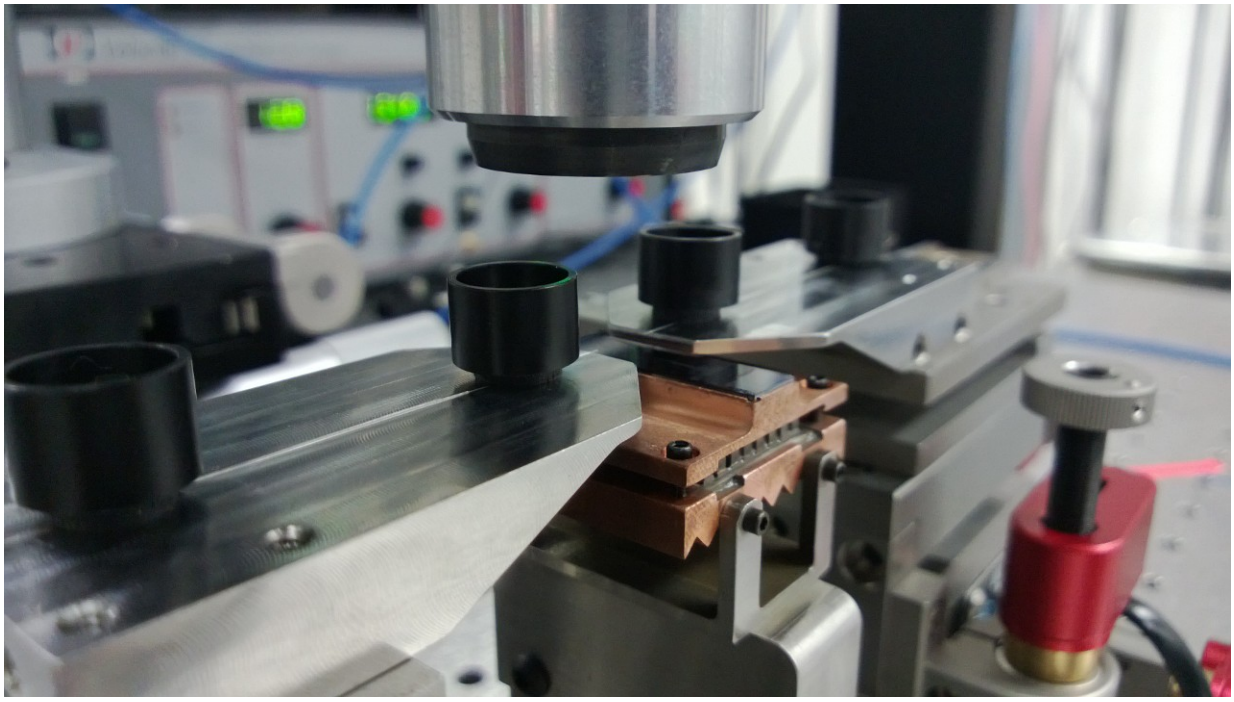


Figure 3.14: The overall chip/fiber-holding solution. The bottom of a microscope lens assembly can be seen at the top of the photo, centered above the silicon chip.

or alternatively when the input is on the bottom arm and the output is on the bottom arm. The cross port refers to a situation when the input is on the top arm and the output is on the bottom arm, or alternatively when the input is on the bottom arm and the output is on the top arm.

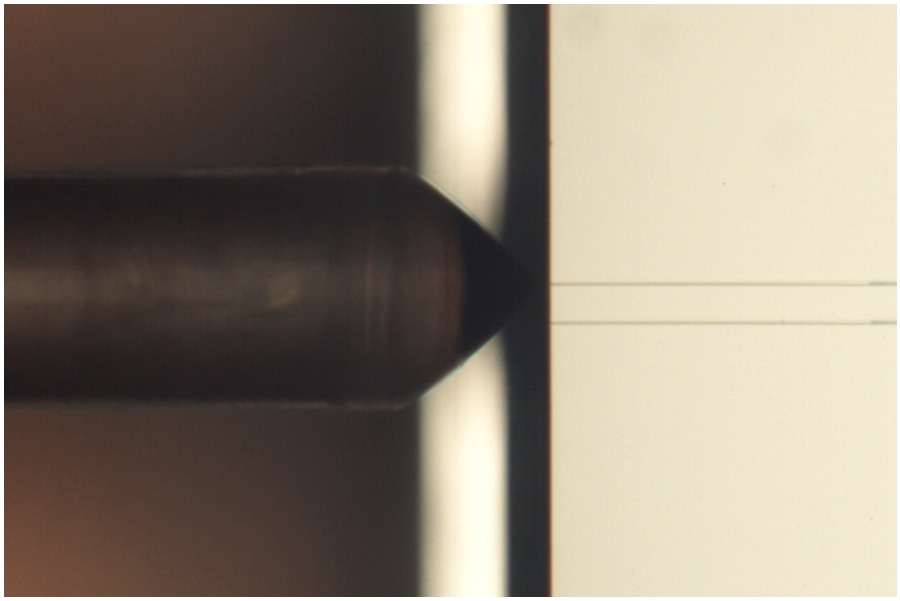


Figure 3.15: A microscope view of the lensed fiber aligned on a waveguide of the MZM. There are in fact two waveguides present in the image, both of which appear as horizontal lines on the right of the image. The lensed fiber is on the left side of the image, and appears to terminate in a conical shape. In this image, the fiber is optimally coupled to the top waveguide.

Table 3.4: All possible combinations of input and output waveguide arms. Throughout the remainder of this document it will be made clear whether the optical input and output are made on the top or bottom waveguides of the MZM.

Input Waveguide	Output Waveguide	Combination Name
Top Arm	Top Arm	Straight
Top Arm	Bottom Arm	Cross
Bottom Arm	Top Arm	Cross
Bottom Arm	Bottom Arm	Straight

3.3 Experimental Setup

In all cases, a tunable laser diode, the New Focus Velocity 6328, was used as a seed laser. The laser has an operating range of 1512.3 nm to 1583.4 nm, a specified linewidth of 300 kHz, and was fiber coupled to a polarization-maintaining (PM) fiber. The output of the laser was then fed to a PriTel PM EDFA with an optical noise figure of <6 dB. The output of the EDFA then went to a Newport Fiber Polarization Controller, whose operation is based on a Babinet-Soleil compensator. While the squeezing mechanism of the polarization controller does not provide precise control of the output, the all-fiber design of the controller has significantly less loss than a free-space quarter-wave, half-wave, quarter-wave polarization rotator. The loss in a free-space polarization controller could be partially compensated by increasing the power to the EDFA, however doing so also increases the noise floor in the optical signal. Prior to the polarization controller, everything was in PM fiber, and everything after the fiber controller is non-PM. With less than 4 m of fiber after the polarization controller, the optical signal remains in a linearly polarized state, with 99 % of power in the TE mode. The previous value was measured by using a polarizing beam splitter and maximizing the amount of power in one arm of the output. The output of the polarization controller then went to a 90:10 fiber splitter, with the 10 % arm going to a power meter and the 90 % arm going to a lensed fiber with a 3.5-μm spot size, and finally the silicon chip. The power meter on the input side is used to 1) determine the amount of power sent to the DUT, and 2) allow AutoLab to tune the input power so that it remains constant throughout the experiment. While maintaining constant power delivered, AutoLab also takes into account the fact that the fiber splitters do not display a constant splitting ratio – the true splitting ratios were experimentally determined, and are given for the 90:10 fiber splitter in Figure 3.16, and for the 99:1 fiber splitter in Figure 3.17. The experimental setup on the input side of the MZM remained constant, with the exception of testing S-parameters, during which the wavelength was fixed and a Santec tunable notch

filter was inserted after the EDFA to reduce noise at other wavelengths. The GSG pads were contacted with an Infinity 50-GHz probe from Cascade Microtech, which was in turn connected to a bias tee with an electrical bandwidth of 50 kHz-65 GHz. The DC input to the bias tee was provided by a Model 523 Precision DC source from Krohn-Hite, with a 10 nV resolution and a 4 ppm accuracy. Different RF sources were connected to the RF input of the bias tee, as dictated by the experiment type. The choice of bias tee proved to be important, as the different bias tees available all had an impact on the RF signal that was transmitted to the DUT. Eye diagrams for each of the available bias tees are given in Figure 3.18, and as the bias tee from Anritsu produced the cleanest output signal, it was the one selected for use in the experimental setup.

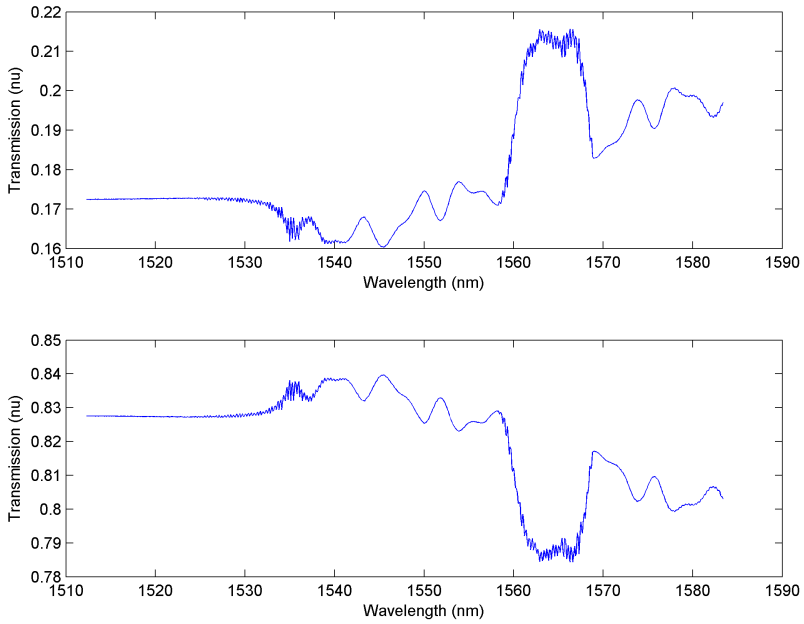


Figure 3.16: Experimentally determined values for the 90:10 fiber splitter, which was used on the input side of the DUT.

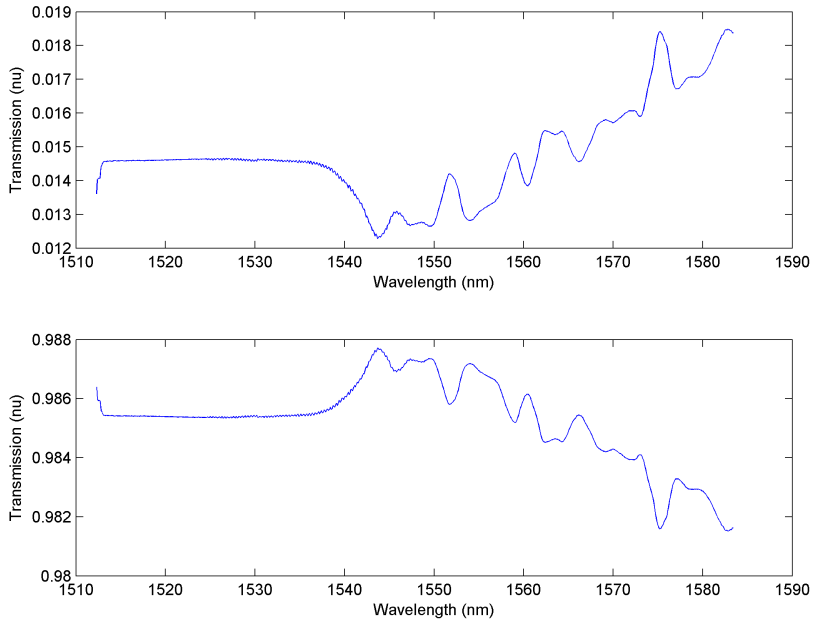


Figure 3.17: Experimentally determined values for the 99:1 fiber splitter, which was used on the output side of the DUT.

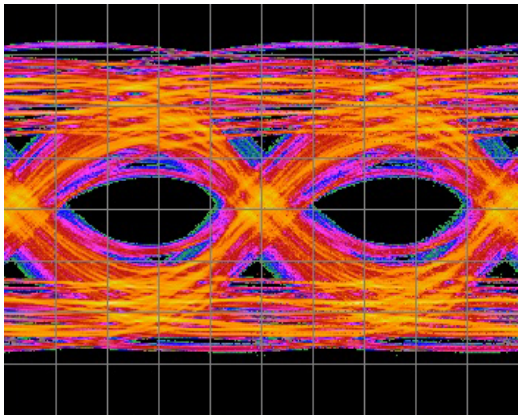
An image of the experimental setup with the GSG probe in use is shown in Figure 3.19. Note that the elastomer magnetic clamps are no longer in use holding the fibers, as doing so impedes the ability to use a GS or SG probe in conjunction with the TWE MZMs. Instead, the fibers are taped to the fiber-holding mounts.

3.3.1 DC Characterization.

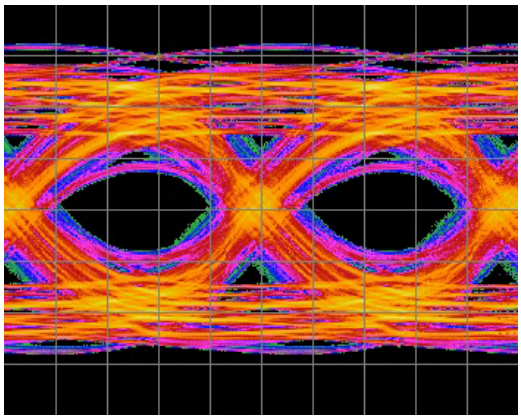
The setup for DC characterization experiments is identical to that described above, with the exception of the output. The output of the MZM is coupled via lensed fiber with a 2.5-μm spot size, and from there goes directly to a power meter.

3.3.2 Bit Error Rate Characterization.

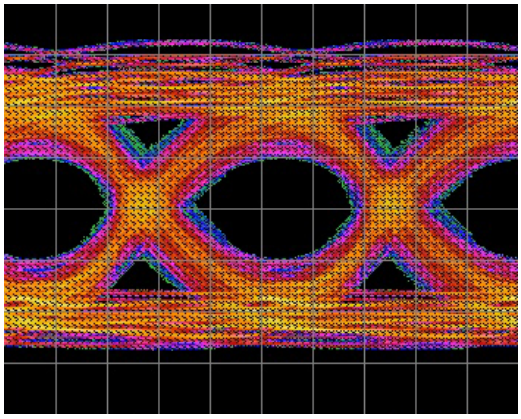
During BER characterization, the RF input to the bias tee is provided by a Tektronix BSA286C BERTScope. Although the stated bandwidth of the BERTScope is 28.6 Gbps,



(a) Picosecond Pulse Labs



(b) Agilent



(c) Anritsu V255

Figure 3.18: The effect of various bias tees on the RF signal of the BERTScope at a bit rate of 26 Gbps using a PRBS-7 test pattern. In these tests, the RF output of the BERTScope was connected directly to the RF input of the bias tee, and the output of the bias tee to the error detector on the BERTScope.

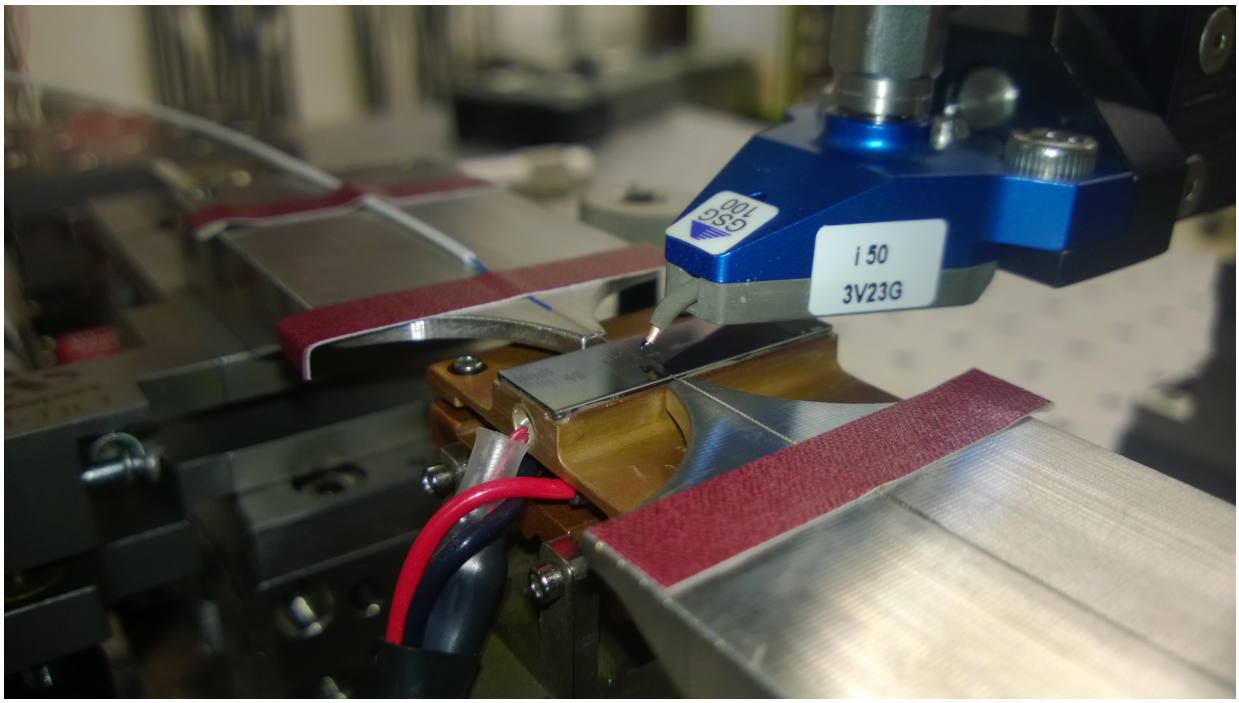


Figure 3.19: The experimental setup shown in use with the Infinity GSG probe. The input fiber is on the upper left, with a blue acrylate coating and a white buffer, and the output fiber is on the lower right with a clear acrylate coating. The input fiber is non-PM with a moderately larger spot size than the PM output fiber.

the error detector is limited to 26 Gbps. The BERTScope has an adjustable frequency and driving voltage, with an upper limit on driving voltage of 2-V_{pp}. The output of the DUT in this case goes directly to an EDFA, and the EDFA outputs to a 99:1 splitter, with the 99 % arm going to a photodetector, and the 1 % arm going to a power meter. The photodetector used was a u^2t Photonics XPDV2120RA, with a bandwidth of 50 GHz and a high damage threshold. The high damage threshold of the photodetector allowed us to avoid using an RF power amplifier between the photodetector and the BERTScope. Indeed, when an RF power amplifier was used, it became the limiting factor in the performance of the experimental setup. The eye diagrams for the available RF power amplifiers are shown in Figure 3.22.

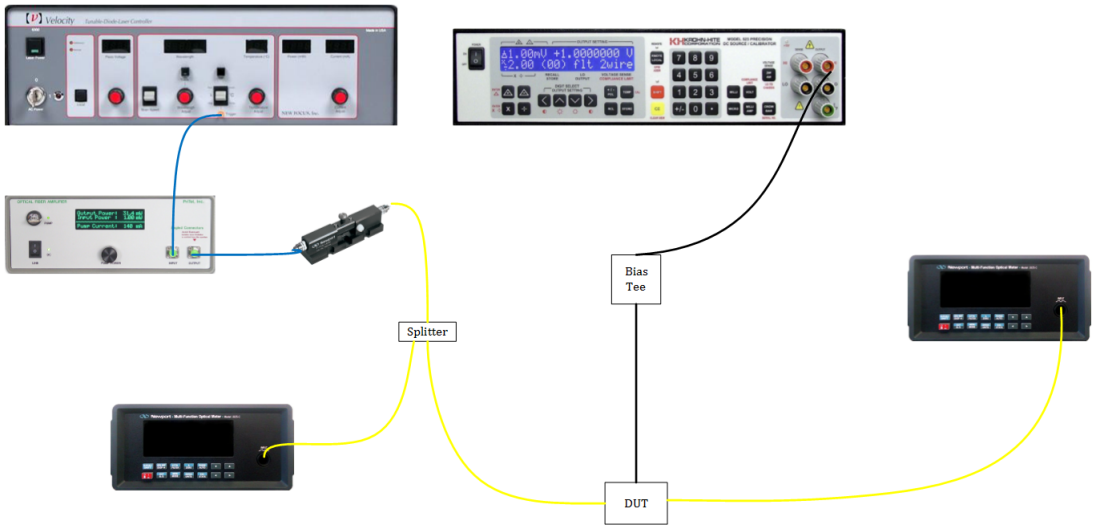


Figure 3.20: The experimental setup during DC characterization experiments.

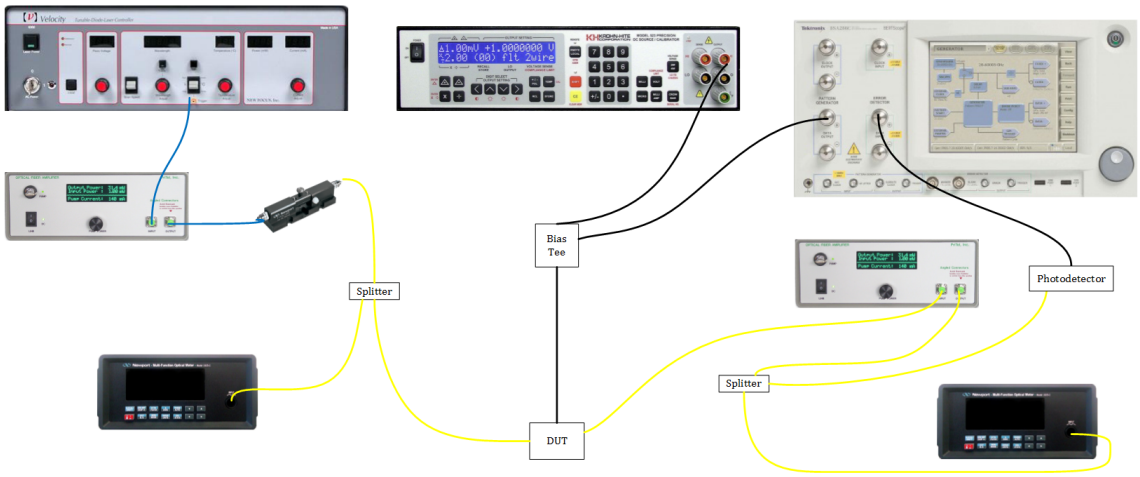


Figure 3.21: The experimental setup during Bit Error Rate Characterization experiments.

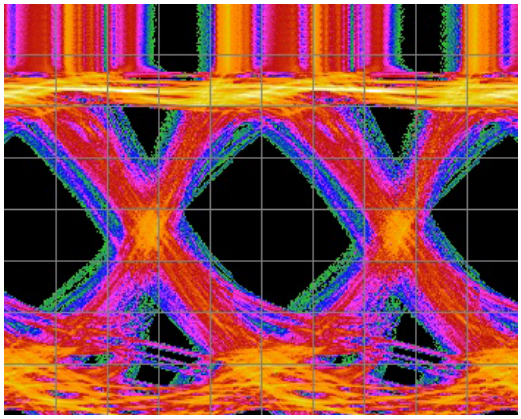
3.3.3 Bandwidth Characterization.

During bandwidth characterization, the RF input to the bias tee is provided by Port 1 of an Agilent E8361C Network Analyzer (PNA) which operates from 10 MHz to 67 GHz. The output of the DUT in this case goes directly to an EDFA, and the EDFA outputs to a 99:1

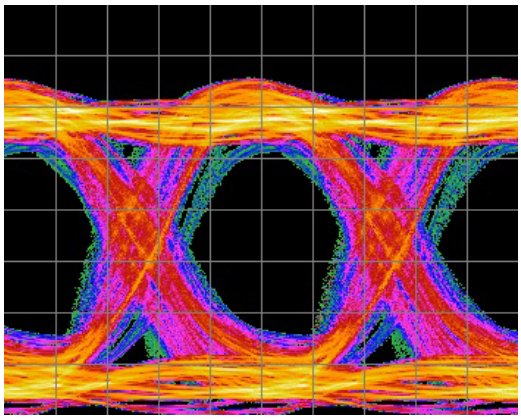
splitter, with the 99 % arm going to an Agilent N4373C Lightwave Component Analyzer (LCA), and the 1 % arm going to a power meter. The electrical output of the LCA was then connected to Port 2 of the PNA. The use of an EDFA was necessitated by the fact that the signal into the LCA was in the noise floor of the LCA's photodetector. Importantly, during S-parameter characterization, the operating wavelength was fixed, and an optical tunable notch filter was used at the output of the polarization controller.

3.3.4 Power-Penalty Measurements.

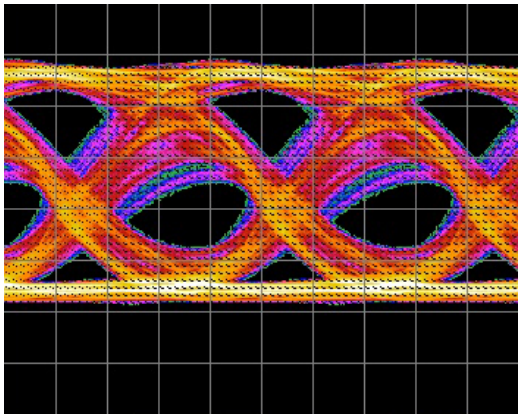
The experimental setup during power-penalty measurements is very similar to the setup during BER Characterization, the exception being a variable optical attenuator (VOA) inserted in-line between the EDFA on the output side of the MZM and the 99:1 fiber splitter. The VOA was manufactured by OZ Optics, fiber-coupled, and has an insertion loss of 1.5 dB. The VOA is adjusted by hand and has two collimated lenses that are well-aligned. An adjustment screw on the body of the VOA varies the position of a beam block between the two collimators, and can therefore affect the amount of power coupled between the input and output collimators.



(a) Agilent 83006A, 10 MHz to 26.5 GHz



(b) Agilent 83051A, 45 MHz to 50 GHz



(c) Picosecond Pulse Labs

Figure 3.22: The effect of various RF power amplifiers on the RF signal of the BERTScope at a bit rate of 26 Gbps using a PRBS-7 test pattern. In these tests, the RF output of the BERTScope was connected directly to the amplifier, and the output of the amplifier to the error detector on the BERTScope.

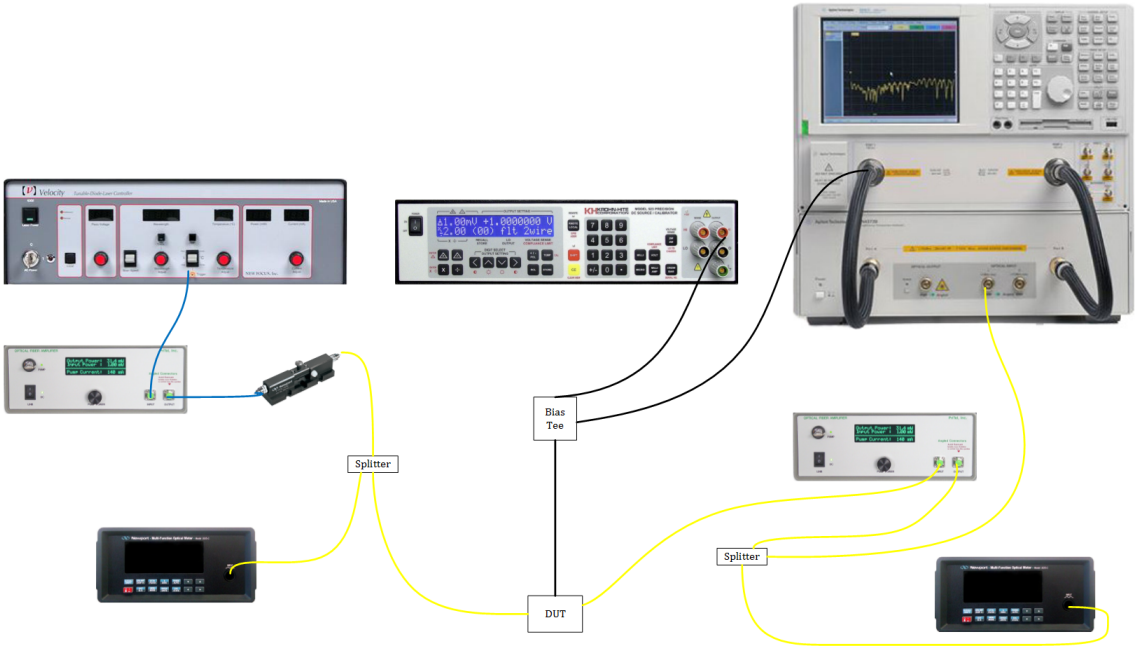


Figure 3.23: The experimental setup during Bandwidth Characterization experiments.

IV. Results and Analysis

4.1 DC Characterization

It is beneficial to start with the simplest presentation of data, and move on to more complex datasets later. As noted in Chapter 2, it is the application of a voltage to the p–n junction that causes a change in the modulation depth of the MZMs under test, and thus during operation of the MZM, the driving voltage changes with the application of a microwave signal. The DC characterization is a screening experiment performed by changing the DC bias of the device, and thus the electric field across the active region. In the push-pull MZM design, the device sees no difference between the application of a DC bias, and the application of a microwave signal, because both signals are applied via the same set of pads. Thus, a number of trials at different DC biases can predict device performance at microwave speeds within the operational bandwidth limits of the device.

Looking at Figure 4.1, the transmission of the MZM does indeed change with the applied voltage. This particular figure is taken from data on the 2000-μm PP MZM, and the blue and green lines show how normalized transmitted power out of the top or bottom arms (Channels A and B) changes with the bias voltage. It was stated earlier that the top and bottom output arms are complementary as a result of the use of a 2x2 adiabatic coupler, and this fact is demonstrated in Figure 4.1. As the power transmitted to Channel A increases, the power transmitted to Channel B decreases. The red line is the sum of the values represented by the blue and green lines, showing that the total output power is not constant across all bias voltages. That fact might initially be surprising, as the complementary nature of the device would lead one to assume that power not being transmitted to Channel A must be transmitted to Channel B, and so the output power must be conserved. It is important to remember that the device is fundamentally a Mach–Zehnder interferometer though, and so it is to be expected that destructive interference will lead to a decrease in total output

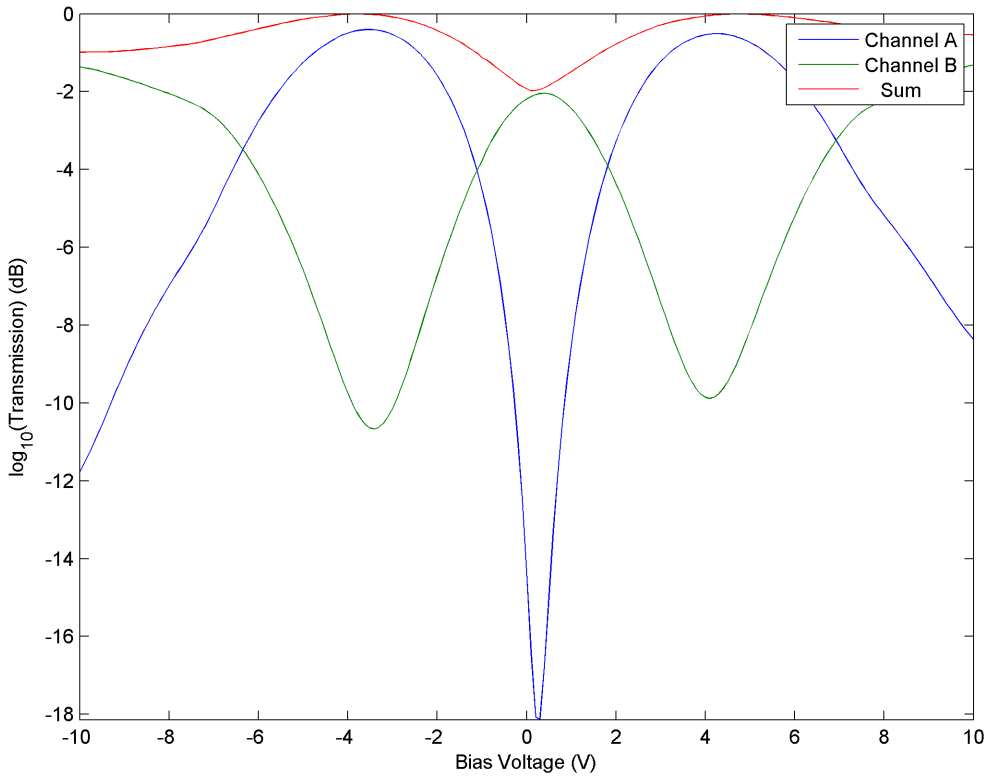


Figure 4.1: The output of the 2000-μm PP MZM, showing the complementary nature of the top and bottom output arms. The output transmission is normalized, and given on a logarithmic scale at λ=1550 nm.

power at some bias voltages. Add to that the effects of FCA, and it becomes clear that the total output power will not be constant. Importantly though, the minima in Figure 4.1 corresponds to the minima in Figure 4.2, which is precisely as expected.

Another consideration is that the data presented here is in fact taken from two experiments. In the first experiment, the input fiber was aligned to the top input arm, and the output fiber was aligned to the top output arm. In the second experiment, the input fiber remained aligned to the top input arm, and the output fiber was aligned to the bottom

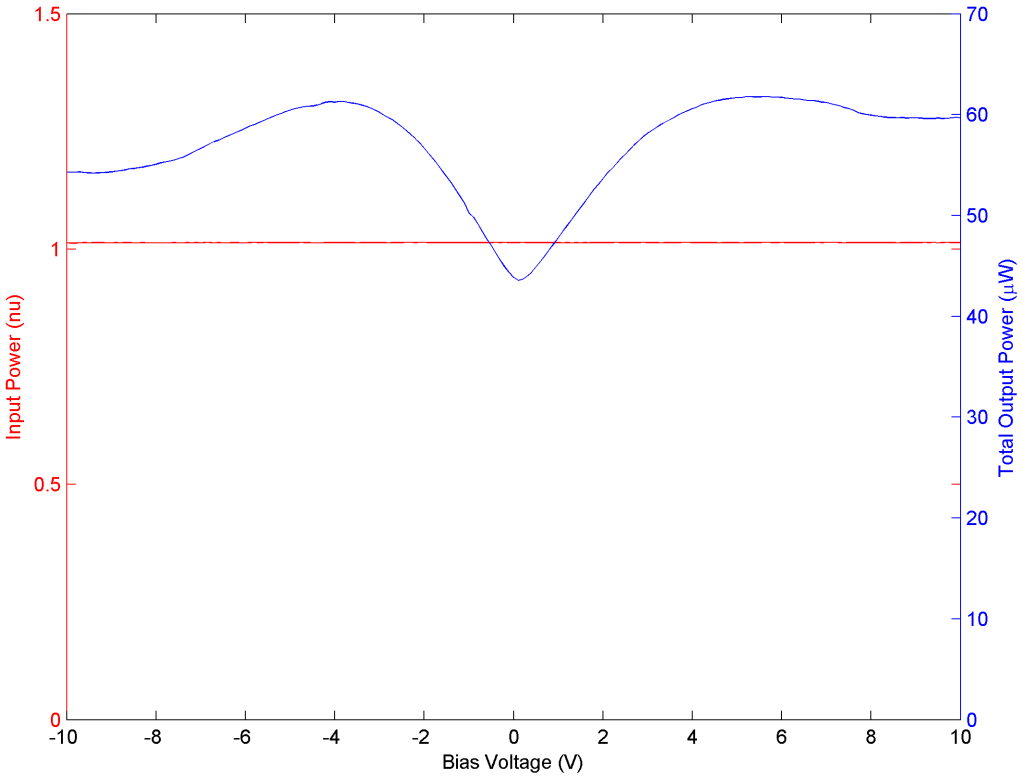


Figure 4.2: A ratio of the input powers to the 2000-μm PP MZM, which shows that the input powers across experiments was roughly equal. The total output power between the two output arms is also shown in terms of real power rather than normalized units.

output arm. Because the data was taken over the course of multiple experiments, it is important that the input power be the same across all experiments so that the results can be compared directly. Figure 4.2 shows the ratio of the input power in the first experiment, to the input power in the second experiment (see the red line). If the ratio of the input powers remains 1 across all bias voltages, then the input power at each bias voltage is exactly the same for both experiments. Any deviation from a value of 1 indicates that the input powers were not exactly the same in both experiments. AutoLab was configured to allow for 1.5 % tolerance for deviation in input power. Smaller tolerances increase the time required to run

an experiment, as more time must be spent tuning the laser power. The given tolerance of 1.5 % is sufficiently small to allow for direct comparison of results without significantly impacting the experimental run time. In this instance, the ratioed value remained constant at 1.014. The blue line in Figure 4.2 shows the total output power, the same data as the red line in Figure 4.1, except in terms of real power and not normalized units.

Returning to Figure 4.1, there are turning points at −3.5 V, 0.4 V and 4.3 V, and all turning points for Channels A and B occur at the same bias voltage, so that when Channel A is at a maxima, Channel B is at a minima, and vice versa. The data also indicates that the value for V_π in this device is 3.9 V. At a length of 2000-μm, that puts the value for $V_\pi L$ at 0.78 V cm. During high-speed operation, the data suggests it would be best to operate the MZM at a bias voltage of −1.55 V so that the device is biased in the linear region of operation, thereby maximizing modulation depth. With a symmetric driving voltage swing of 3.9-V_{pp} around the bias point of −1.55 V, the modulation depth will be at its maximum. If taking the output from Channel A, the logic is inverted, so that the application of a positive voltage swing reduces the output power, whereas taking the output from Channel B produces normal logic, and the application of a positive voltage swing increases the output power. Whereas the values given in Figure 4.1 show the normalized transmission on a logarithmic scale, Appendix A gives the same values on a linear scale. If operating at the specified bias voltage with the specified driving voltage, an ER of 17.73 dB is possible on Channel A, and an ER of 8.63 dB on Channel B.

The MZM was tested from −10 V to 10 V because the junction begins to break down outside this range. The nature of the output is very much sinusoidal, as one would initially expect; however, as can be seen by looking at Figure 4.1, the phase efficiency of the device decreases with the application of increasingly larger voltages. In other words, while it might only take 3.9 V for the first π phase shift, the second π phase shift would be at greater than $2 \cdot 3.9 = 7.8$ V. The decrease in phase efficiency is clearly evidenced by the fact that

there is not another turning point at either −7.4 V or 8.2 V; behavior which is caused by the fact that the optical mode is centered on the junction — therefore the junction is initially depleted of carriers in the region where the optical mode has greatest intensity, and further expansion of the depletion region overlaps with areas where the optical mode is less intense. As such, the effective index of the waveguide changes less drastically with increasingly large voltages. The reduction in phase efficiency is not necessarily a disadvantage. In fact, the reduction in phase efficiency speaks to the optimization of device design and the fabrication process – anything more than a single turning point is extraneous, as operating at larger voltages would not only be inconsistent with CMOS supply levels in the International Technology Roadmap for Semiconductors (ITRS), but would also result in greater power dissipation.

The sinusoidal nature of this specific device's operation is in part due to the length of the device, as a reduction in the length of the device necessitates larger voltages to achieve a π phase shift. The breakdown voltage is fixed regardless of active region length, so longer devices can achieve greater phase shifts before the junction breaks down. While the output of the 2000-μm PP MZM looks to be sinusoidal, smaller devices will have an output that looks less sinusoidal, and perhaps more like a $\mathrm{sinc}^2\, x$ function. The same two sets of data as presented above in Figures 4.1 and 4.2 are presented for the 1000-μm PP MZM in Figures 4.3 and 4.4. That is to say, Figure 4.3 shows the complementary nature of the outputs of the 1000-μm PP MZM in logarithmic form, while Figure 4.4 shows the relationship of the input power between experiments, and the total output power expressed as real power.

By way of comparison with the 2000-μm PP MZM, the 1000-μm PP MZM does not have nearly as sinusoidal an output, and has only two turning points on the output, as opposed to three. The reduction in the number of turning points is attributable to the decrease in active region length, as mentioned in the previous paragraph. At a voltage of 10 V, the depletion region width is expanded from 69.44 nm at 0 V to 239.65 nm, and

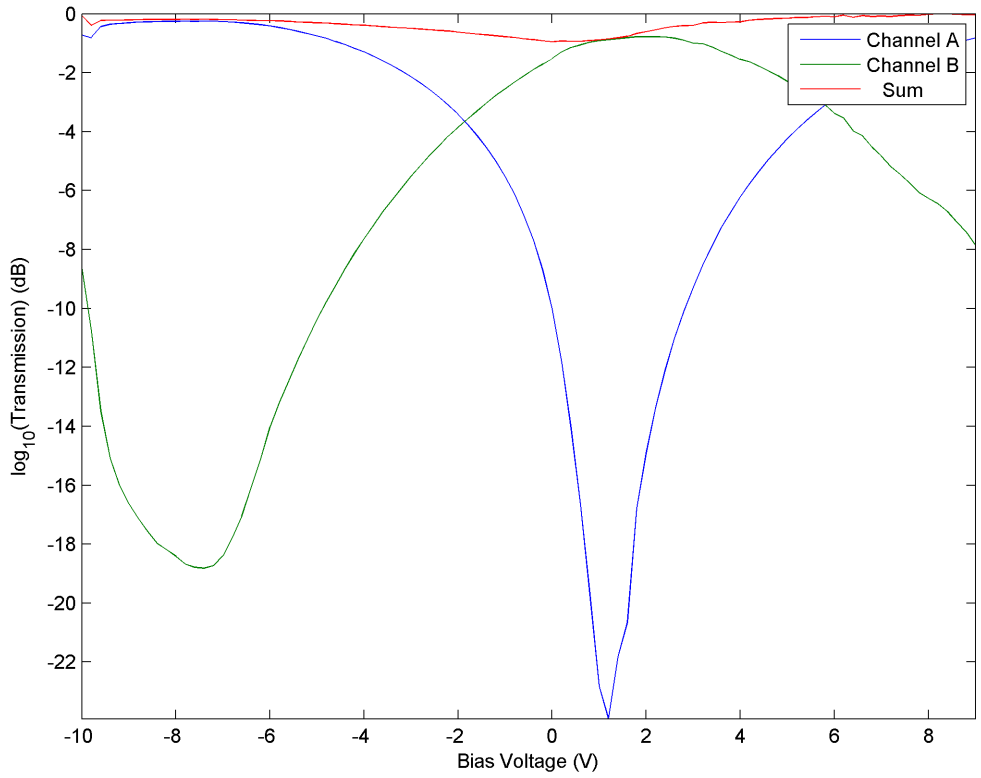


Figure 4.3: The output of the 1000-μm PP MZM, showing the complementary nature of the top and bottom output arms. The output transmission is normalized, and given on a logarithmic scale at λ=1550 nm.

with the optical mode profile (see Section 5.1), there would only be a marginal change in the effective index of the waveguide with the application of a larger voltage (the ridge waveguide is only 240 nm high). With turning points located at −7.6 V and 1.2 V, V_π = 8.8 V and $V_\pi L$ = 0.88 V cm. Because the device design is identical, with the exception of the length of the active region, the value of V_π should be roughly twice that of the 2000-μm PP MZM, and the value of $V_\pi L$ should be nearly equal. That $V_\pi L$ is not entirely equal may be attributed to minor fabrication imperfections, despite the fact that both devices were fabricated on the same chip — fabrication variances such as junction misalignment

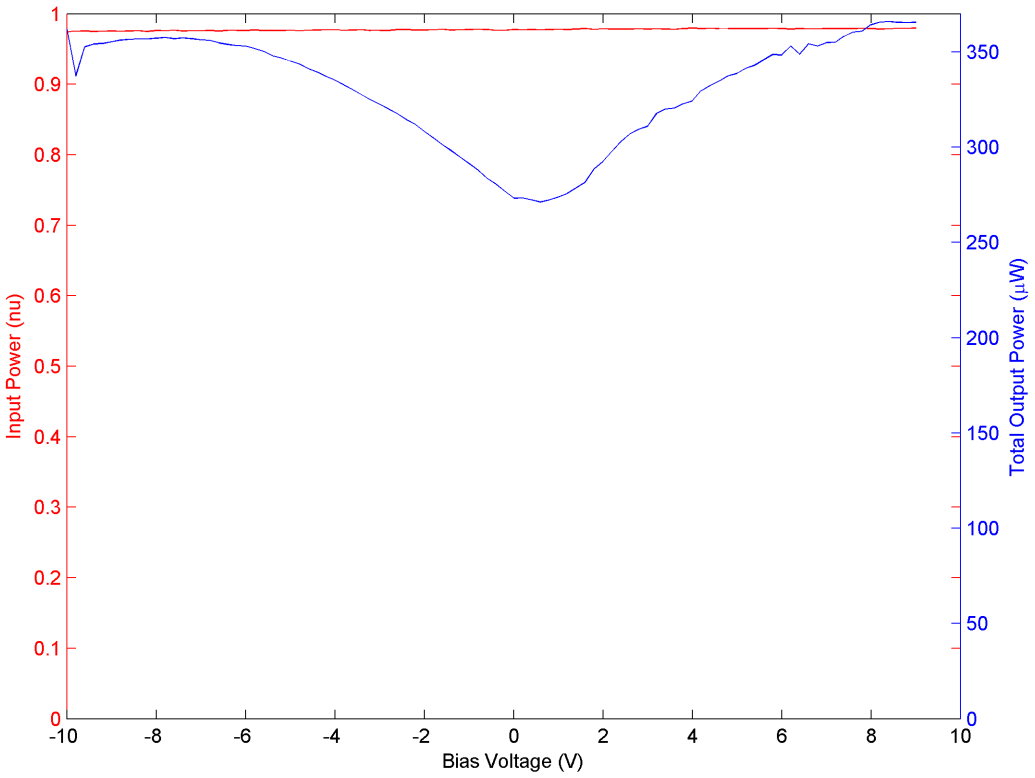


Figure 4.4: A ratio of the input powers to the 1000-μm PP MZM, which shows that the input powers across experiments was roughly equal. The total output power between the two output arms is also shown in terms of real power rather than normalized units.

and differences in dopant density could be the cause. To maximize ER, the 1000-μm PP MZM should be biased at −3.2 V, and driven with 8.8-V_{pp}, to achieve an ER of 24.24 dB on Channel A, and an ER of 18.05 dB on Channel B.

It is also possible to match the ER of the 2000-μm PP MZM with the 1000-μm PP MZM by altering the bias and driving voltages, thereby realizing power savings. By biasing the MZM at 0.3 V and driving with a 1.8-V_{pp} signal, Channel A can achieve an ER of 17.82 dB, and by biasing the MZM at −1.5 V and applying a 6.2-V_{pp} driving signal, an ER of 8.64 dB on Channel B is possible. It is, however, important that the MZM be kept

out of injection mode during high-speed operation. To do otherwise would decrease the maximum achievable bit rate, and the ER at high-speeds as a result of free-carrier lifetimes. Therefore, the above bias and driving voltages should be adjusted appropriately to ensure that the p–n junction is never forward-biased by a value larger than its built–in potential. Altering the bias and driving voltages to keep from exceeding the built–in potential trades ER at all speeds for a higher achievable bit rate.

As mentioned before, the DC bias is applied to the center pin of the TWE MZM (via the GSG pads) and the microwave signal is applied to the top and bottom electrodes (via the GS/SG pads). Although the active regions of both the PP MZMs and the TWE MZM are identical, as implemented during the course of this thesis work, the TWE MZM uses a single-arm modulation scheme. This is due to the fact that there is always a DC bias applied to the central electrode of the TWE MZM (the center pin in Figure 2.14), whereas the PP MZMs leave the central electrode floating. Because the central electrode is not floating during high-speed experimentation on the TWE MZM, the electric field between the ground electrode and the central electrode remains constant, whereas the electric field between the signal electrode and the central electrode varies. Therefore, only one arm of the MZM is actually modulated, whereas the other arm remains at some constant phase shift — the exact value of the constant phase shift being a function of the applied bias voltage. Certainly if the central electrode was left floating, then the TWE MZM would also have a single-drive push-pull modulation scheme. During DC characterization experiments though, the traveling-wave signal electrode was left floating, and the ground and bias electrodes were connected via the GSG pads. In this manner, there was no electric field across one arm of the MZM, and the electric field across the other arm of the MZM varied with the applied DC bias. From this perspective, single-arm modulation was achieved, and as such the DC characterization has direct applicability to the high-speed performance of

the device. Additionally, there is equivalence between the DC characterization of the TWE MZM and the DC characterization of the PP MZMs.

Figure 4.5 provides transmission data between $-10\,\mathrm{V}$ to $10\,\mathrm{V}$ at $1550\,\mathrm{nm}$. The data speaks to the complementary nature of the outputs, but this should not be surprising, as the TWE MZM uses the same 2x2 adiabatic coupler design as the PP MZMs do, and the adiabatic couplers are the cause of the complementary behavior. Looking at Figure 4.5, far better ERs are achieved using Channel A rather than Channel B. Operating on Channel A corresponds to using the cross output port. Note that due to the device design and doping profile, the application of a positive bias voltage to the center pin places the device in depletion mode.

Figure 4.6 shows that the input power was roughly identical for both experiments, and that the total output power changes significantly with voltage, from nearly $65\,\mu\mathrm{W}$ at its peak to almost $10\,\mu\mathrm{W}$ at its minimum, which is a level of destructive interference unseen in either the 1000-μm or 2000-μm PP MZMs. In the case of the PP MZMs, the percent change in total output power was not nearly as large, as can be verified by looking at Figures 4.2 and 4.4. Physically, the explanation for this is difficult to imagine, as the active regions and couplers are the same in both the PP and TWE designs, and though the modulation scheme changes from push-pull to single-arm, the interference between the two arms should be just as complete at a π phase shift. The lack of destructive interference in the PP design could be attributed to a velocity and impedance mismatch, but in the DC case, neither of these should be factors.

Importantly, the above results were given at a single wavelength, namely $1550\,\mathrm{nm}$. It is also possible to look at the effect that wavelength has on the transmission of the MZM. To do so adds a third dimension to the dataset, and it is therefore necessary to make use of surface and contour plots. The contour and surface plots each present the same data, the only difference being that the contour plot is a 2D plot of the two factors (bias voltage and

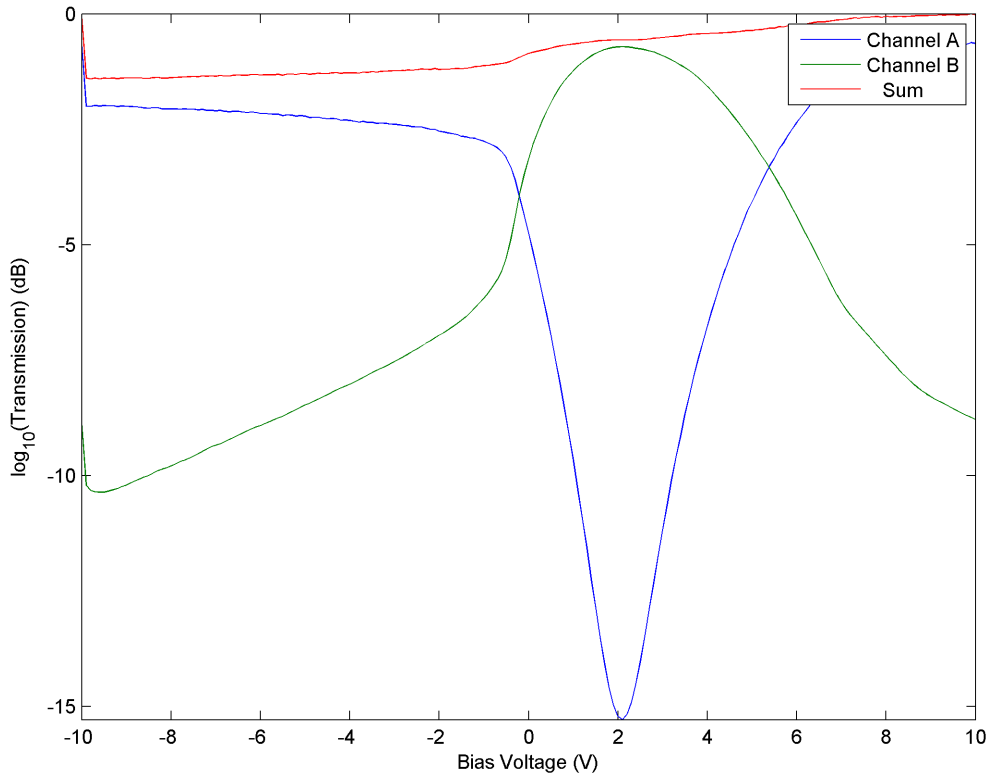


Figure 4.5: The output of the 1500-μm TWE MZM, showing the complementary nature of the top and bottom output arms. The output transmission is normalized, and given on a logarithmic scale at λ=1550 nm.

wavelength) with the response (transmission) plotted as color, and the surface plot is a 3D plot of the three variables, with transmission also designated by color. Surface plots with a logarithmic scale are presented in this chapter, while surface plots with a linear scale and contour plots using both linear and logarithmic scales are given in the Appendix, Section A. The reader is encouraged to refer to these figures, as a different presentation of the data may be beneficial.

Figure 4.7 consists of surface plots presenting the transmission of the 1000-μm PP MZM as a function of both wavelength and bias voltage, for both the straight and cross

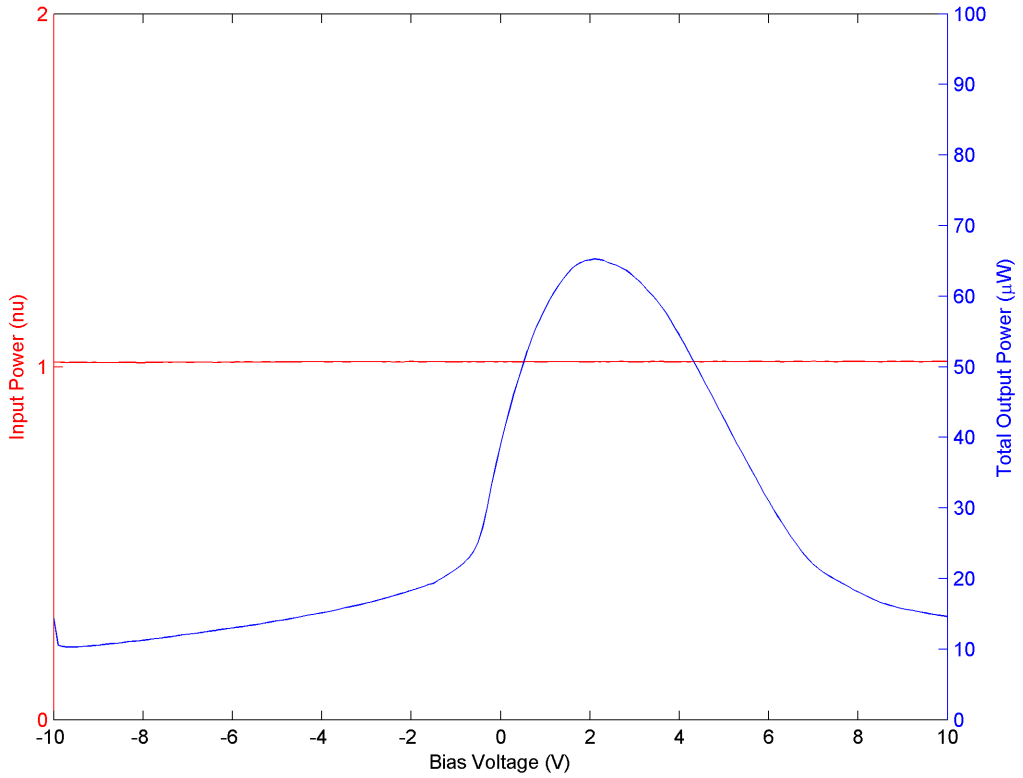


Figure 4.6: A ratio of the input powers to the 1500-μm TWE MZM, which shows that the input powers across experiments was roughly equal. The total output power between the two output arms is also shown in terms of real power rather than normalized units.

ports. The data from Channel A shown in Figure 4.3 corresponds to a cross-section of the data from Figure 4.7a, the cross-section being across a single wavelength, namely 1550 nm. The same idea holds true for Channel B shown in Figure 4.3, as it is a cross-section of Figure 4.7b.

The minimal variance in the voltage at which the minimum and maximum transmission occurs translates into a minimal change in the ideal bias voltage and driving voltage across all wavelengths. Practically, this means the drive circuitry for the modulator does not have to be designed for a specific wavelength, leading to larger flexibility during

implementation, and the laser source does not need to be tuned to a specific operating wavelength to ensure optimal performance of the MZM.

As can be seen from the surface and contour plots, the behavior is roughly similar across all wavelengths with the exception of a minor ripple that appears to be a function of wavelength. This ripple will be discussed in Section 4.5; however for now, the reader should ignore the ripple's presence. Such a characteristic — the device's uniform behavior across a 43 nm range in wavelength — serves as a testament to the device's broadband optical bandwidth. This is a feature that is important for real-world applications, and validates the use of adiabatic couplers rather than directional or MMI couplers. Additionally, the length symmetry between the arms of the MZM means that there is no inherent Free Spectral Range (FSR) to the device, and all told, the device has a much broader optical range than would a resonant-cavity device.

Considering the straight port, from a macroscopic perspective, the maximum transmission decreases linearly with wavelength, from 1540 nm to 1583 nm, with the maximum transmission decreasing by 40 % over the same range. Over that range of wavelengths, the minimum transmission occurs at smaller bias voltages, with a mean value of $\mu = 1.14\,\mathrm{V}$ and a standard deviation of $\sigma = 0.1095\,\mathrm{V}$. This then results in the ideal DC bias point decreasing with an increase in operating wavelength. The maximum transmission decreases with wavelength, which contradicts expectations based solely upon the fact that FCA is anticipated to decrease with an increase in wavelength, or rather as the photons become less energetic. It may be that scattering dominates the process, as opposed to FCA, as scattering was expected to increase with an increase in wavelength due to the wavelength dependence of the confinement factor. There is also a wavelength dependence to the optical power meters, and although AutoLab was configured to set the wavelength being measured by the optical power meter (which presumably compensated for changes in

responsivity), more experimentation should be done to rule this out. As of now, the cause of the decrease in maximum transmission is inconclusive.

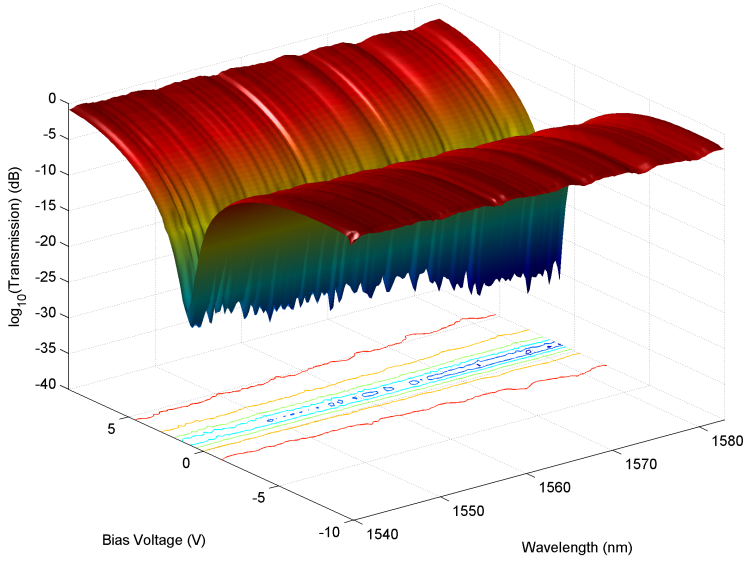
The decrease in maximum transmission also accompanies an unexpected increase in ER with an increase in wavelength. The maximum and minimum transmission and ER at each wavelength can be seen in Figure 4.8a. The mean ER is 24.96 dB, with a standard deviation of 1.90 dB. The minimum ER of 21.00 dB occurs at 1543.25 nm, and the maximum ER of 29.69 dB occurs at 1581.75 nm. Noting that the ER is greater than 20 dB for all wavelengths is once more a testament to the device's broad optical bandwidth, and the general device design. The achievement of such ERs also indicates that there are minimal concerns to be had with regard to power imbalance in the two arms of the MZM. The data is summarized in Table 4.1.

While the above discussion centered on the output from the straight port, the output from the cross port displays similar trends, with the obvious exception being that transmission is complementary. Figure 4.7b shows the transmission response data from these experiments. The ER as a function of wavelength is given in Figure 4.8b. The behavior is still similar across wavelengths, as the maximum transmission decreases with an increase in wavelength, and V_π has minimal variance. Data concerning minimum and maximum transmission as well as ER is presented in Table 4.1. Once again, note that the minima and maxima are very much aligned at the same voltage between the two output arms, so that the outputs are indeed complementary in nature. However, as pointed out in the previous discussion, the ER is lower when taking the output from the cross port than it would be if taking the output from the straight port. The need for normal or inverted logic, and bit rate requirements would dictate which arm should be used during normal operation. When stressing the modulator by operating at higher data rates, it would be beneficial to use the straight port, as it has the higher ER, and would therefore result in fewer bit errors at a given bit rate.

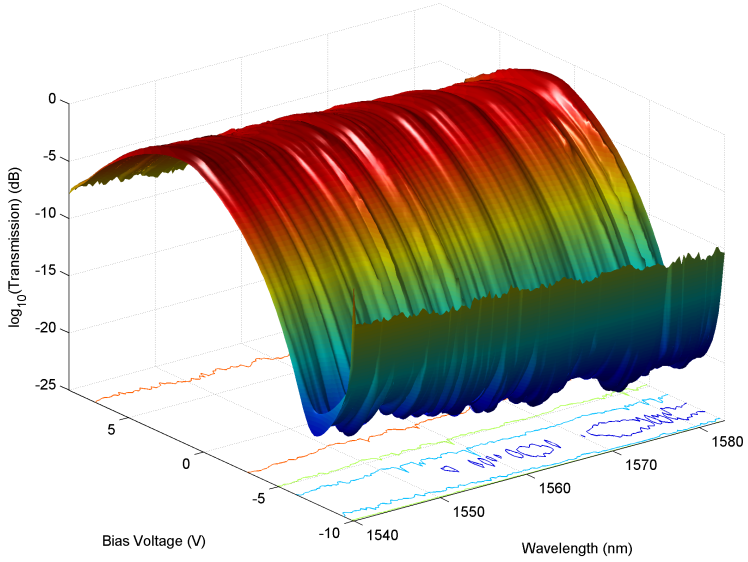
Table 4.1: Summary of the DC characterization data for the 1000-μm PP MZM.

Metric	Straight Port	Cross Port
μ, Minimum Transmission Voltage	1.14 V	−7.79 V
σ, Minimum Transmission Voltage	0.11 V	0.30 V
μ, Maximum Transmission Voltage	−7.83 V	1.52 V
σ, Maximum Transmission Voltage	0.21 V	0.19 V
μ, V_π	8.96 V	9.31 V
σ,V_π	0.20 V	0.37 V
μ, ER	24.96 dB	18.57 dB
σ, ER	1.90 dB	0.72 dB
Minimum ER	21.00 dB at 1543.25 nm	16.71 dB at 1545.75 nm
Maximum ER	29.69 dB at 1581.75 nm	20.10 dB at 1565 nm

Returning to the 2000-μm PP MZM, the results of the same DC characterization experiments are presented as they were for the 1000-μm PP MZM above. Figure 4.9 presents the transmission for the straight and the cross port. The ER as a function of wavelength is shown in Figure 4.10. Summary statistics for the straight and cross ports are contained in Table 4.2. Notably, the V_π figure was less than half that of the 1000-μm PP MZM, as would be expected since the active region doubled in length. The ER of the 2000-μm PP MZM was lower though — it may be that the ER decreases due to an increase in power imbalance between the arms of the MZM due to waveguide loss [59]. Although Table 4.2 states that the maximum transmission for the straight port occurs at 0.24 V, the maximum transmission actually occurs at −10 V; however operating at such large voltages is both impractical and unnecessary, considering that −6 V is beyond the first turning point. As a result, data <−6 V was ignored during the analysis. Note that the ideal DC bias point decreases as the operating wavelength increases, as was the case with the 1000-μm PP

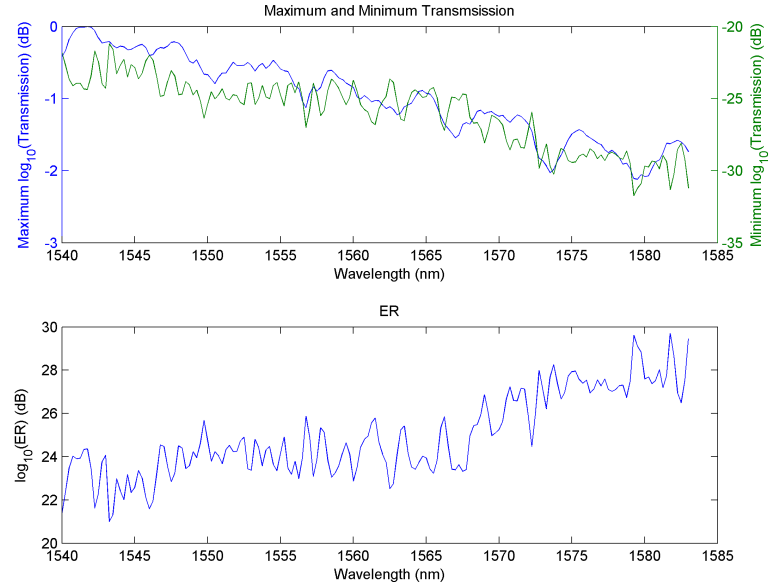


(a) Straight Port

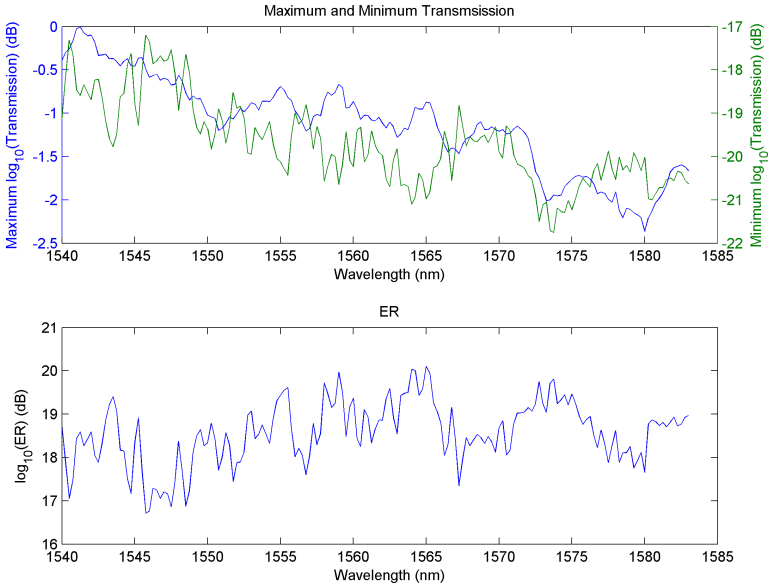


(b) Cross Port

Figure 4.7: The transmission of the 1000-µm PP MZM as a function of wavelength and bias, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.



(a) Straight Port



(b) Cross Port

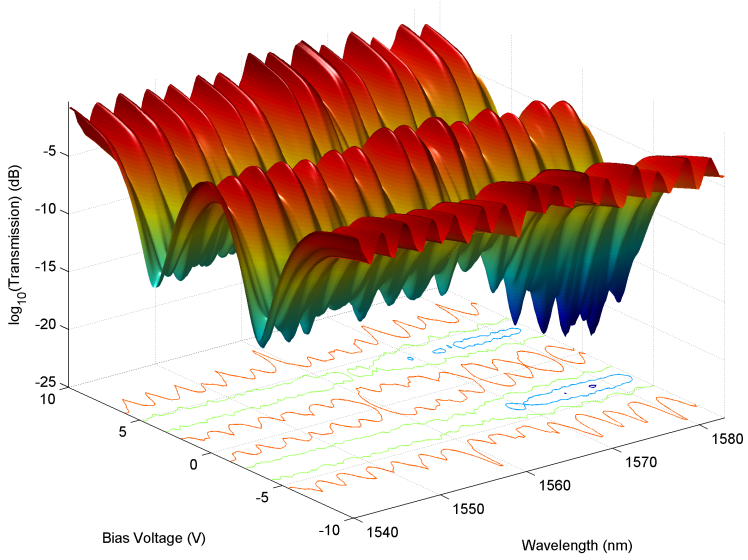
Figure 4.8: The maximum transmission, minimum transmission, and maximum ER of the 1000-μm PP MZM as a function of wavelength, presented on a logarithmic scale.

Table 4.2: Summary of the DC characterization data for the 2000-μm PP MZM.

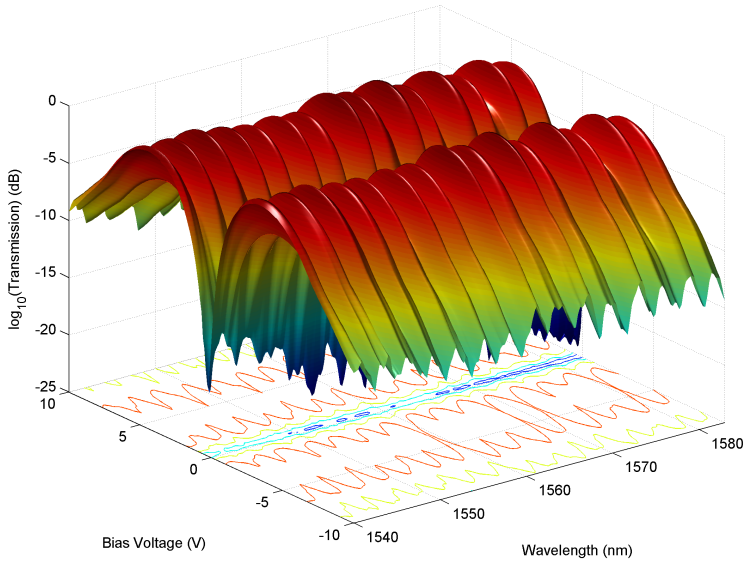
Metric	Straight Port	Cross Port
μ, Minimum Transmission Voltage	−3.43 V	0.23 V
σ, Minimum Transmission Voltage	0.09 V	0.05 V
μ, Maximum Transmission Voltage	0.34 V	−3.60 V
σ, Maximum Transmission Voltage	0.06 V	0.10 V
μ, V_π	3.77 V	3.83 V
σ,V_π	0.08 V	0.08 V
μ, ER	10.90 dB	17.26 dB
σ, ER	2.83 dB	2.73 dB
Minimum ER	5.99 dB at 1559.25 nm	10.79 dB at 1542 nm
Maximum ER	18.02 dB at 1578.25 nm	22.08 dB at 1582 nm

MZM. The standard deviation of the ER was also significantly impacted by the ripple, as the depth of the ripple in this data set was significantly larger than it was for the 1000-μm PP MZM. Unlike the 1000-μm PP MZM, the 2000-μm PP MZM has a higher ER on the cross port than on the straight port; however, the port with the central null is the port that has the higher ER in both cases. This behavior will be explained later. Once again, note the similar behavior across all wavelengths, meaning that this device too has a broad optical bandwidth, as did the 1000-μm PP MZM.

For the 1500-μm TWE MZM, a surface plot is once more given for transmission as a function of both wavelength and bias voltage, with Figure 4.11 covering the behavior of the straight port and cross ports. ER as a function of wavelength is given in Figure 4.12. The presence of the ripple makes it unclear as to whether or not the maximum transmission and wavelength are inversely proportional, as they were with the PP MZMs; however,

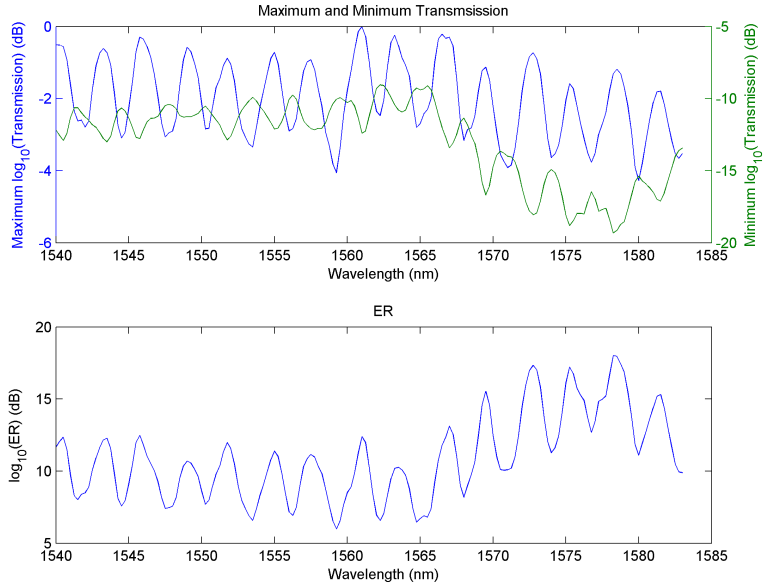


(a) Straight Port

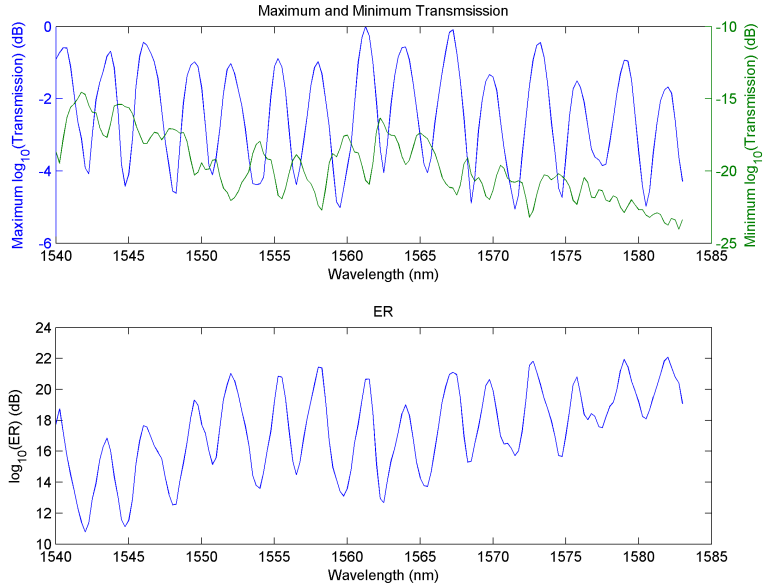


(b) Cross Port

Figure 4.9: The transmission of the 2000-µm PP MZM as a function of wavelength and bias, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.



(a) Straight Port



(b) Cross Port

Figure 4.10: The maximum transmission, minimum transmission, and maximum ER of the 2000-μm PP MZM as a function of wavelength, presented on a logarithmic scale.

it is assumed that would be true, as the behavior was caused by an increase in optical propagation losses. It is noted that the location of the turning point decreases with respect to voltage with an increase in wavelength, as was the case with the PP MZMs. Summary statistics are given in Table 4.3 for the straight and cross ports. The entire voltage range, from −10 V to 10 V, was not considered for Channel A during analysis when determining the value for V_π and its associated ER. The maximum transmission voltage occurs at 10 V, and there is less than a 1 dB difference in ER between 7 V to 10 V, therefore, data was only considered from −10 V to 7 V. Rather than considering the entire voltage range for Channel B, voltages less than −1 V were not considered. Without discarding data below −1 V, the minimum transmission voltage would have been found at roughly −9.5 V, thereby skewing the value for V_π, despite the fact that operating between the maximum at 2.1 V and the minimum at −9.5 V would operate the device in injection mode. Using this process, an ER of 13.95 dB is possible on Channel A, with a V_π of 4.9 V, and an ER of 8.08 dB on Channel B, with a V_π of 7.9 V. The summary statistics make clear the fact that the ERs between straight and cross ports are radically different, as the cross port output has a mean ER of 18.97 dB as compared to the straight port's mean ER of only 8.87 dB. Interestingly, the minimum ER for the cross port, 10.52 dB at 1548.00 nm, is actually larger than the maximum ER for the straight port, 10.33 dB at 1581.5 nm. Moreover, the larger ER of the cross port is achieved with a lower driving voltage, as V_π for the cross port is only 5.17 V on average, compared to the straight port's 7.94 V. Once again, the depth of the ripple affects the standard deviation of ER as can be seen from the summary statistics.

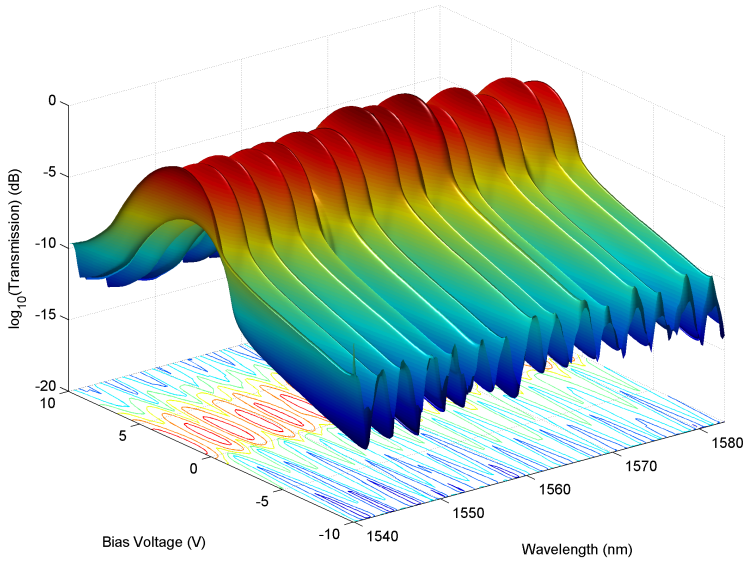
Additionally, the output port with a central null has the largest ER, which corresponds to the most complete destructive interference. It is unclear why that level of ER is only present in one output arm and not the other, as the interference occurs in the interaction region of the couplers, before power is split to either port. It may be that the interference is more complete in either the even or odd waveguide mode, and because that mode is always

Table 4.3: Summary of the DC characterization data for the 1500-μm TWE MZM.

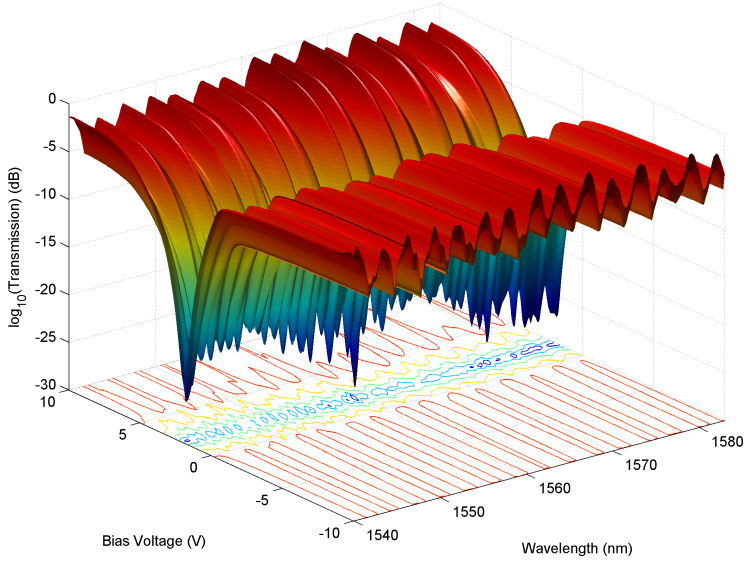
Metric	Straight Port	Cross Port
μ, Minimum Transmission Voltage	10 V	1.83 V
σ, Minimum Transmission Voltage	0.00 V	0.19 V
μ, Maximum Transmission Voltage	2.06 V	7 V
σ, Maximum Transmission Voltage	0.12 V	0 V
μ, V_π	7.94 V	5.17 V
σ,V_π	0.12 V	0.18 V
μ, ER	8.87 dB	18.97 dB
σ, ER	0.57 dB	2.93 dB
Minimum ER	7.24 dB at 1541.75 nm	10.52 dB at 1548.00 nm
Maximum ER	10.33 dB at 1581.5 nm	25.61 dB at 1579.25 nm

sent to only one output port, the maximum ER is different between the two output arms. There was no evidence to suggest that the active regions perform more efficiently for a particular mode, but it may be that a difference in confinement factor for the two modes creates a power imbalance, leading to less complete destructive interference.

Returning to the DC characterization of the PP MZMs, the mean value for V_π of the 1500-μm TWE MZM (mean of the mean values for V_π of the straight and cross ports) is 6.56 V, and is between the value for V_π of the 1000-μm or 2000-μm PP MZMs, with the 1000-μm at 9.14 V, and the 2000-μm at 3.80 V. It is not entirely surprising that V_π for the 1500-μm TWE MZM is greater than the 2000-μm PP MZM, as the 2000-μm has a longer active region (which reduces the necessary voltage to achieve a π phase shift), and the series push-pull modulation scheme is noted to have only a marginal effect on the half-wave voltage. By the same token, it is not astonishing that V_π for the 1000-μm PP MZM

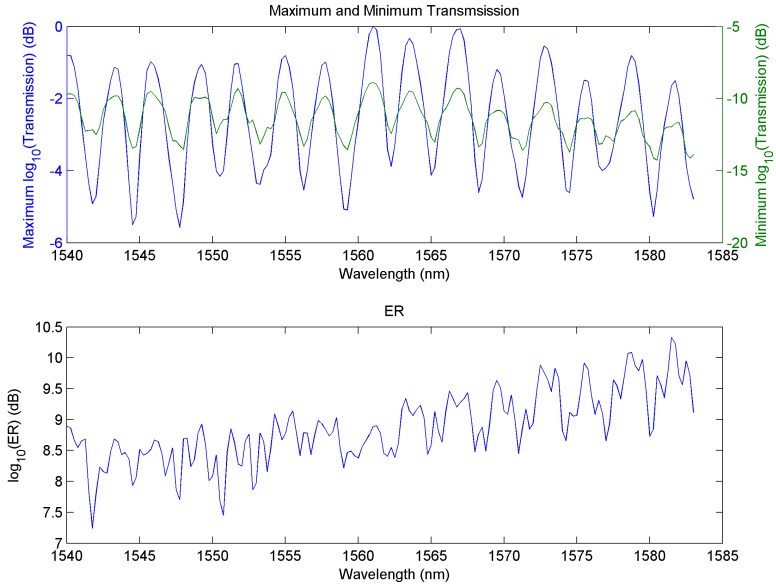


(a) Straight Port

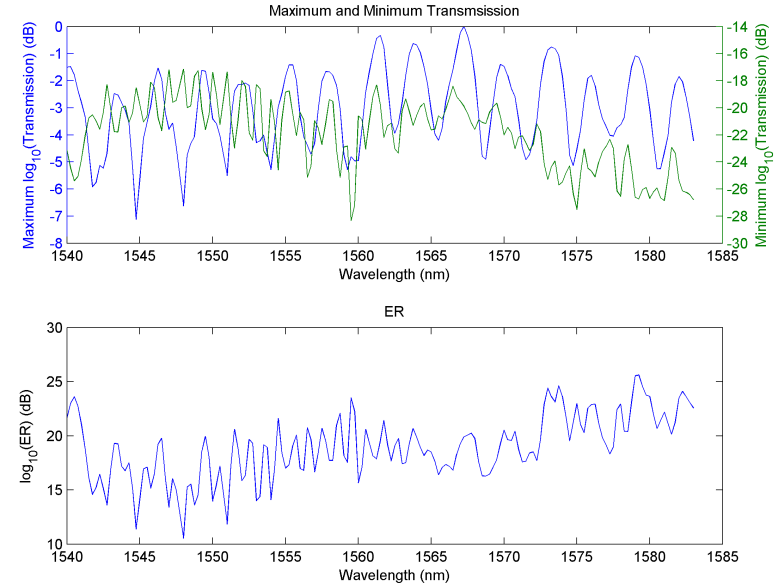


(b) Cross Port

Figure 4.11: The transmission of the 1500-μm TWE MZM as a function of wavelength and bias, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.



(a) Straight Port



(b) Cross Port

Figure 4.12: The maximum transmission, minimum transmission, and maximum ER of the 1500-μm TWE MZM as a function of wavelength, presented on a logarithmic scale.

is greater than the 1500-μm TWE MZM, as the active region is shorter, and the series push-pull scheme serves primarily to increase modulation bandwidth.

IL is defined as the maximal element, m, of the set S, given by

$$S = \left\{ 10 * \log_{10} \left[\frac{P_{\text{out, straight}} + P_{\text{out, cross}}}{\left(P_{\text{in, straight}} + P_{\text{in, cross}} \right) / 2} \right] \right\} \tag{4.1}$$

such that $m \in S$ and $m \geq s$ for all $s \in S$, where s is given in dB. Using this definition, the IL of the 1000-μm PP MZM is -8.78 dB, and the IL of the 2000-μm PP MZM is -15.56 dB, leading to the conclusion that the optical loss in the active region is significant compared with the optical loss in the remainder of the ridge waveguide. The IL of the TWE MZM was calculated to be -13.93 dB, which lies between the IL of the 1000 μm and 2000 μm PP MZMs, and further supports the conclusion that the active region dominates optical propagation losses. Of this total loss, 0.4 dB can be attributed to the adiabatic couplers, 6 dB to the coupling at the facets, and 0.01 dB mm^{-1} to the silicon wire waveguides, and the remainder to the active region [60] [15] [34].

4.2 Bandwidth Characterization

Bandwidth characterization of the PP MZMs was performed by measuring the S_{11} and S_{21} responses of the device. The bandwidth of the device is ultimately dictated by the S_{21} response of the device, however as described earlier, reflections off the transmission line will occur at the probe, if the device is not properly impedance matched, thereby leading to poor power efficiency. Reflections will also occur off the end of the transmission line if the transmission line is left unterminated. Such reflections will negatively impact the bandwidth of the device, as described in Section 2.1.5. If the S_{11} response of the device is largely attenuated across the range of frequencies of interest (large return loss/low reflection), then the device is properly impedance matched and terminated in said frequency range. The S_{21} response of the device is affected by a number of parameters mentioned earlier – the junction resistance, junction capacitance, transmission line velocity matching, and transmission line impedance matching. Devices are typically benchmarked by their S_{21} 3 dB and 6 dB performances. A reference line is drawn at -3 dB on the S-parameter figures, and the data is normalized to the lowest frequency measured, 500 MHz. Because S-parameter measurements are small signal measurements, and test the responsivity of the active region and the electrodes, the combination of input and output arms was not thought to be particularly important. Still though, data was collected for the straight and cross ports in S_{21} tests. S_{11} measurements are purely electrical, and so the combination of optical input and optical output arms is irrelevant. The data for the S_{21} plots was taken by averaging the results of five S-parameter measurements and then plotted using a moving average to further smooth the data, while the S_{11} tests simply used the average of five measurements.

The S_{11} responses, at multiple bias points for the 1000-μm PP MZM are given in Figure 4.13, for the 2000-μm PP MZM in Figure 4.14, and for the 1500-μm TWE MZM in Figure 4.15. The data is summarized in Table 4.4. As can be seen for the PP MZMs, there is almost no return loss, whereas for the TWE MZM, there is significant return loss across

the entire frequency range. The insignificant return loss at low frequencies for the PP MZM is simply due to the unterminated design of the electrodes, and the increase in return loss with frequency is simply due to impedance mismatch. The lack of return loss in the PP cases will negatively affect bandwidth, as there is very little return loss at low frequencies, where most of the power is in PRBS patterns. By contrast, the TWE MZM has a minimum return loss of 10.44 dB at a bias of 4 V. Because the capacitance of the p–n junction in the active region is a loaded capacitance and affects the device impedance, it is a given that the bias voltage affects the S_{11} measurements. Forward biasing the TWE MZM results in an unfavorable modification to capacitance, and thus the least return loss.

Table 4.4: Summary of the S_{11} response data for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs. The bias voltage in the table is the bias voltage which results in the absolute largest minimum return loss.

Device	Bias (V)	Min Return Loss (dB)
1000 PP	0	0.32
2000 PP	-6	0.54
1500 TWE	4	10.44

The S_{21} response, when biased at multiple operating points for the 1000-μm PP MZM, is given in Figures 4.16 and 4.17, for the 2000-μm PP MZM in Figures 4.18 and 4.19, and for the 1500-μm TWE MZM in Figures 4.20 and 4.21. The $f_{3\,\mathrm{dB}}$ and $f_{6\,\mathrm{dB}}$ bandwidths of these devices is summarized and reported in Tables 4.5 and 4.6.

As is shown in the given figures, the bias voltage does affect the bandwidth of the device. The device bandwidth increases as the bias voltage decreases, or in other words, the device is farther into depletion mode. This fact is not surprising, as the drift velocity is

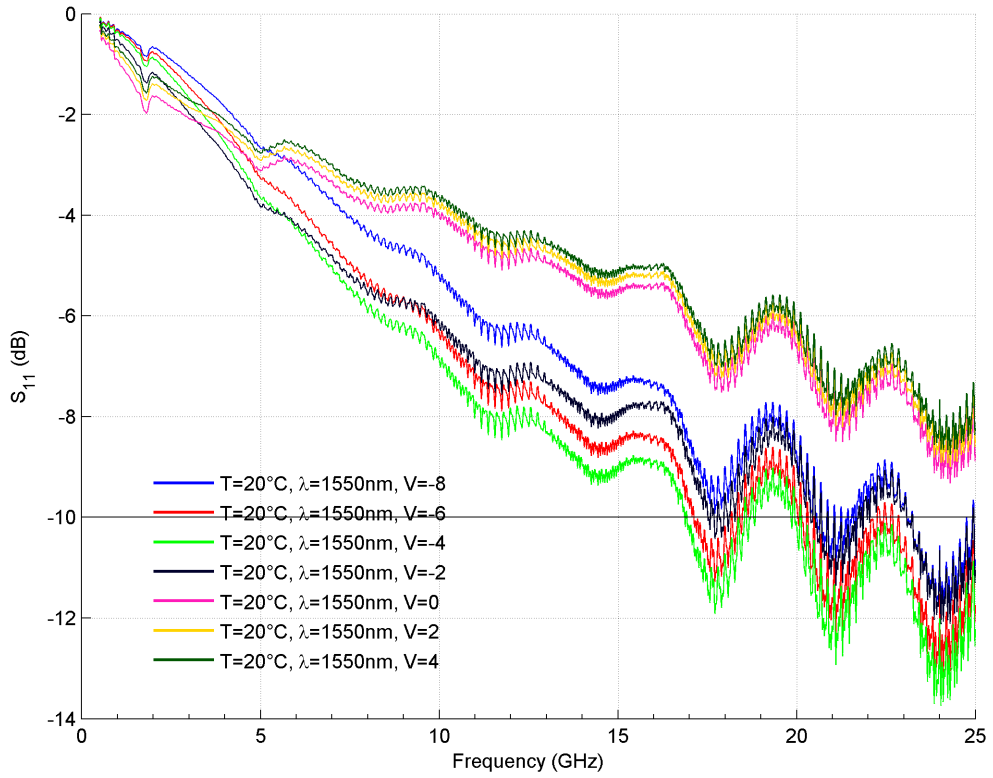


Figure 4.13: S_{11} for the 1000-μm PP MZM at a variety of bias voltages. The best device performance is achieved using a 0 V bias, and results in a minimum return loss of 0.32 dB.

proportional to the electric field strength, which increases proportionally with the voltage applied across the junction — *i.e.* with the bias voltage — and so the faster carrier transport will increase device responsivity. In the case of a lumped-element model, this behavior would be explained by looking at the equation for p–n junction capacitance, as reverse biases decrease the junction capacitance, and thus decrease the *RC* constant time. As can be verified by looking at Figures 4.16, 4.17, 4.18, and 4.19, S_{21} responses for the PP MZMs are nearly identical whether positively or negatively biased (e.g. S_{21} response for 2 V bias = S_{21} response for −2 V bias). This is simply because the diodes (p–n junctions

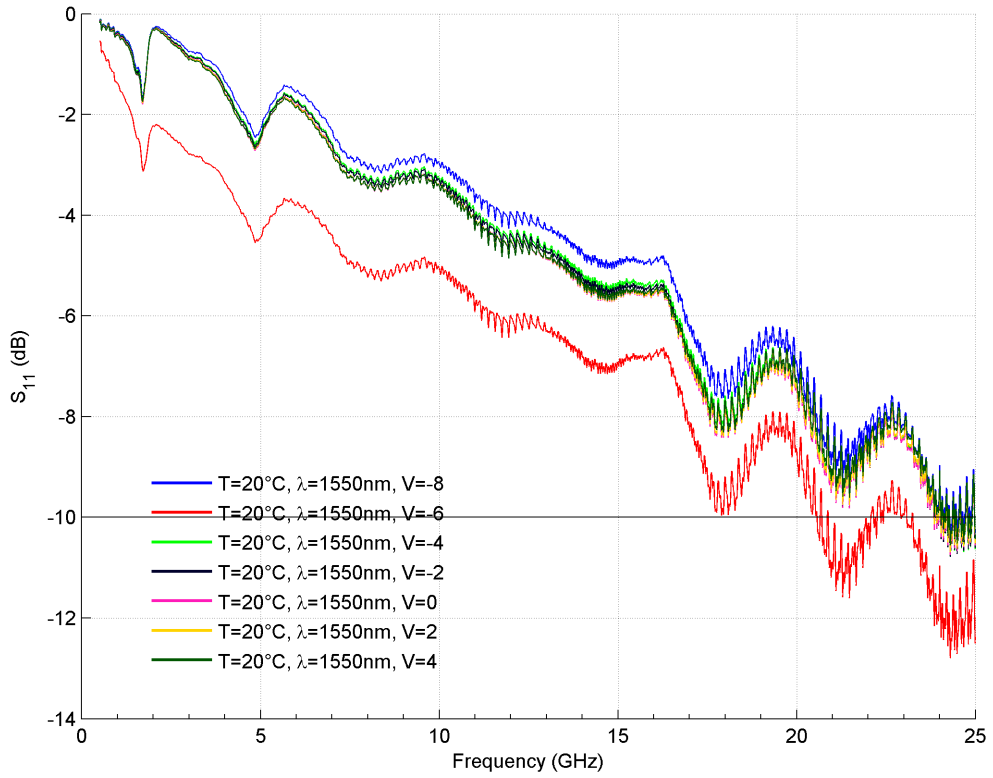


Figure 4.14: S_{11} for the 2000-μm PP MZM at a variety of bias voltages. The best device performance is achieved using a −6 V bias, and results in a minimum return loss of 0.54 dB.

in the active region) are facing in opposite directions, so positively or negatively biasing the junction always results in one junction being forward-biased, and the other reverse-biased. Additionally, it is important to note that S-parameter measurements only suggest an optimal bias voltage, and perhaps a bias voltage with a large bandwidth has a poor ER when using a large driving signal. For example, the S_{21} test would suggest using a bias voltage of −3 V for the straight port, and −5.5 V for the cross port of the 1000-μm PP MZM. However, the BER tests, which used a large driving signal, actually used voltages of −2.5 V and −4.3 V for the straight and cross ports, respectively. Likewise, the 2000-μm PP

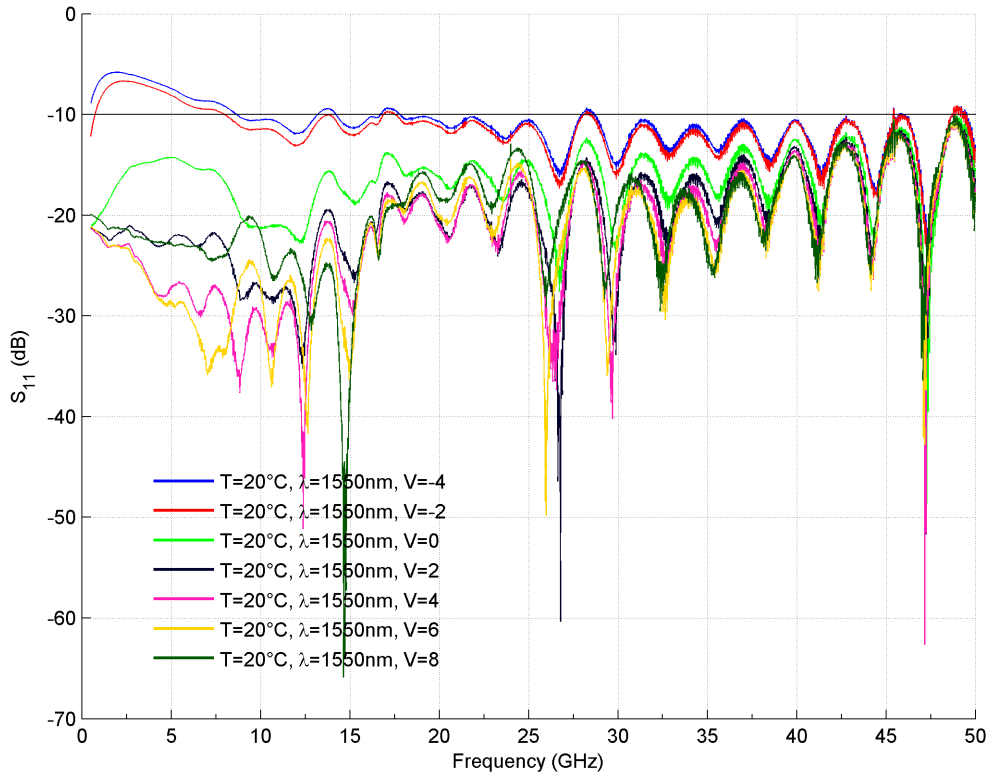


Figure 4.15: S_{11} for the 1500-µm TWE MZM at a variety of bias voltages. The best device performance is achieved using a 4 V bias, and results in a minimum return loss of 10.44 dB.

MZM used voltages of −2.6 V and −2.45 V for the straight and cross ports, respectively, during BER tests, despite the suggested voltages of −3.5 V and −3 V from the S_{21} tests. The 1500-µm TWE MZM used a bias voltage of 4 V rather than 8 V as suggested by the S_{21} tests, though the logic behind the decision to operate at 4 V is different.

One would think that because the 2000-µm PP MZM has twice the capacitance of the 1000-µm PP MZM, it would have half the bandwidth, but looking at Table 4.5 suggests otherwise. This would indeed be the case if only looking at the 0 V bias case, for which the 1000-µm PP MZM has a 5.30 GHz/5.18 GHz $f_{3\,\mathrm{dB}}$ cutoff frequency (straight/cross ports),

Table 4.5: Summary of the S_{21} response data for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs. The bias voltage in the table is the bias voltage which results in the largest 3 dB and 6 dB bandwidths.

Device (μm)	Port	Bias (V)	$f_{3\,\mathrm{dB}}$ (GHz)	$f_{6\,\mathrm{dB}}$ (GHz)
1000 PP	Straight	-3	9.27	19.43
1000 PP	Cross	-5.5	8.73	18.82
2000 PP	Straight	-3.5	3.94	9.62
2000 PP	Cross	-3	3.54	7.97
1500 TWE	Straight	8	24.44	44.35
1500 TWE	Cross	8	24.30	47.21

Table 4.6: Summary of the S_{21} response data for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs. The ranges that are given are defined as the difference between the minimum and maximum values for the respective 3 dB and 6 dB bandwidths, based on the bias voltages which produce the minimum and maximum bandwidths.

Device (μm)	Port	$f_{3\,\mathrm{dB}}$ Range (GHz)	$f_{6\,\mathrm{dB}}$ Range (GHz)
1000 PP	Straight	3.97	7.81
1000 PP	Cross	3.56	7.46
2000 PP	Straight	1.47	4.17
2000 PP	Cross	1.15	2.63
1500 TWE	Straight	22.25	39.20
1500 TWE	Cross	22.27	42.14

and the 2000-μm PP MZM has a 2.48 GHz/2.39 GHz $f_{3\,\mathrm{dB}}$ cutoff frequency (straight/cross ports). The applied bias should modify the capacitance identically for both the 1000-μm and 2000-μm PP MZMs though, as the junction capacitance is based on the junction doping

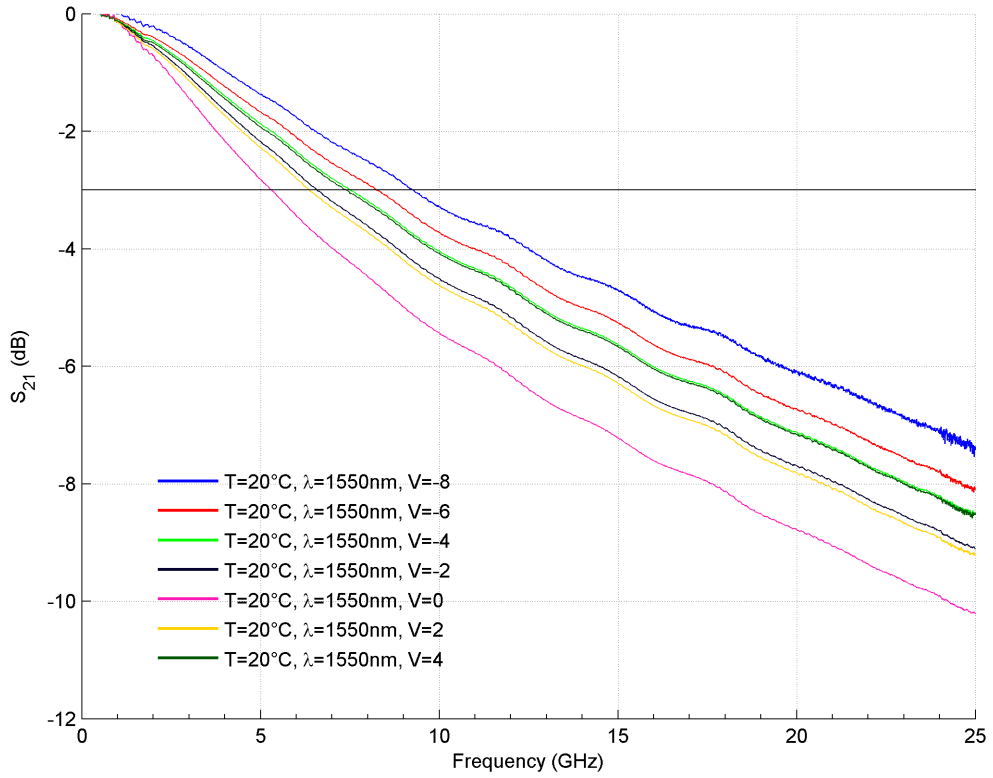


Figure 4.16: S_{21} for the 1000-µm PP MZM at a variety of bias voltages and 1550 nm, using the straight port. The best device performance is achieved using a -3 V bias, and results in $f_{3\,\mathrm{dB}} = 9.27$ GHz and $f_{6\,\mathrm{dB}} = 19.43$ GHz.

profile. That the junction capacitances do not change equally is a fact that is not easily explained.

While the PP MZMs S_{21} plots (Figures 4.16, 4.17, 4.18, and 4.19) do not show the severe effects of free-carrier lifetime limitations on bandwidth, the TWE MZM plots (Figures 4.20 and 4.21), do show large bandwidth differences for the injection and depletion

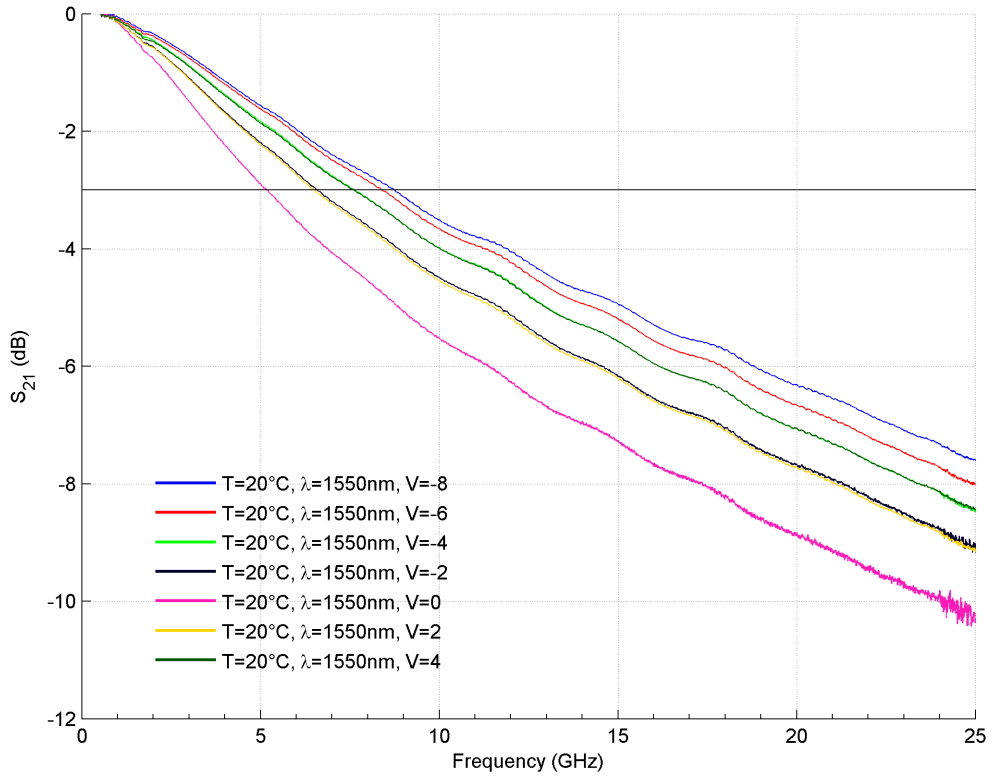


Figure 4.17: S_{21} for the 1000-μm PP MZM at a variety of bias voltages and 1550 nm, using the cross port. The best device performance is achieved using a −5.5 V bias, and results in $f_{3\,\mathrm{dB}} = 8.73\,\mathrm{GHz}$ and $f_{6\,\mathrm{dB}} = 18.82\,\mathrm{GHz}$.

mode cases. This is because, as mentioned previously, the PP MZM always has one diode reverse-biased and the other forward-biased, regardless of bias voltage, while the TWE MZM configuration has both diodes either forward-biased or both reverse-biased. Table 4.6 illustrates a 22 GHz difference in 3 dB cutoff frequencies between forward and reverse-biased operation.

A point that can be further emphasized by the inclusion of eye diagrams, is the fact that the MZM does indeed transmit an optical signal at frequencies higher than 10 Gbps, which was the highest bit rate for which an eye diagram was given in the PP case. Figure

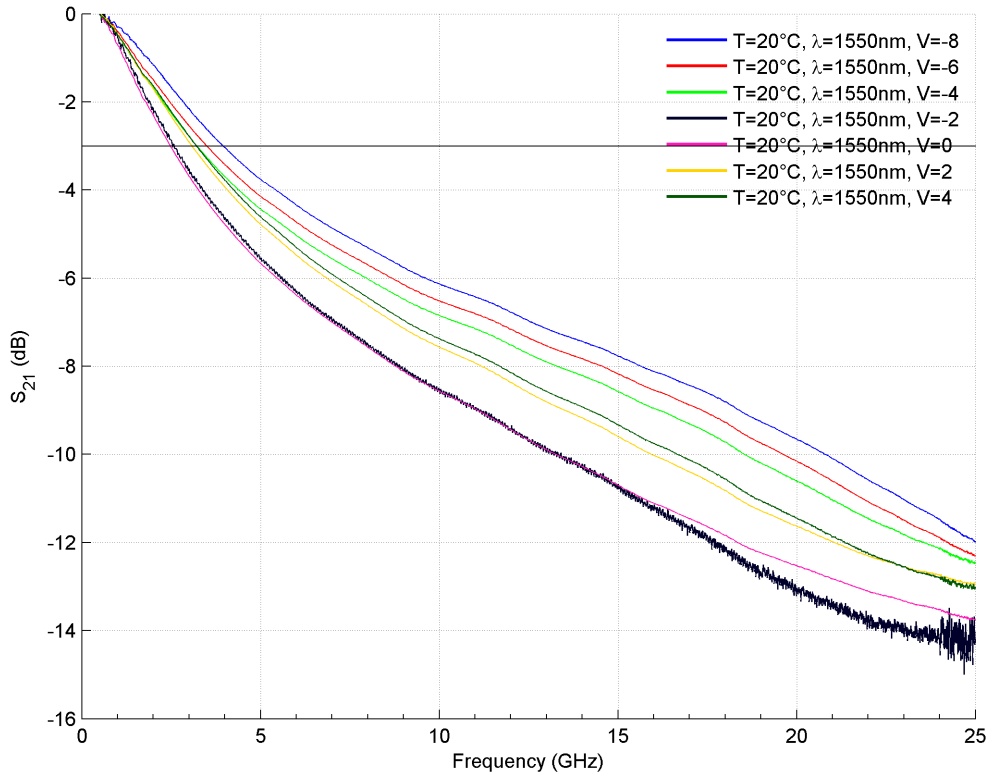


Figure 4.18: S_{21} for the 2000-μm PP MZM at a variety of bias voltages and 1550 nm, using the straight port. The best device performance is achieved using a −3.5 V bias, and results in $f_{3\,\mathrm{dB}} = 3.94$ GHz and $f_{6\,\mathrm{dB}} = 9.62$ GHz.

4.22 shows eye diagrams using a 1010 pattern, which is a pattern that spans a much smaller frequency spectrum than would a PRBS pattern. As a result, the eye diagram is very much open compared to what the MZM would transmit if using a PRBS pattern at the same bit rate. The fact that there are two eyes present on the diagram is the result of an improperly calibrated delay line on the comparators of the BERTScope.

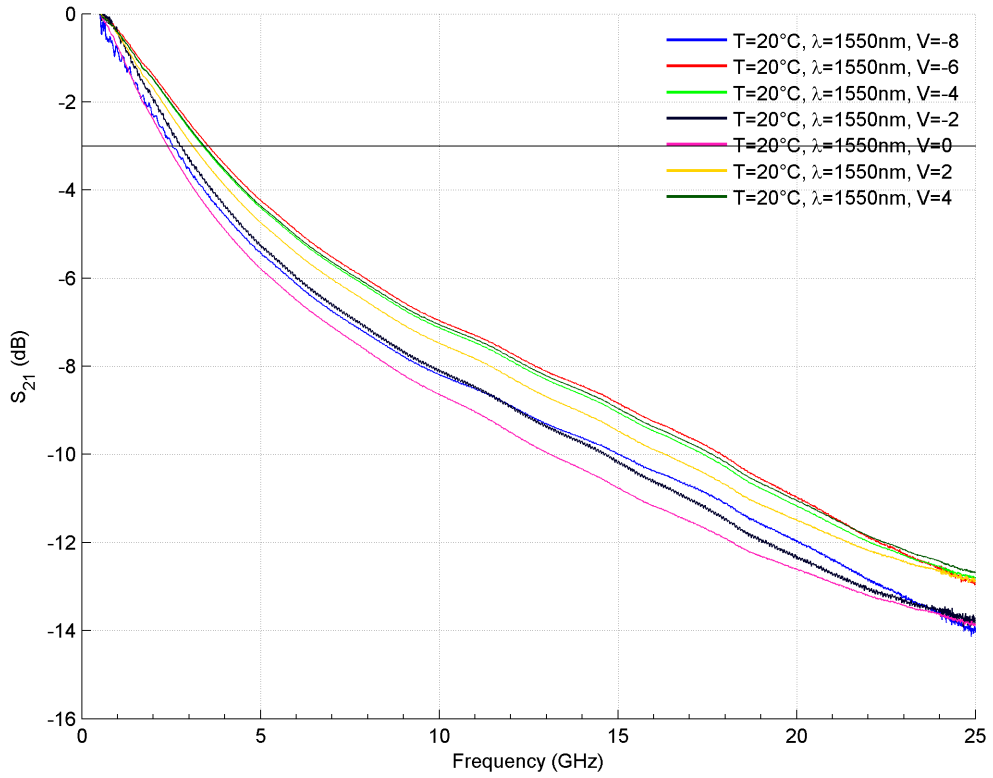


Figure 4.19: S_{21} for the 2000-μm PP MZM at a variety of bias voltages and 1550 nm, using the cross port. The best device performance is achieved using a -3 V bias, and results in $f_{3\,\mathrm{dB}} = 3.54$ GHz and $f_{6\,\mathrm{dB}} = 7.97$ GHz.

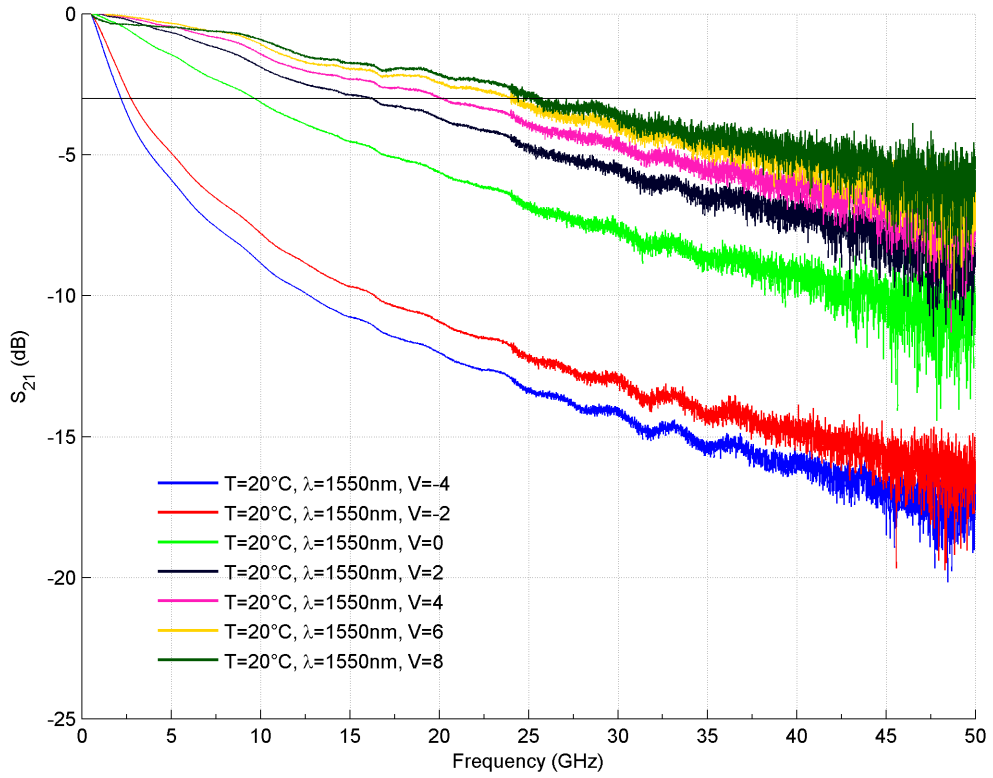


Figure 4.20: S_{21} for the 1500-µm TWE MZM at a variety of bias voltages and 1550 nm, using the straight port. The best device performance is achieved using an 8 V bias, and results in $f_{3\,\mathrm{dB}} = 24.44$ GHz and $f_{6\,\mathrm{dB}} = 44.35$ GHz. The large amount of noise at roughly 25 GHz is due to a combination of the PNA internally changing to a mixer with more noise, and the low level of input optical power to the LCA.

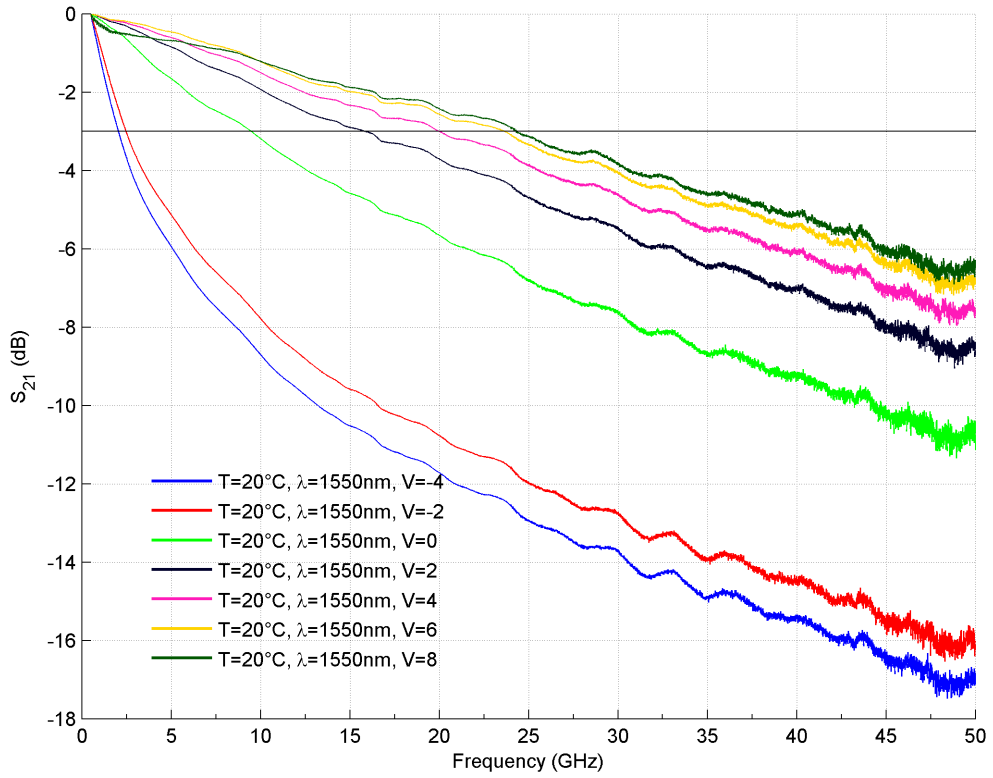
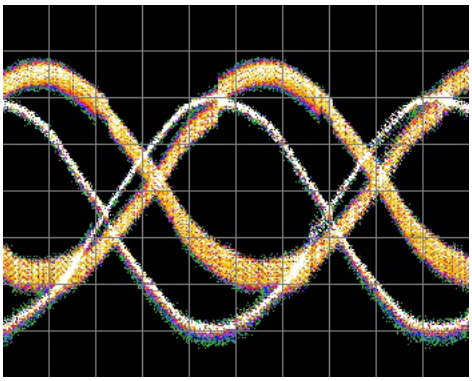
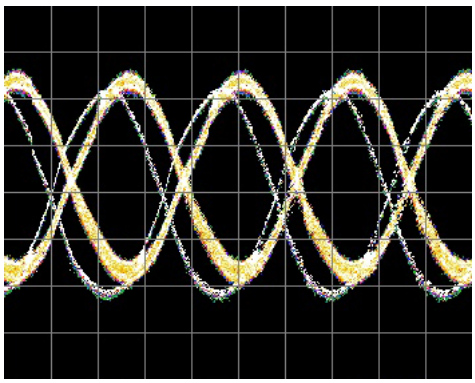


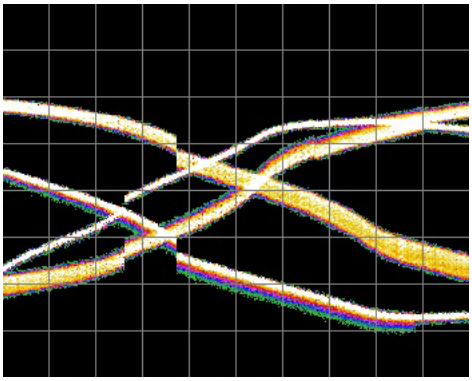
Figure 4.21: S_{21} for the 1500-µm TWE MZM at a variety of bias voltages and 1550 nm, using the cross port. The best device performance is achieved using an 8 V bias, and results in $f_{3\,\mathrm{dB}} = 24.30\,\mathrm{GHz}$ and $f_{6\,\mathrm{dB}} = 47.21\,\mathrm{GHz}$.



(a) 1000-μm PP MZM with 14 mV per division



(b) 2000-μm PP MZM with 18 mV per division



(c) 1500-μm TWE MZM with 23 mV per division

Figure 4.22: Eye diagrams for the two PP MZMs using a 1010 test pattern at 26 Gbps. The fact that the eye is wide open demonstrates that the MZMs can indeed transmit at high frequencies, but the maximum error-free bit rate is limited by the fact that the MZM has a different transmission response at each frequency. As a result, when using a PRBS pattern that consists of several frequencies (i.e. a PRBS square wave), the eye closes.

4.3 Bit Error Rate Characterization

During normal operation, the PP MZM is biased at a fixed DC voltage and driven with a symmetric, ideally square, wave, which is the electrical data signal. The electrical driving signal alternately drives the phase shift, and thus drives the optical output of the MZM between high and low states. The previous DC characterization section suggest the optimum values for the bias and driving voltages – the ideal bias point being equidistant from the minimum and maximum transmission voltages, and the ideal driving voltage being V_π. If one selects something other than this bias point and uses V_π as the driving voltage, the MZM will be overmodulated, and the ER reduced from its maximum value. If one selects the ideal bias point, but uses a driving voltage greater in magnitude than V_π, the MZM will also be overmodulated, with a commensurate decrease in ER. Using the ideal bias point with a driving voltage less in magnitude than V_π results in undermodulation, and decreases ER. Additionally, the combination of bias and driving voltage should be chosen to ensure that the device remains in depletion mode operation to avoid free-carrier lifetime limitations. S_{21} plots also suggested the proper bias voltage based on bandwidth limitations, and coupled with the DC characterization, the ideal bias point can be chosen. It is possible to drive the device slightly into forward bias, with a voltage less than the built-in voltage, which causes little current to flow and thereby avoids free-carrier lifetime limitations [61]. As demonstrated earlier though, one can also trade ER for power savings by operating at lower bias voltages or using smaller swings on the driving voltage.

To operate at the fastest possible data rates, the device must use the ideal bias and driving voltages; however, the operational requirement for data rates may be such that wider tolerances are afforded to the bias and driving voltages. Error-free operation results in the most stringent tolerances for bias and driving voltages, but it is also possible to trade the transmission accuracy (i.e. intentionally increase the number of errors in the transmitted

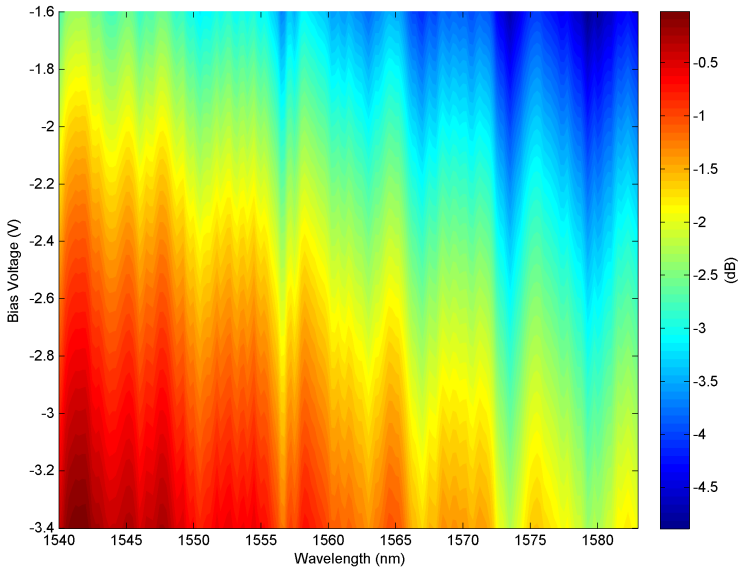
signal) for wider tolerances. While the increase in transmission errors is not always viable, receivers employing forward error correction can typically tolerate a BER of $\leq 10^{-3}$ [16].

Contour plots (Figures 4.23, 4.24, and 4.25) and ER have been broken out from the DC characterization data for the bias and driving voltages selected during the BER characterization. Values for ER are presumed *a priori* to not be identical between the DC characterization and eye diagrams, because of the use of an EDFA at the output of the MZM, as well as the use of a different photodiode. However, the ERs are similar insomuch as the 1000-μm PP MZM cross port has a higher ER in both tests, and the 2000-μm PP MZM straight port has a higher ER. The summary statistics for the expected values from DC characterization are given in Table 4.7.

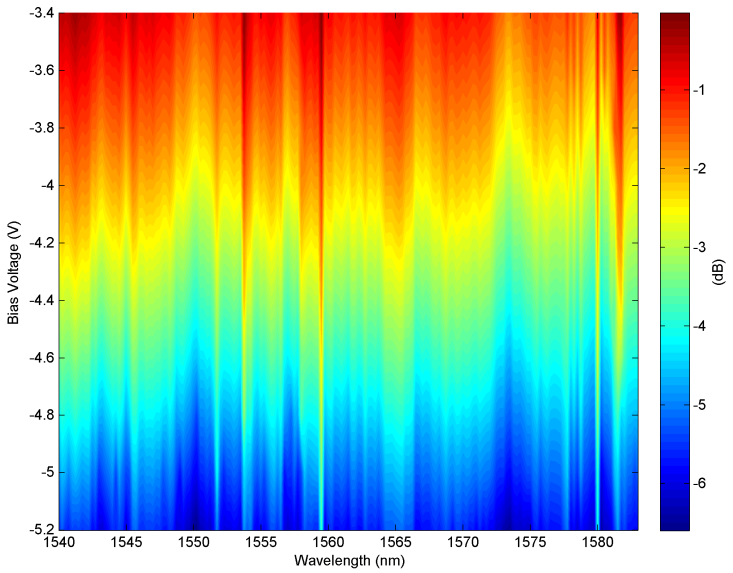
Table 4.7: Expected ERs from the DC characterization, when operating at the same conditions (bias voltage and driving voltage) as used during the BER tests. V_H designates the voltage that produces a high level output optical power, and V_L the low level output optical power.

Device	Port	Bias (V)	V_H (V)	V_L (V)	μ, ER (dB)	σ, ER (dB)	ER at 1550 nm
1000 PP	Straight	-2.5	-3.5	-1.5	2.41	0.06	2.33
1000 PP	Cross	-4.3	-3.3	-5.3	4.71	0.30	5.11
2000 PP	Straight	-2.6	-1.6	-3.6	7.18	2.67	5.18
2000 PP	Cross	-2.45	-3.45	-1.45	2.42	0.22	2.29
1500 TWE	Straight	4	3	5	2.05	0.14	1.87

Figures 4.26 and 4.27 show the range of bias voltages over which the PP MZMs can operate, for a given speed and a given maximum error rate. Figure 4.26 is for the 1000-μm PP MZM, and Figure 4.27 is for the 2000-μm PP MZM. The dashed lines show the minimum possible bias voltage, and the solid lines show the maximum possible bias

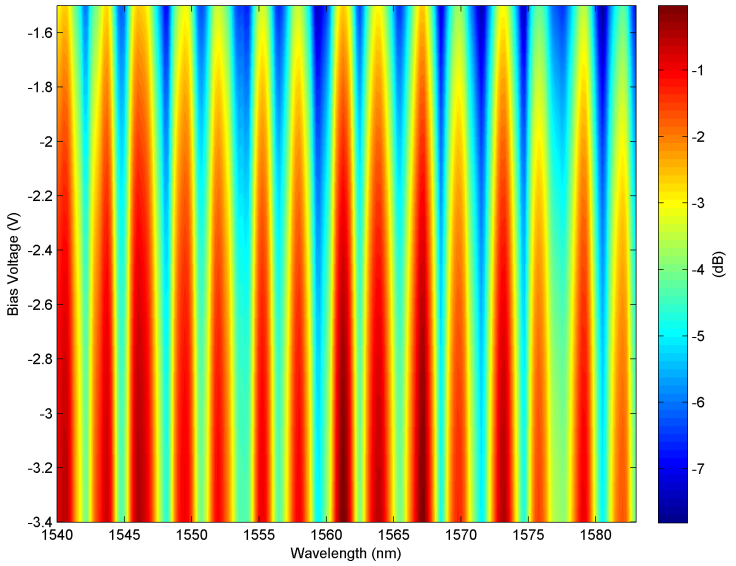


(a) Straight Port

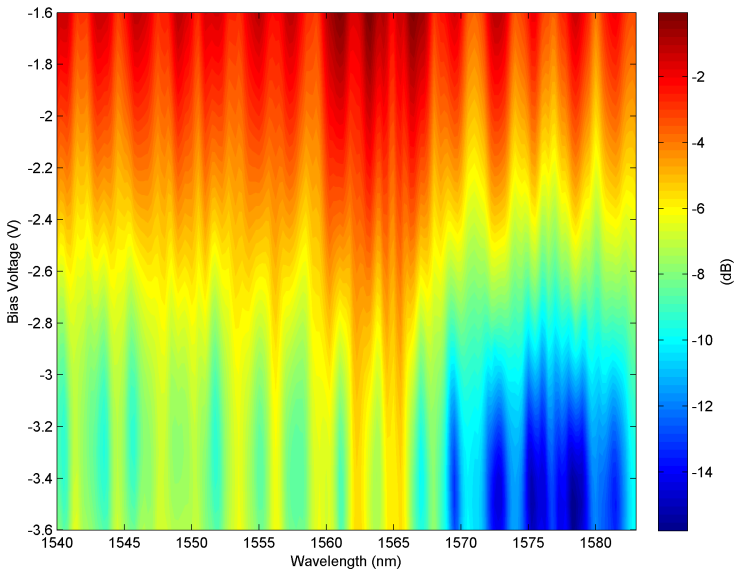


(b) Cross Port

Figure 4.23: The transmission of the 1000-μm PP MZM as a function of wavelength and bias for the specific bias and operating voltages used during BER experimentation, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.



(a) Straight Port



(b) Cross Port

Figure 4.24: The transmission of the 2000-μm PP MZM as a function of wavelength and bias for the specific bias and operating voltages used during BER experimentation, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.

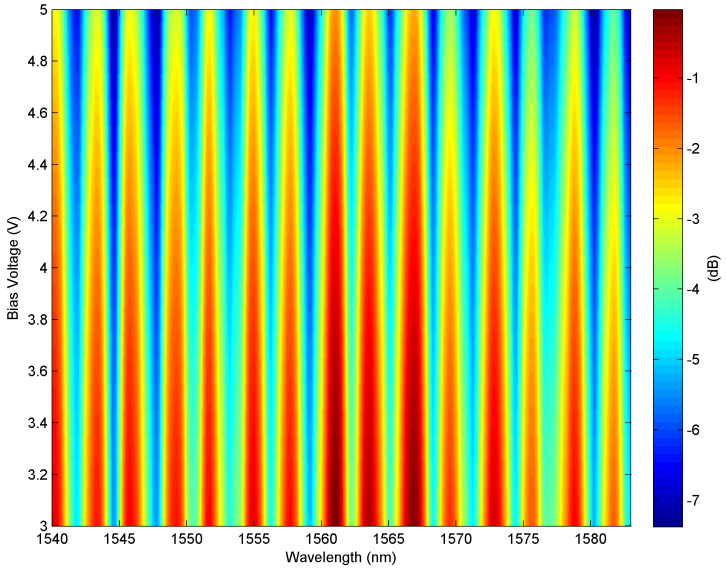


Figure 4.25: The transmission of the 1500-µm TWE MZM as a function of wavelength and bias for the specific bias and operating voltages used during BER experimentation, taken on the straight port, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.

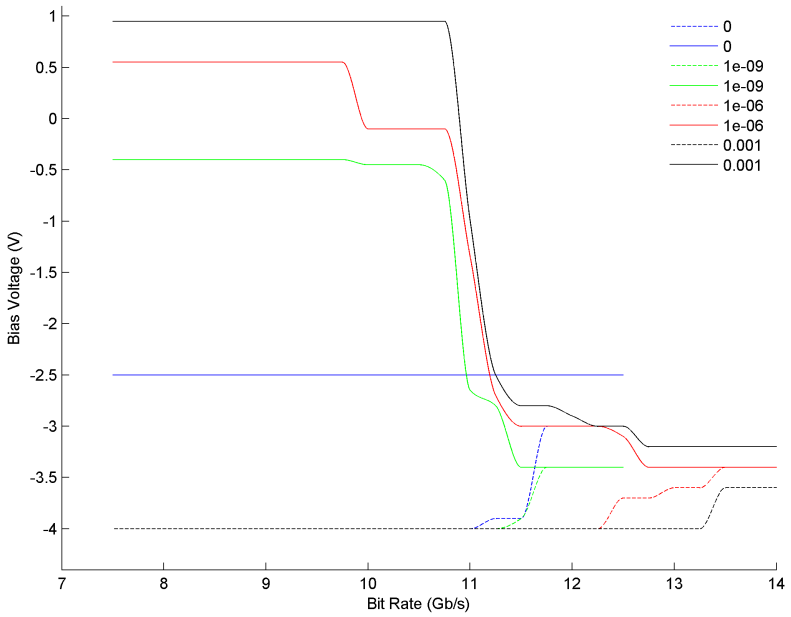
voltage. The contour lines are all color coded, with each color representing a different maximum bit error rate. To operate at a given bit rate under a specified bit error rate, one must bias the MZM somewhere in the range of the solid and dashed lines of the appropriate color. The data was taken using a 2-V_{pp} driving voltage, which was the maximum output voltage the BERTScope can provide. Certainly, larger driving voltages would increase the bias voltage tolerances. It is only logical that the contour lines for larger BERs lie at values that are greater in magnitude than those with small BERs. As the bit rate increases, the range of possible driving voltages tapers down.

Table 4.8: The max data rates achieved at the specified BERs, taken from the data presented in Figure 4.26 for the 1000-μm PP MZM.

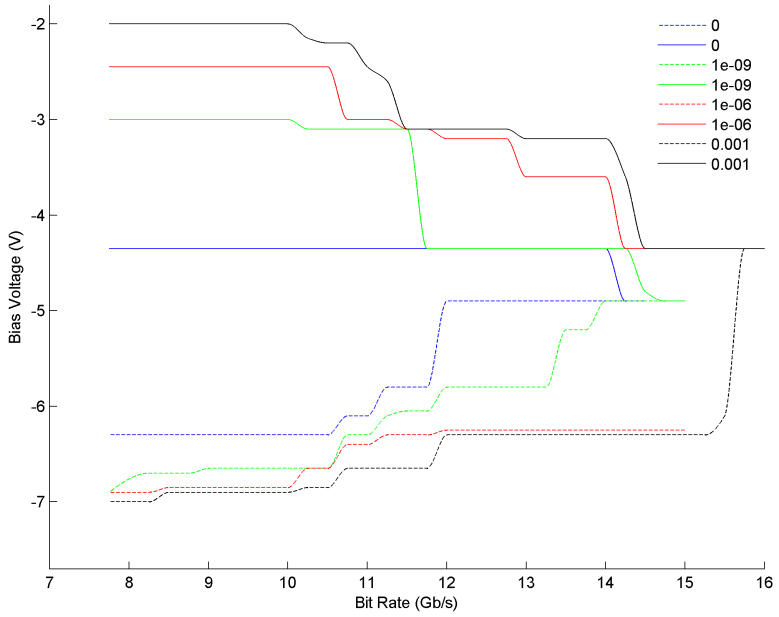
Maximum BER	Max Bit Rate, Straight (Gbps)	Max Bit Rate, Cross (Gbps)
0	12.5	14.5
1×10^{-9}	12.5	15
1×10^{-6}	14	15
1×10^{-3}	14	16

Table 4.9: The max data rates achieved at the specified BERs, taken from the data presented in Figure 4.27 for the 2000-μm PP MZM.

Maximum BER	Max Bit Rate, Straight (Gbps)	Max Bit Rate, Cross (Gbps)
0	10.75	10.25
1×10^{-9}	11	11
1×10^{-6}	11.25	11
1×10^{-3}	11.75	11

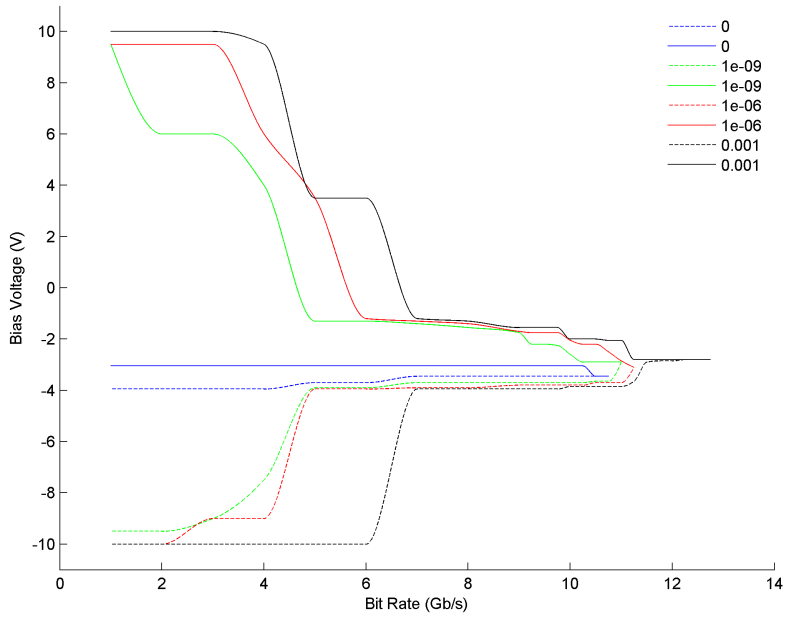


(a) Straight Port

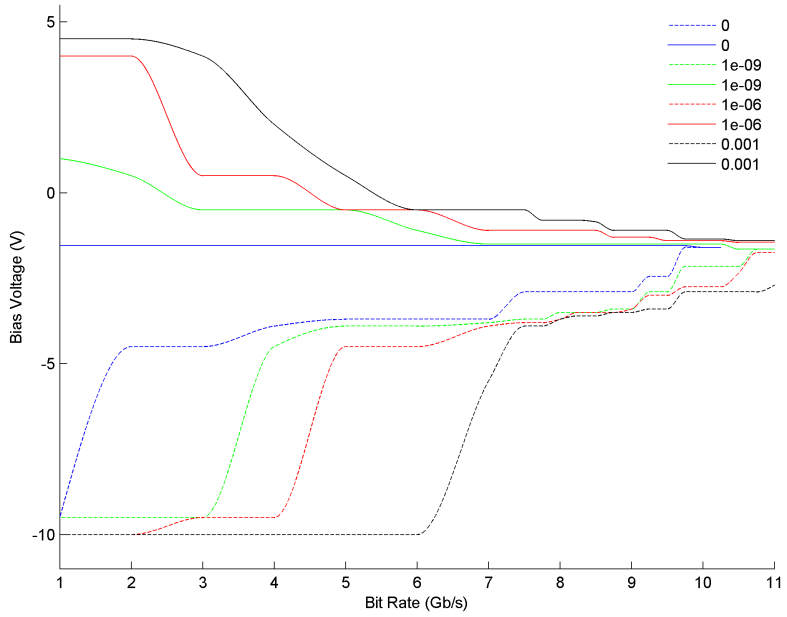


(b) Cross Port

Figure 4.26: BER contours for the 1000-µm PP MZM as a function of bias voltage, taken from the straight and cross ports at 1550 nm.



(a) Straight Port



(b) Cross Port

Figure 4.27: BER contours for the 2000-µm PP MZM as a function of bias voltage, taken from the straight and cross ports at 1550 nm.

For a BER test, the BERTScope must decide upon a detector threshold value and a delay time (measured in seconds after the signal increases or decreases by 10 % from the low or high logic voltage levels, respectively) at which to latch in the received electrical signal. The ideal threshold value is expected to be the halfway point between the high and low logic voltage levels, and the delay time is ideally expected to be half the period of the bit rate. However, it is possible that the lowest BER may be found in a location other than previously mentioned, due to amplitude noise and timing jitter. The BERT therefore searches for the threshold and delay values that minimize the BER, starting with the expected values and moving outward from there. Experimentation showed that the manufacturer's algorithm which optimized the threshold and delay values was not entirely deterministic, and significant changes in BER would result from small deviations in optimal threshold voltage and delay times (on the order of a few millivolts or picoseconds, respectively), which could be achieved by either manually setting threshold and delay values or multiply running the optimization algorithm. The lack of accuracy in the algorithm to optimize threshold voltage and delay time resulted in BER contours that are not differentiable, unlike what one would expect. The contours one would expect are differentiable, and smoothly taper from a value large in magnitude to one small in magnitude. For the most part, the BER contours are as expected, however, Figure 4.26a is illogical from the perspective that starting at 11 Gbps, the 0 BER contour is at a larger bias voltage than the 1×10^{-9} BER contour, and shortly thereafter has a bias voltage higher than all other BER contours. Perhaps another oddity of the figures is that the contours overlap, as they do for the minimum bias voltage BER contours in Figure 4.26a up to 11 Gbps. These oddities could be removed with multiple trials at each sampling point, and fundamentally amount to the effects of experimental error, i.e. noise.

Tables 4.8 and 4.9, also give the maximum bit rate achieved for each BER contour. The fact that the 1000-μm PP MZM displays higher data rates than the 2000-μm PP MZM

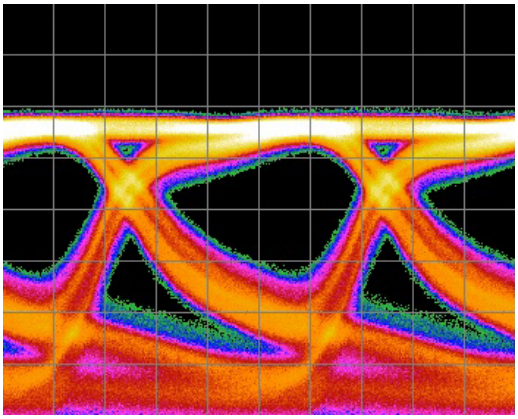
for each contour is no surprise, as the 1000-μm PP MZM has half the capacitance, and thus should achieve twice the bandwidth. In these figures and tables, the doubling in bandwidth of the 1000-μm PP MZM does not translate to twice the maximum bit rate of the 2000-μm PP MZM, because the driving voltage is not set to its optimum value of V_π. Additionally, the data for the 1000-μm PP MZM begins at 7 Gbps, whereas the data for the 2000-μm PP MZM begins at 1 Gbps. This is merely due to time constraints, but the contours starting at 7 Gbps can safely be projected back to 1 Gbps, though testing the 1000-μm PP MZM starting at 1 Gbps would likely reveal wider operating tolerances than simply projecting the data at 7 Gbps.

A point that these figures emphasize, is that high-speed operation forces one to operate the device at negative bias, which one would think corresponds to depletion mode as opposed to injection mode. This is demonstrated by the fact that all contours eventually lie below 0 V. The maximum bias voltage also steps down in voltage with the bit rate, far faster than the minimum bias voltage steps up, indicating the device's propensity to operate at high-speeds while negatively biased. As can also be seen in the figures, error-free operation is only achieved while the device is operated at negative bias. What is important to remember though, is that the PP devices always have one diode forward-biased, and the other reverse-biased, and the S_{21} response shows that the bandwidth should be identical for both cases. As such, this behavior is not clear. To be certain, operating at different bias points would change the logic from normal to inverted; however, these figures amalgamate data from multiple experiments where the photodetector was connected to both the normal and inverted error analyzer inputs.

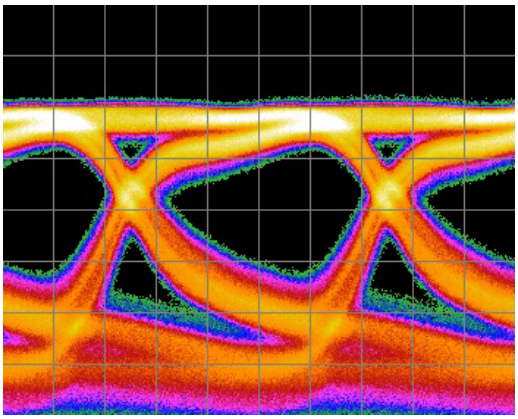
The fundamental point that these figures emphasize is that wider operating tolerances are possible when the modulator is not stressed by its bandwidth limitations. At low data rates, a reduction in rise time, fall time, or ER, that is caused by sub-optimally biasing the modulator has little effect, as there is still a large amount of time between the conclusion

of one transition and the start of the next transition in the bit pattern. At higher data rates, optimally biasing the modulator becomes necessary, as bandwidth limitations of the device already detract from performance, and negative performance effects caused by a less than ideal bias point cannot be endured.

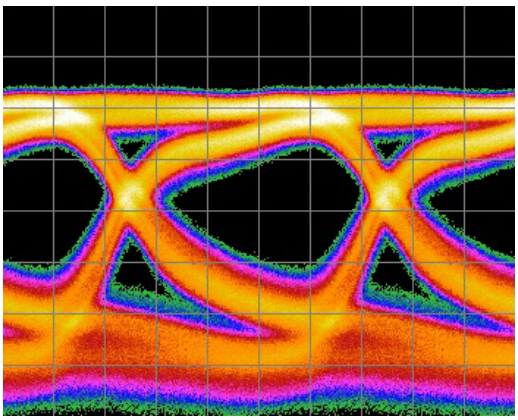
Due to resource constraints, BER contours are not available for the 1500-μm TWE MZM as they were available for the PP MZMs. While the BER contours provide good information on how one can achieve power savings by varying the bias voltage within a given tolerance, the tests were initially conducted in an effort to confirm the ideal bias voltage at which to operate the device, as suggested from the DC characterization and S_{21} tests. In order to achieve the same goal for the TWE MZM, testing was performed on both output arms at a single bit rate of 10 Gbps while sweeping bias voltages from −10 V to 10 V with a 2-V_{pp} driving voltage. The resultant data was analyzed to confirm the best operating bias voltage for each output arm. The TWE MZM operated error-free at 10 Gbps over a wide range of bias voltages, from −0.5 V to 8.5 V — ultimately, a bias voltage of 4 V was chosen, as it produced a relatively symmetric eye with the largest ER. Some eye diagrams from this testing are given in Figure 4.28. Looking at the 6 V eye would suggest that the eye is more symmetric than the 4 V, and this is indeed the case; however, the additional 2 V of power results in higher energy consumption. Despite the fact that a bias voltage of 4 V was used, the device doping means that a positive voltage on the center pin reverse biases the junctions in the two arms. During the course of testing, the BERTScope was unable to lock onto the pattern produced when operating off the cross output port. This is unfortunate, because as noted in the previous section, the cross port appears to produce a significantly larger ER than the straight port, which would translate to fewer errors at any given bit rate. As a result, BER data can only be presented for the straight port. It is not clear whether the inability to synchronize the BERT with the data stream produced by the modulator is a consequence of the experimental setup, or of the modulator design.



(a) Bias voltage of 2 V with 39 mV per division, ER = 7.21 dB, Eye Height = 80.2 mV



(b) Bias voltage of 4 V with 38 mV per division, ER = 6.82 dB, Eye Height = 78.2 mV



(c) Bias voltage of 6 V with 35 mV per division, ER = 6.40 dB, Eye Height = 75.1 mV

Figure 4.28: Eye diagrams for the 1500-µm TWE MZM PP MZM, using the straight port, taken at 2 V, 4 V and 6 V and 1550 nm.

It is also possible to plot the minimum bit error rate for each PRBS test pattern as a function of the bit rate. This is done by sweeping the bit rate while recording the BER. During the test, the bias voltage is set to its optimum value, that is the bias voltage at which the maximum error-free bit rate was achieved from the previous data sets, and the driving voltage was set to 2-V_{pp}. The wavelength was also fixed at 1550 nm for this testing. The test pattern itself was changed, so that the minimum achievable BER is presented for the PRBS-7, 11, 15, 20, 23, and 31 test patterns.

The experiment results are shown in Figure 4.29 for the 1000-μm PP MZM, and Figure 4.30 for the 2000-μm PP MZM. The expectation is that the higher the value of *n* in the test pattern, the larger the BER at a given bit rate, so in all cases, the PRBS-31 pattern should have a BER greater than or equal to the BER of the PRBS-7 pattern. From this perspective, there are several values which appear to be erroneous. For example, looking at Table 4.10, the maximum achievable error-free bit rate using a PRBS-7 pattern on the straight port is 9.5 Gbps; however the very same MZM operating under the same conditions achieves error-free operation at 10.75 Gbps using a PRBS-31 pattern. The seemingly erroneous value is once again due to the inaccuracy involved with optimizing the threshold voltage and delay time. It can be safely assumed that if the PRBS-31 pattern can achieve error-free operation at 10.75 Gbps, all other patterns can achieve error-free operation at a minimum of 10.75 Gbps. In fact, the very same table lists the PRBS-11 pattern as being error-free at 12 Gbps, so the PRBS-7 pattern should also be error-free at 12 Gbps. The lack of separation between contour lines at high BERs can also be attributed to the problems with algorithm, as the algorithm reaches its decision in less than four seconds, which is less time than necessary to arrive at an optimal decision. As expected, the 1000-μm PP MZM generally achieves faster error-free operation than the 2000-μm PP MZM, in spite of the reduced ER for the 1000-μm PP MZM when operated at the same driving voltage as the 2000-μm PP MZM, once more due to the higher bandwidth and lower capacitance.

Pattern contours for the TWE MZM are available, as shown in Figure 4.31 and Table 4.12. Once again, the driving voltage was limited to 2-V_{pp}, and the bit rate was swept using PRBS-7, 11, 15, 20, 23, and 31 test patterns. As is expected, the maximum error-free bit rate decreases as the pattern length increases, from its maximum of 16 Gbps using a PRBS-7 pattern to 11 Gbps using a PRBS-31 pattern. With a higher bandwidth on the TWE MZM, it is expected that the TWE MZM would outperform the PP MZM at each test pattern. The performance of the TWE MZM is better than or equal to the performance of either of the PP MZMs, with the exception of the 2000-μm PP MZM operating on the straight port and using a PRBS-23 pattern – in this instance the PP MZM was error-free until 11.75 Gbps whereas the TWE MZM was error-free until 11.5 Gbps. The difference between the two is minor though, and the TWE MZM may have been error-free until 11.75 Gbps, being potentially masked by the fact that the BER testing for the TWE MZM used 0.5 Gbps steps and not 0.25 Gbps steps in order to save time. Additionally, as is the case with many experimental errors during the BER tests, a better choice of threshold voltage and delay times may have resulted in fewer errors in transmission, and therefore extended the error-free transmission for the PRBS-23 pattern beyond 11.5 Gbps.

Table 4.10: The maximum error-free data rates achieved when using the specified test pattern, taken from the data presented in Figure 4.29 for the 1000-μm PP MZM.

Test Pattern	Max Bit Rate, Straight (Gbps)	Max Bit Rate, Cross (Gbps)
PRBS-7	9.5	13
PRBS-11	12	11
PRBS-15	10.75	12.75
PRBS-20	8.5	11.25
PRBS-23	8.25	10
PRBS-31	10.75	7.75

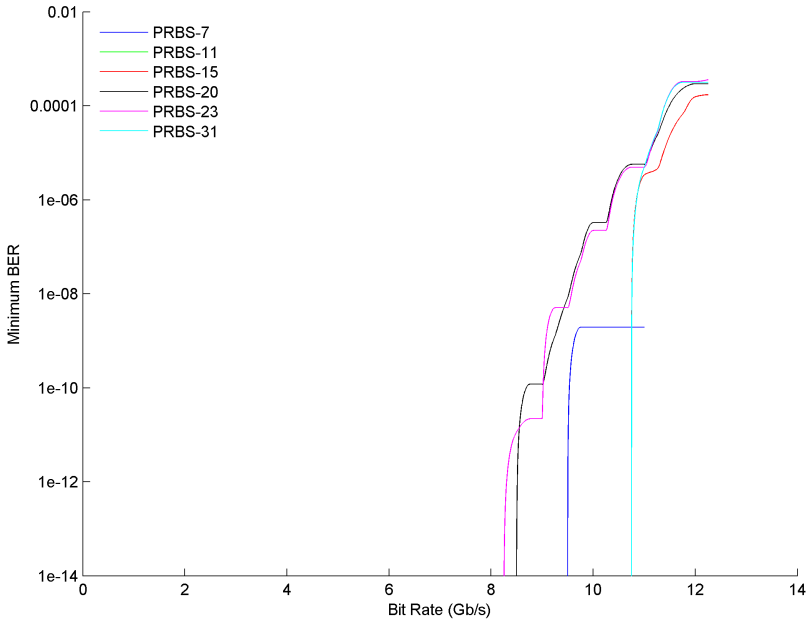
Table 4.11: The maximum error-free data rates achieved when using the specified test pattern, taken from the data presented in Figure 4.30 for the 2000-μm PP MZM.

Test Pattern	Max Bit Rate, Straight (Gbps)	Max Bit Rate, Cross (Gbps)
PRBS-7	11.25	10.5
PRBS-11	10.75	9.25
PRBS-15	8.5	10
PRBS-20	10.75	8
PRBS-23	11.75	7.5
PRBS-31	10.75	8

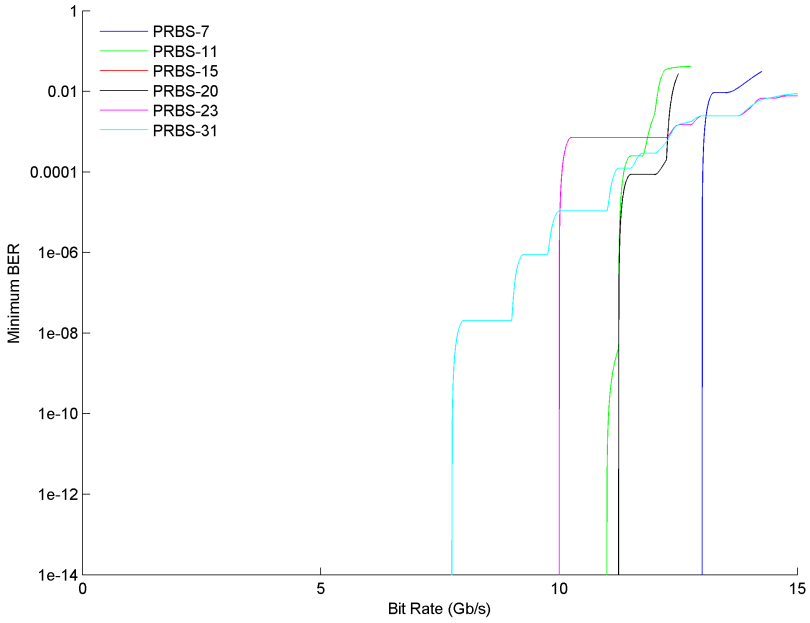
Table 4.12: The maximum error-free data rates achieved when using the specified test pattern, taken from the data presented in Figure 4.31 for the 1500-μm TWE MZM, using the straight port.

Test Pattern	Max Bit Rate (Gbps)
PRBS-7	16
PRBS-11	14
PRBS-15	12.75
PRBS-20	11.5
PRBS-23	11.5
PRBS-31	11

As explained earlier, eye diagrams intuitively show device performance, but do not show events with a low probability. Their inclusion here is to highlight device turn-on/turn-off transient times, and to depict the increase in amplitude noise and timing jitter with bit rate. Figures 4.32 and 4.33 show eye diagrams for the 1000-μm PP MZM at 0.5 Gbps,

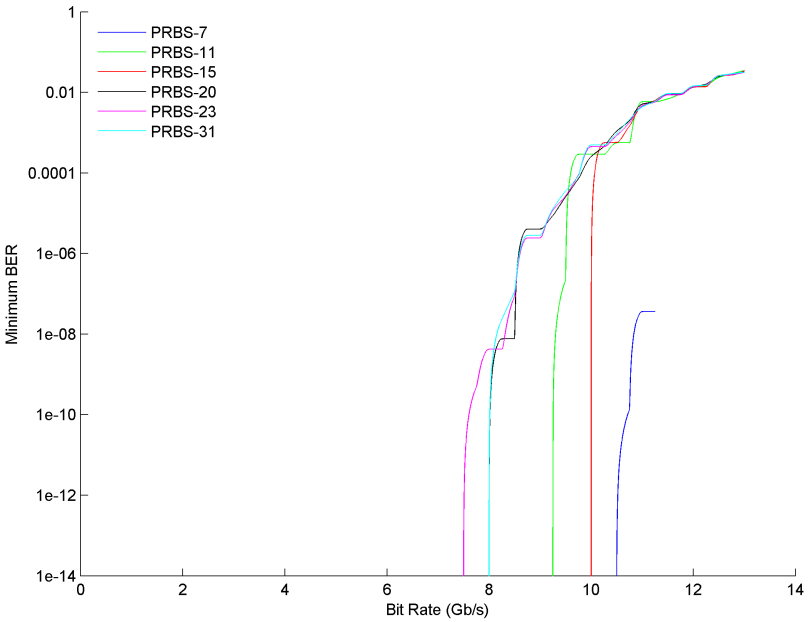


(a) Straight Port

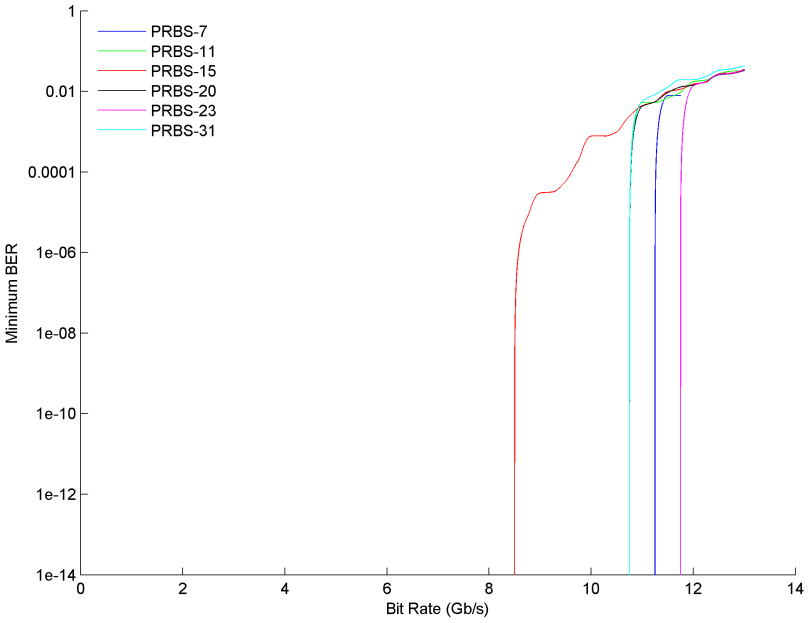


(b) Cross Port

Figure 4.29: Test pattern contours for the 1000-µm PP MZM as a function of bit rate, taken at 1550 nm.



(a) Straight Port



(b) Cross Port

Figure 4.30: Test pattern contours for the 2000-μm PP MZM as a function of bit rate, taken at 1550 nm.

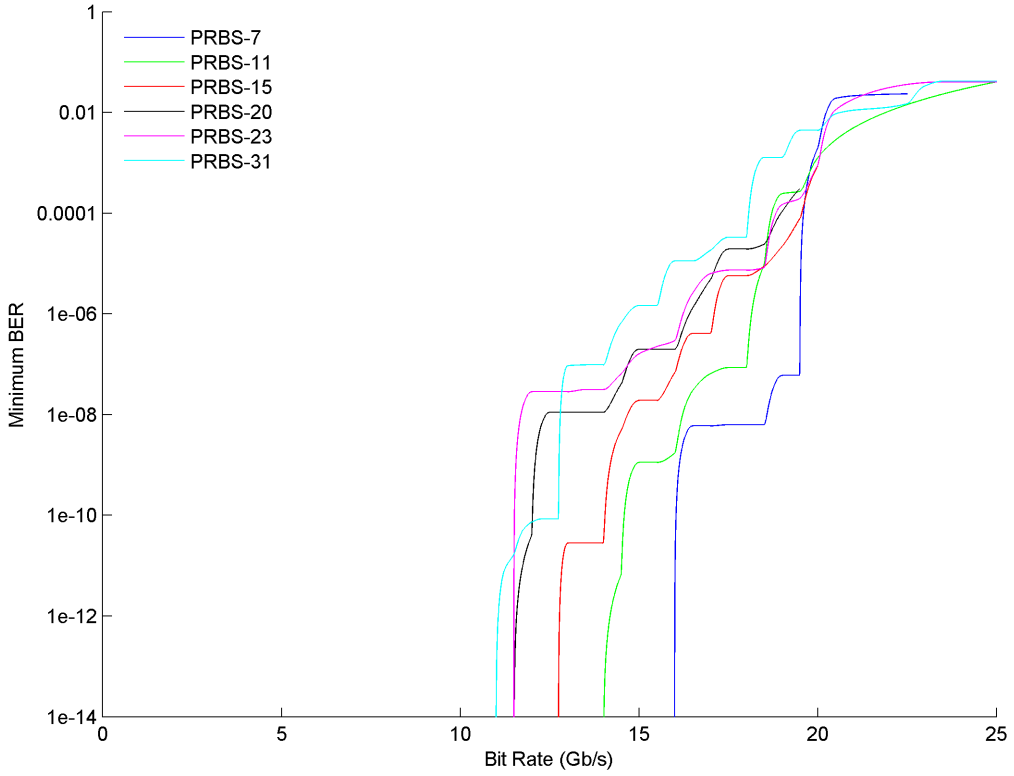


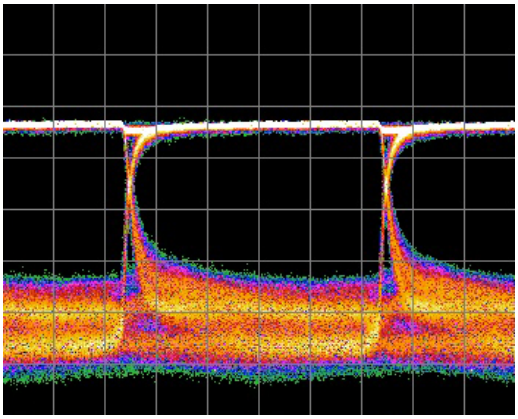
Figure 4.31: Test pattern contours for the 1500-μm TWE MZM as a function of bit rate, taken from the straight port at 1550 nm.

5 Gbps and 10 Gbps, and Figures 4.34 and 4.35 show eye diagrams for the 2000-μm PP MZM at 0.5 Gbps, 5 Gbps and 10 Gbps. Both PP MZMs are error-free at all speeds despite what appears to be a minimally open eye at 10 Gbps. As would be expected, the eye closes down as the bit rate increases. The eye diagrams at high-speed appear to be distorted because of the fact that the eye diagrams were not produced on a sampling oscilloscope, but instead using the BERTScope which does not have an ADC, and instead uses two comparators with delay lines that were not properly calibrated. The amplitude noise and timing jitter also increase with a commensurate rise in data rates. Quantitatively, it is

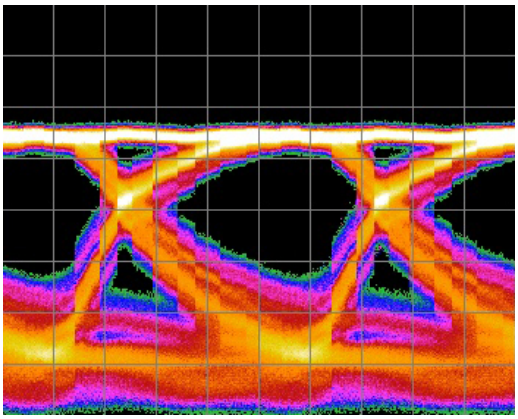
possible to explain the eye closure in terms of an increase in the deterministic components of noise and jitter, which include periodic noise/jitter (PN/PJ), data dependent noise/jitter (DDN/DDJ), and bounded uncorrelated noise/jitter (BUN/BUJ) [62]. Fundamentally, the root cause of an increase in these forms of noise and jitter is the bandwidth limitation of the MZM. Unfortunately, the BERTScope did not reliably capture statistics for jitter, and so it is not possible to track the progression of jitter with increasing data rates.

One also notices that the eye diagram changes from the ideally square output to a sinusoid, albeit a skewed sinusoid, as the bit rate increases. This is a result of the bandwidth limitation of the device. The Fourier expansion of a square wave results in an infinite series of discretely spaced components at harmonics of the fundamental frequency, and it is the high frequency harmonics of the signal that give the square wave its characteristic edges. As the high frequency harmonics are attenuated, the corners begin to round and the square wave looks increasingly like a sinusoid. As the MZM approaches its bandwidth limitation, the high frequency harmonics of the electrical signal are attenuated, and the output changes from the ideal square wave to a sinusoid.

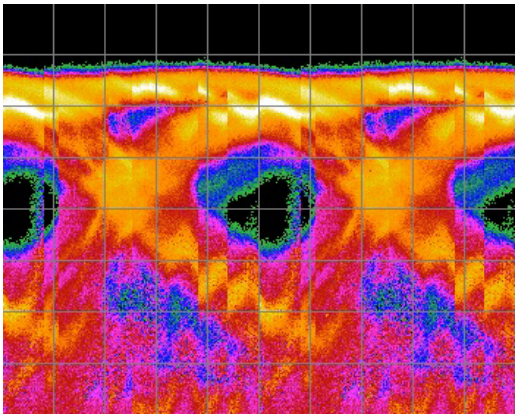
Using the definition of ER as given in Section 2.1.2, the ER is inversely proportional to the bit rate, as would be expected. ERs are reported for each eye in the figure itself. The ER is a potentially misleading figure, as ER for the cross port of the 2000-μm PP MZM at 0.5 Gbps is 9.91 dB, and the ER at 10 Gbps decreases to only 8.27 dB. This would give the impression that the eye diagrams are similar, however the definition of ER as the difference between the mean high and low power levels severely understates the eye closure at 10 Gbps. Another metric that provides some insight about the opening of the eye is the eye height, and is a value reported in mV in each figure. Unlike the ER, the eye height presents a quantitative measure of how the eye closes with increasing data rates, and the fact that the eye height for the cross port of the 2000-μm PP MZM decreases from



(a) 0.5 Gbps with 54 mV per division, ER = 7.31 dB, Eye Height = 165.1 mV

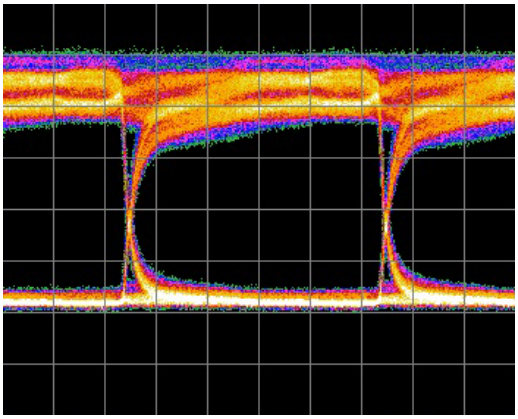


(b) 5 Gbps with 42 mV per division, ER = 4.71 dB, Eye Height = 97.1 mV

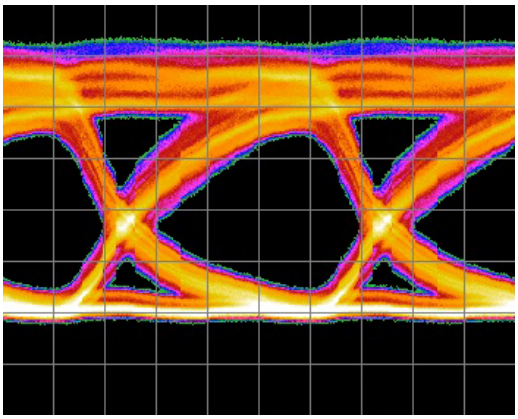


(c) 10 Gbps with 29 mV per division, ER = 4.64 dB, Eye Height = 42.3 mV

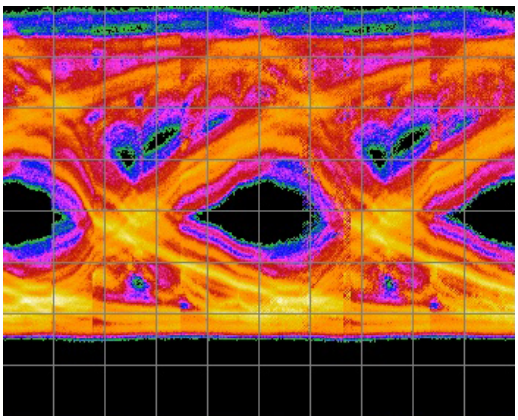
Figure 4.32: Eye diagrams for the 1000-µm PP MZM, using the straight port, taken at 0.5 Gbps, 5 Gbps and 10 Gbps and 1550 nm. All data rates are error-free in operation.



(a) 0.5 Gbps with 59 mV per division, ER = 10.1 dB, Eye Height = 197.2 mV

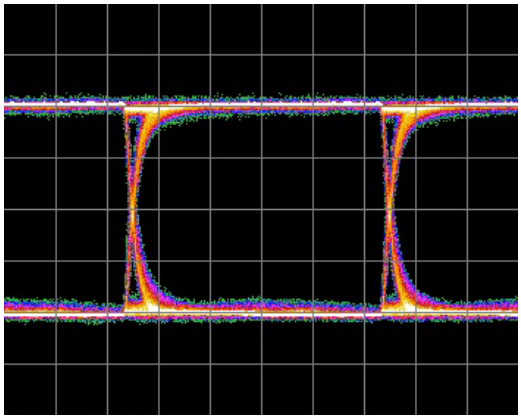


(b) 5 Gbps with 53 mV per division, ER = 8.2 dB, Eye Height = 140.8 mV

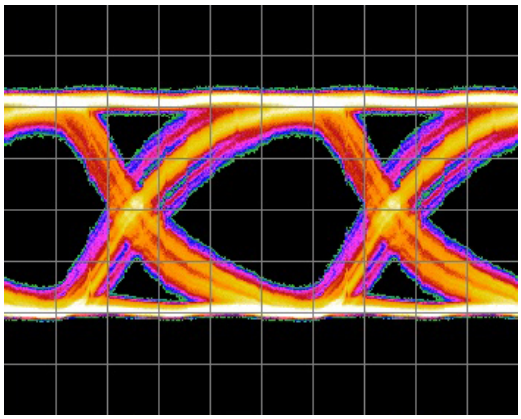


(c) 10 Gbps with 46 mV per division, ER = 7.2 dB, Eye Height = 72.3 mV

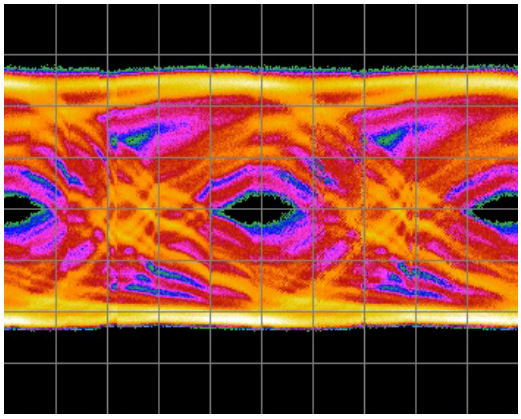
Figure 4.33: Eye diagrams for the 1000-μm PP MZM, using the cross port, taken at 0.5 Gbps, 5 Gbps and 10 Gbps and 1550 nm. All data rates are error-free in operation.



(a) 0.5 Gbps with 54 mV per division, ER = 11.3 dB, Eye Height = 212 mV

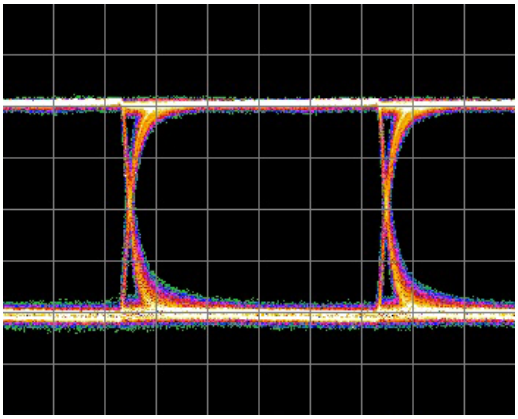


(b) 5 Gbps with 54 mV per division, ER = 9.6 dB, Eye Height = 148.1 mV

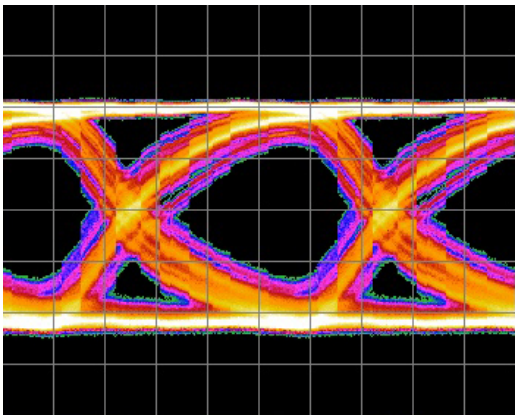


(c) 10 Gbps with 48 mV per division, ER = 8.7 dB, Eye Height = 49 mV

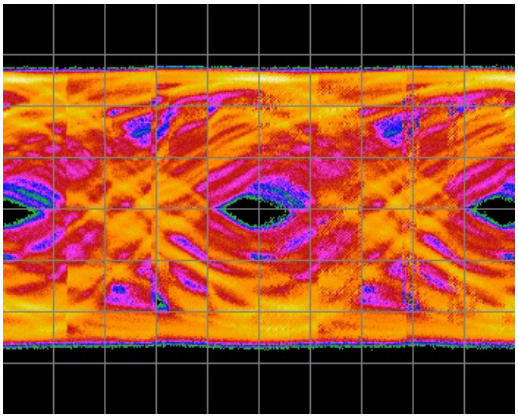
Figure 4.34: Eye diagrams for the 2000-µm PP MZM, using the straight port, taken at 0.5 Gbps, 5 Gbps and 10 Gbps and 1550 nm. All data rates are error-free in operation.



(a) 0.5 Gbps with 63 mV per division, ER = 9.91 dB, Eye Height = 243 mV



(b) 5 Gbps with 61 mV per division, ER = 9.41 dB, Eye Height = 141.2 mV



(c) 10 Gbps with 52 mV per division, ER = 8.27 dB, Eye Height = 36.3 mV

Figure 4.35: Eye diagrams for the 2000-μm PP MZM, using the cross port, taken at 0.5 Gbps, 5 Gbps and 10 Gbps and 1550 nm. All data rates are error-free in operation.

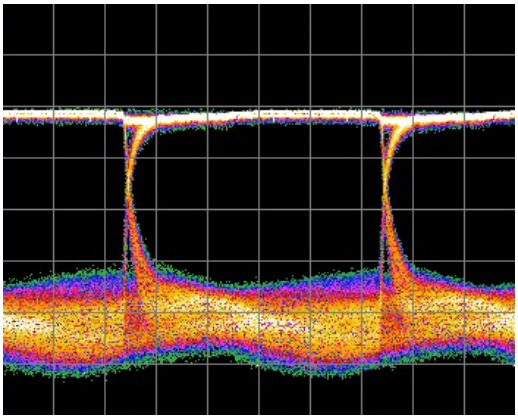
243 mV at 0.5 Gbps to 36.3 mV at 10 Gbps corresponds well with the eye diagrams. The eye height drives the requirement for precision in the optical receiver's threshold circuitry, which likely drives expenses, so operating a receiver using the modulator at 10 Gbps likely is more expensive than at 0.5 Gbps.

From the eye diagrams, the 1000-μm PP MZM cross port displays a higher ER than the straight port, which contradicts expectations based on the full DC characterization. The 2000-μm PP MZM has a higher ER using the straight port rather than the cross port, which also contradicts expectations established by the full DC characterization. Note though, the summary statistics, particularly the value for ER, was calculated as the difference in transmission between the minimum and maximum transmission voltages, irrespective of whether those voltages were located in the forward bias region. Operation at such voltages would have arbitrarily limited bandwidth due to free-carrier lifetimes, and so alternate operating conditions were chosen.

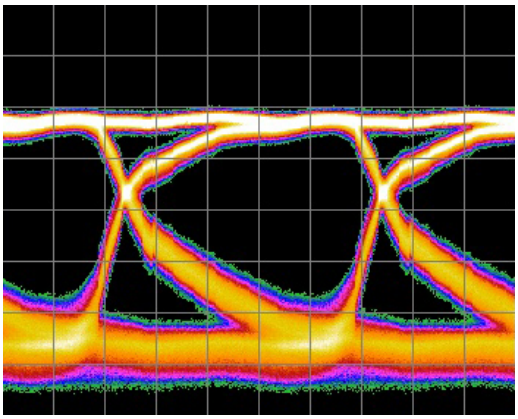
Eye diagrams are also shown in Figures 4.36 and 4.37, for speeds of 0.5 Gbps, 5 Gbps, 10 Gbps, 15 Gbps, 18 Gbps and 20 Gbps for the 1500-μm TWE MZM. The eyes appear to open even up until 20 Gbps, however, as explained earlier, an eye diagram does not reveal performance limitations that a BER test may show. The increase in amplitude noise and timing jitter results in eye closure, and is a consequence of bandwidth limitations. Despite the apparent opening at 18 Gbps and 20 Gbps, the two have BERs of 6.25×10^{-9} and 1.99×10^{-3}, respectively. It is difficult to accept that a change of only 2 Gbps would have impacted device performance by six orders of magnitude, and therefore the BERTScope at 20 Gbps may not have used the appropriate threshold voltage and delay time. Additionally, as can be seen in the eye diagrams, there is a large imbalance between rise and fall times, with fall times markedly longer. This imbalance leads to significant intersymbol interference, which limits the maximum bit rate. If the device was operated in injection

mode, the asymmetry in rise/fall times could be attributed to carrier dynamics and carrier lifetimes, but that is not the case here, as the device was operated in depletion mode.

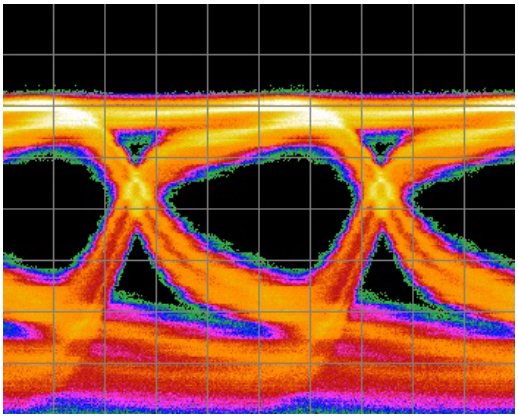
While the TWE MZM did not always outperform the ER of the PP MZMs at low data rates such as 0.5 Gbps — the 1000-μm PP MZM has an ER of 10.1 dB on the cross port, and the 2000-μm PP MZM has an ER of 11.3 dB on the straight port — the TWE MZM does have the largest eye height at 10 Gbps; eye height being the metric which best describes the performance of the MZM when it is bandwidth limited. At an ER of 8.21 dB at 0.5 Gbps, the performance of the MZM at low data rates is not far off from the performance of either of the PP MZMs. The TWE MZM achieved an error-free bit rate of 16 Gbps, which exceeds the maximum error-free bit rate of the 1000-μm and 2000-μm PP MZMs by 3 Gbps and 4.75 Gbps, respectively. The limitation of 2-V_{pp} driving voltage would have increased these margins, by fully driving the MZM to its maximum and minimum output optical levels.



(a) 0.5 Gbps with 50 mV per division, ER = 8.21 dB, Eye Height = 165.7 mV

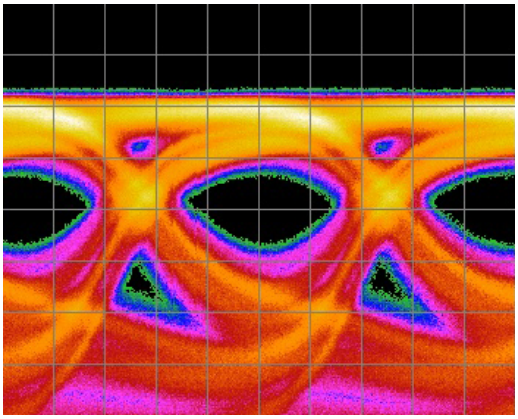


(b) 5 Gbps with 43 mV per division, ER = 7.08 dB, Eye Height = 115.5 mV

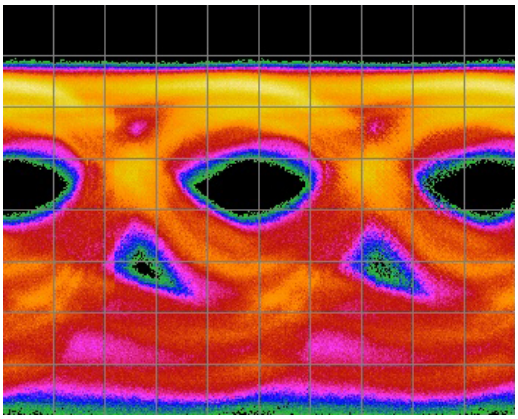


(c) 10 Gbps with 36 mV per division, ER = 6.72 dB, Eye Height = 74.9 mV

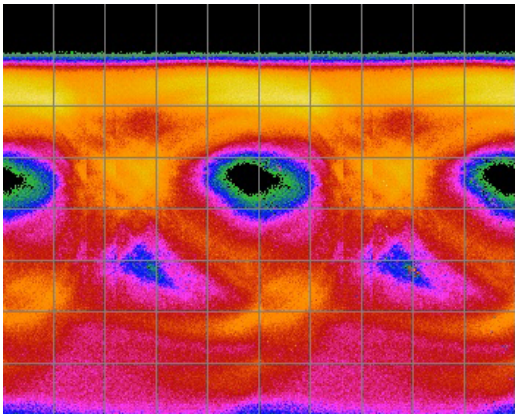
Figure 4.36: Eye diagrams for the 1500-μm TWE MZM PP MZM, using the straight port, taken at 0.5 Gbps, 5 Gbps and 10 Gbps and 1550 nm. All data rates are error-free in operation.



(a) 15 Gbps with 33 mV per division, ER = 6.09 dB, Eye Height = 49 mV



(b) 18 Gbps with 32 mV per division, ER = 6.04 dB, Eye Height = 38.4 mV



(c) 20 Gbps with 31 mV per division, ER = 5.90 dB, Eye Height = 25.7 mV

Figure 4.37: Eye diagrams for the 1500-µm TWE MZM, using the straight port, taken at 15 Gbps, 18 Gbps and 20 Gbps and 1550 nm. Although the eye appears to have a sufficiently large opening in all cases, the MZM is only error-free at 15 Gbps, and has a BER of 6.25×10^{-9} and 1.99×10^{-3} at 18 Gbps and 20 Gbps, respectively.

4.4 Power-Penalty Measurements

Power-penalty cures are another important measurement to make when characterizing optical modulators for communications applications. They show the effect of received power on BER and can be taken at different data rates or over different transmission distances. The expectation is that to ensure error-free operation, the amount of received power needs to increase with either bit rate or transmission distance. The figures given for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs — Figures 4.38, 4.40, and 4.42, respectively — show the BER versus received power at 3 Gbps, 5 Gbps and 10 Gbps at 1550 nm. Strictly speaking, the data shown in these plots is not BER versus received power, but rather Q-factor versus received power, with the labels on the y-axis at the Q-factor corresponding to the BER listed. Original data for BER versus received power is given in Appendix A, and appears to follow an exponential trend. The figures in the appendix plot the actual data points as circles, and use the data to compute an exponential fit, which is also plotted. Deviations from the exponential fit can once again be attributed to the inaccuracy caused by the BERTScope choosing a suboptimal threshold voltage or delay time. This inaccuracy was compensated for by oversampling at each optical power level and discarding statistical outliers, i.e. taking multiple BER measurements at a given power level, which forced the BERTScope to re-optimize threshold voltage and delay time with each BER measurement.

The importance, and intuitiveness of these power-penalty measurements, is the fact that it takes more power at the receiver in order to accurately receive a high-speed signal than it would to accurately receive a low power signal. This is simply explained by the fact that the eye height, and therefore the difference between high and low logic levels, decreases with an increase in bit rate. Increases in data rates contribute to the effects

Table 4.13: Mean and standard deviations for the power penalties for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs. The labels 3-5 Gbps are meant to correspond to the power penalty incurred when choosing to operate at 5 Gbps instead of 3 Gbps, and the same idea holds true for the 3-10 Gbps state.

Device	μ, 3-5 Gbps (dBm)	σ, 3-5 Gbps (dBm)	μ, 3-10 Gbps (dBm)	σ, 3-10 Gbps (dBm)
1000 PP	1.20	0.09	2.68	0.28
2000 PP	0.78	0.02	4.08	0.73
1500 TWE	1.46	0.13	3.30	0.09

of noise and jitter, and increases in transmission distance result in heightened degrees of dispersion, all which serve to decrease the opening of the optical eye. Further decreasing the optical power also serves to decrease the difference between high and low logic levels, leading to inaccurate quantization. Consequently, the amount of power at the receiver will need to increase in order to maintain sufficient differentiation between high and low logic levels.

As can be seen from Figures 4.39 and 4.41 and the summary statistics given in Table 4.13, the power-penalty between different data rates on the same device is not constant as expected, with the one exception being the 2000-μm PP MZM 3-5 Gbps curve. There are, however, two discernible trends: the first and most obvious is a larger power-penalty between the 3-10 Gbps curves than between the 3-5 Gbps curves, and the second is the power-penalty curves begin to converge at high BER. The former has already been explained, and the latter is mathematically attributable to the non-constant power-penalty between curves. Physically though, the mechanism for such behavior is indeterminate.

The fact that the 2000-μm PP MZM has a minimal power-penalty between 3-5 Gbps simply speaks to the fact that the device isn't bandwidth limited at 5 Gbps, and so the curves should be expected to be nearly atop one another. One then questions why the case is not the same for the 1000-μm PP MZM, which has twice the bandwidth. That can be attributed

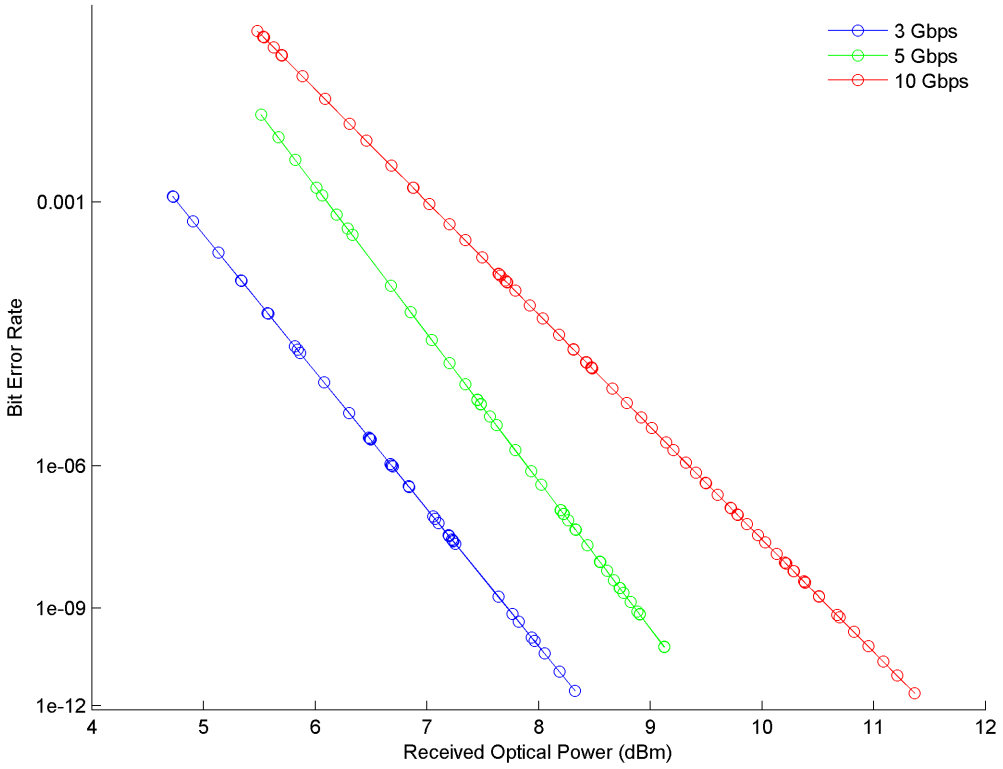


Figure 4.38: Power-penalty curves for the 1000-µm PP MZM at 3 Gbps, 5 Gbps and 10 Gbps using an end-to-end connection, taken from the cross port at 1550 nm.

to the fact that the 1000-µm PP MZM is not bandwidth limited, but is limited in ER by the 2-V_{pp} driving signal. When reducing the active region length by half, the half-wave voltage doubles, and so the fact that both the 1000-µm and 2000-µm PP MZMs were tested using the same driving voltage is disadvantageous to the 1000-µm PP MZM, and shows in the results of this test.

At a BER of 1×10^{-9}, the TWE MZM displays a power-penalty of 1.33 dBm between 3 Gbps and 5 Gbps, and a power-penalty of 3.38 dBm between 3 Gbps and 10 Gbps. When comparing the power-penalty curves of the two PP MZMs (Figures 4.38 and 4.40) and the

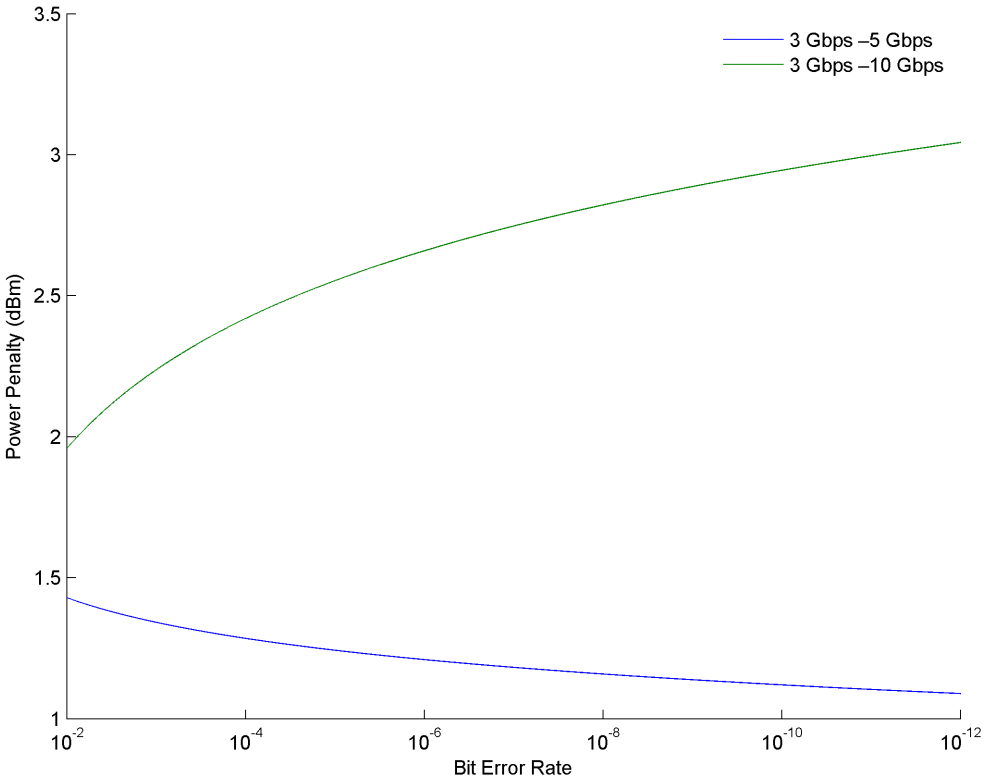


Figure 4.39: Power-penalty measurements for the 1000-μm PP MZM for 5 Gbps and 10 Gbps compared to power at 3 Gbps using an end-to-end connection, taken from the cross port at 1550 nm.

power-penalty curve of the TWE MZM in Figure 4.42, significantly less optical power is required at the receiver when transmitting with the TWE MZM, as opposed to transmitting with either of the PP MZMs. In fact, the TWE MZM requires only slightly more power at the receiver at 10 Gbps than the 2000-μm PP MZM does at 3 Gbps, and slightly less power than the 1000-μm PP MZM does at 3 Gbps. This result is surprising, as it was assumed the efficiency advantages of the push-pull modulation scheme over the single-arm modulation scheme would result in lower power-penalty for the PP MZMs at low bandwidths, when both the PP and TWE MZMs are not bandwidth limited. This result

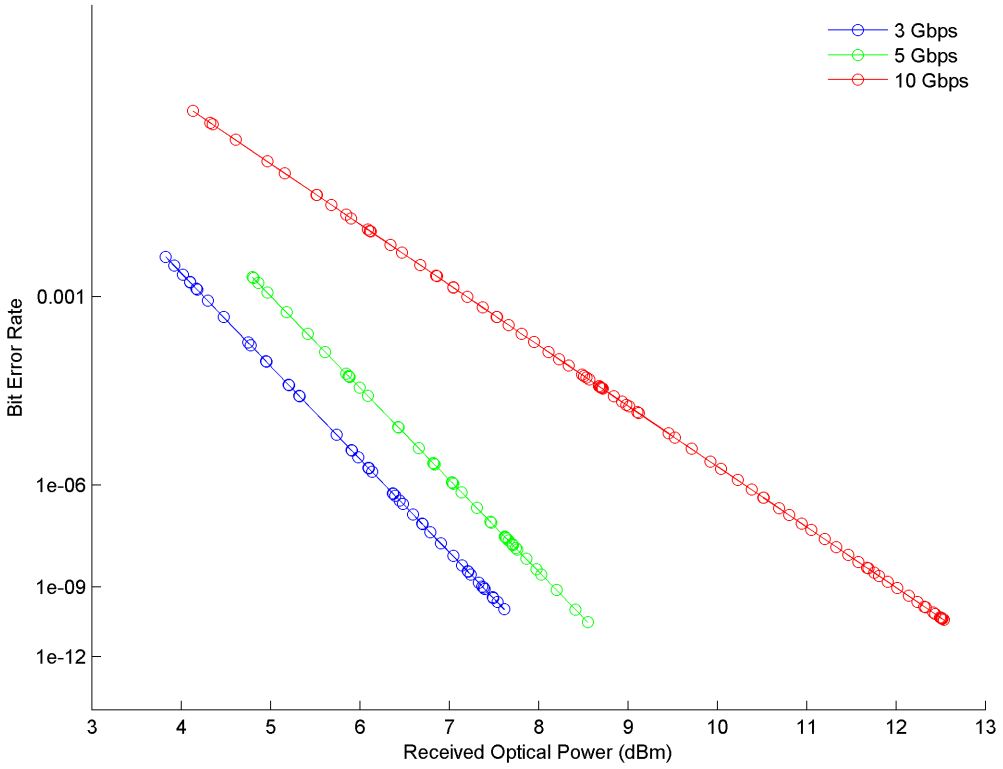


Figure 4.40: Power-penalty curves for the 2000-μm PP MZM at 3 Gbps, 5 Gbps and 10 Gbps using an end-to-end connection, taken from the straight port at 1550 nm.

speaks to the importance of the velocity matching and impedance matching conditions, even at the relatively low speed of 3 Gbps.

When the power-penalty curves for the 2000 nm PP MZM at 3 Gbps and 5 Gbps were nearly one atop the other, it was stated that this was evidence of the fact that the device was not bandwidth limited at 5 Gbps. The 1500-μm TWE MZM certainly isn't bandwidth limited at 5 Gbps; however, it may be limited in ER, as was the case with the 1000 nm PP MZM. The limitation in ER then arises from the use of the straight port, the reduction in active region length, and to a minor degree, the change to a single-arm modulation scheme.

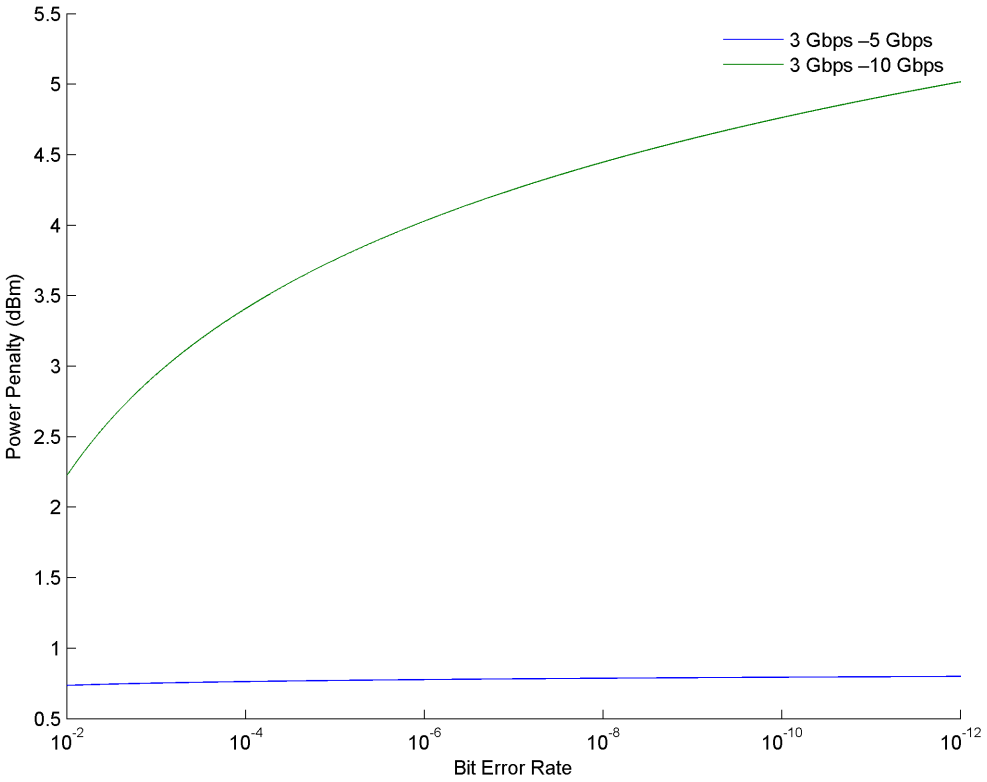


Figure 4.41: Power-penalty measurements for the 2000-μm PP MZM for 5 Gbps and 10 Gbps compared to power at 3 Gbps using an end-to-end connection, taken from the straight port at 1550 nm.

Once more, the power-penalty between curves is not constant, but the trend is that the power-penalty between the 3-10 Gbps curve is larger than that of the 3-5 Gbps curve. This is the same as was the case for the PP MZM, and is not at all shocking. Though the TWE MZM does require less absolute power at the receiver than either of the PP MZMs at any given speed, the power penalties for the TWE are greater than either PP for the 3-5 Gbps case, and greater than the 3-10 Gbps case for the 1000 nm PP MZM, but within 1σ of the 2000 nm PP MZM case.

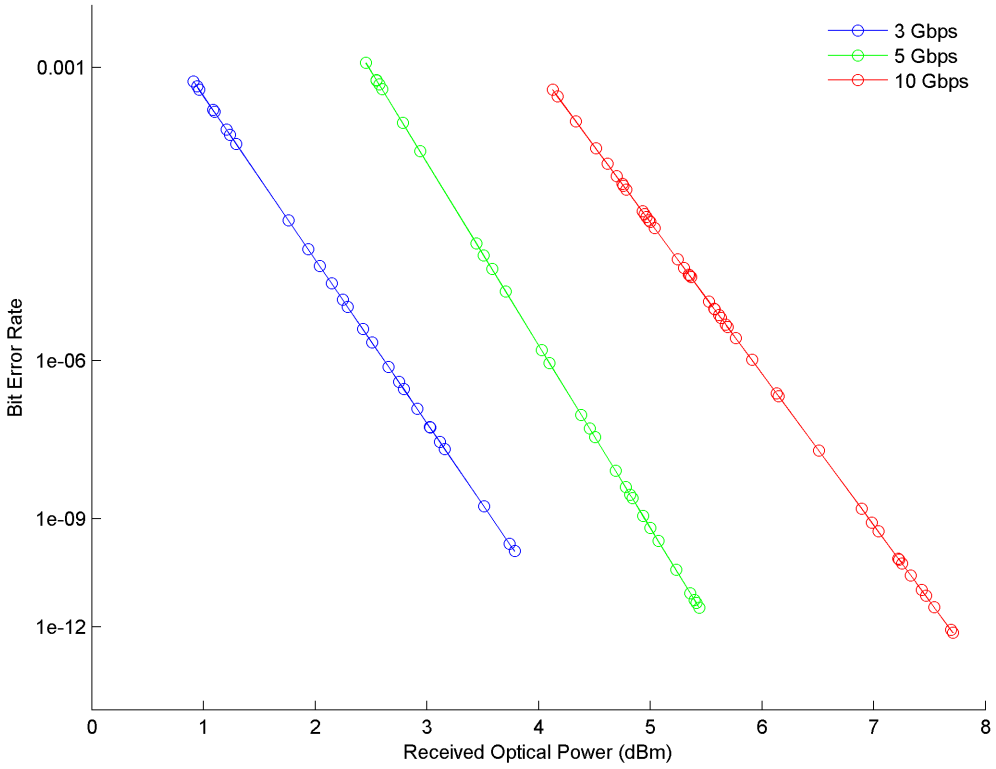


Figure 4.42: Power-penalty measurements for the 1500-μm TWE MZM at 3 Gbps, 5 Gbps and 10 Gbps using an end-to-end connection, taken from the straight port at 1550 nm.

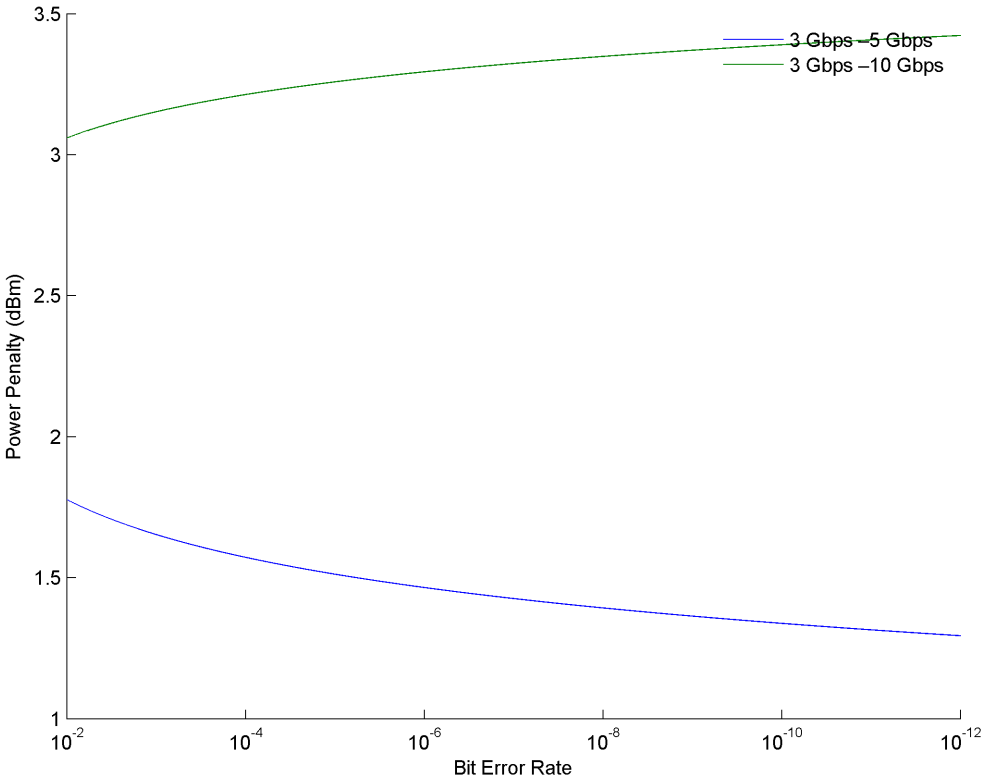


Figure 4.43: Power-penalty measurements for the 1500-μm TWE MZM for 5 Gbps and 10 Gbps compared to power at 3 Gbps using an end-to-end connection, taken from the straight port at 1550 nm.

4.5 Polarization Dependence

As previously noted in the section on DC characterization, there is a ripple that appears in the transmission response of the device, and that ripple seems to vary with the operating wavelength. The MZM was designed to operate over a wide range of wavelengths, as adiabatic couplers have a large optical bandwidth, and the active region of the MZM does not have a large wavelength dependence. The design then, apparently conflicts with the results from the DC characterization. In truth, the ripple is not a function of the operating wavelength, but is in fact a function of the input optical polarization to the MZM. The polarization dependence of the MZM arises due to two components with distinct polarization dependencies: the adiabatic couplers, and the birefringence of the silicon wire waveguide.

Because the adiabatic couplers are polarization dependent, and only display a constant splitting ratio for the TE mode, when the TE mode is not aligned with the principal axes of the silicon waveguide, the power is split unevenly between the arms of the MZM waveguide. A condition for completely constructive or destructive interference is that the power in the two arms of the MZM be equal before recombination.

Additionally, it is a well known fact that the internal stresses in a waveguide can turn a normally uniaxial crystal into a birefringent crystal. Within the birefringent waveguide, two effective indices are set-up, one for the TE mode, and one for the TM mode. Each effective index is modified independently by the plasma-dispersion effect. With the two modes seeing two different effective indices, it only makes sense that at certain wavelengths, the two modes will arrive at the interaction region after the active region, with a slight relative phase shift between the two. As such, the two modes may destructively interfere. At other wavelengths, the two modes will have no apparent phase shift between them, and as such, constructively interfere. It is noted in [63], that in demanding telecommunications applications such as in optical data modulators, a level of birefringence less than 10^{-4} (less

than 10^{-4} difference between effective indices for the two modes) is often necessary for good performance. Though it may not be easily apparent, Section 5.1 demonstrates the effects of small changes in the effective index, on the order of 10^{-4}.

To confirm that the ripple is indeed a function of polarization and not the operating wavelength, the polarization was tuned using a Newport Fiber Polarization Controller. As mentioned earlier, the method by which the polarization controller operates is not accurate, and leads to an inability to present discrete results for the polarization dependence of the MZM. It is however possible to show how the ripple depth and period vary with different input polarizations through a number of contour plots, shown in Figures 4.44 and 4.45. To determine the proper input polarization, the operating wavelength was swept with the power meter measuring the output power of the MZM connected to an oscilloscope. When the input polarization is roughly aligned with the principal axes of the silicon waveguide, the output power should be relatively flat across all wavelengths, and this would appear as a flat voltage response on the oscilloscope. The response is only relatively flat, and not entirely flat, as the gain spectrum of the laser is not flat, and AutoLab was not used to tune the power of the laser. In practical applications, the MZMs would either be integrated into larger photonics chips, in which case the optical input is likely from another waveguide on-chip, or the MZM would be a stand-alone device that is individually packaged, and the optical input is fiber-coupled. In the former case, the likelihood is that the optical input polarization is already properly aligned, as most devices are optimized to interact with the TE mode. In the latter case, the onus for alignment of the fiber's output polarization is on the fabrication facility, which would then epoxy or weld the fiber into place. In either case, the burden of aligning the input polarization is not on the end user, but rather on the device manufacturer. Finally, it is also possible to integrate a polarizing element prior to the adiabatic coupler (which would then increase optical insertion loss), or simply operate with and accept the presence of the ripple, which affects ER.

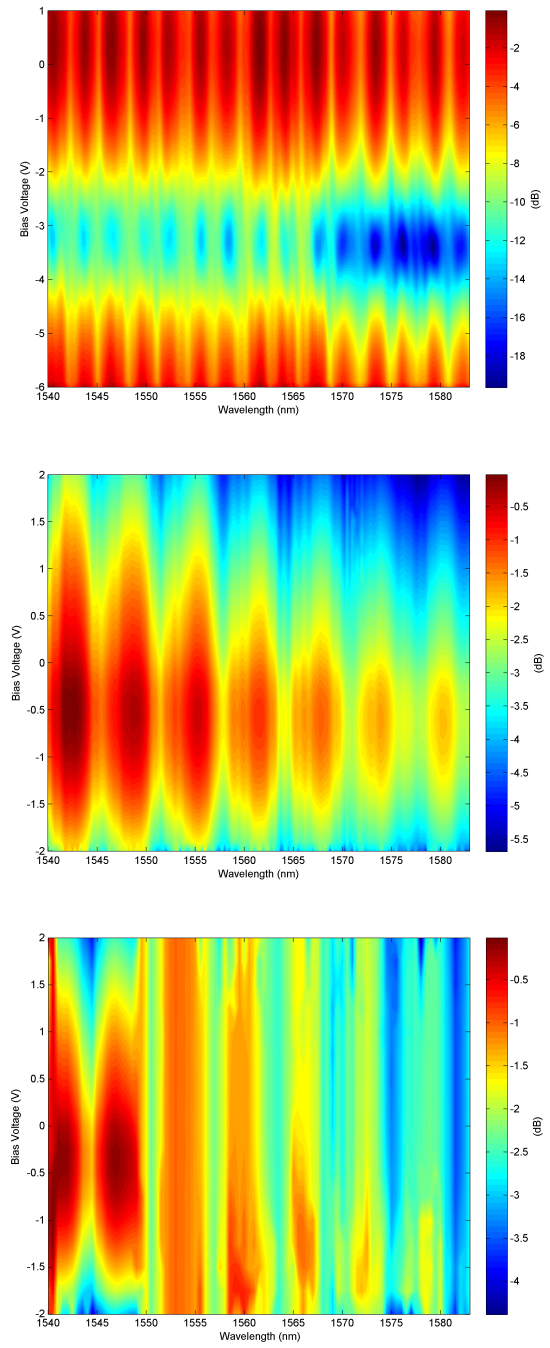


Figure 4.44: DC characterization contour plots for the 2000-μm PP MZM using a variety of input optical polarizations, which demonstrate the device's sensitivity to input optical polarization as a result of the adiabatic couplers.

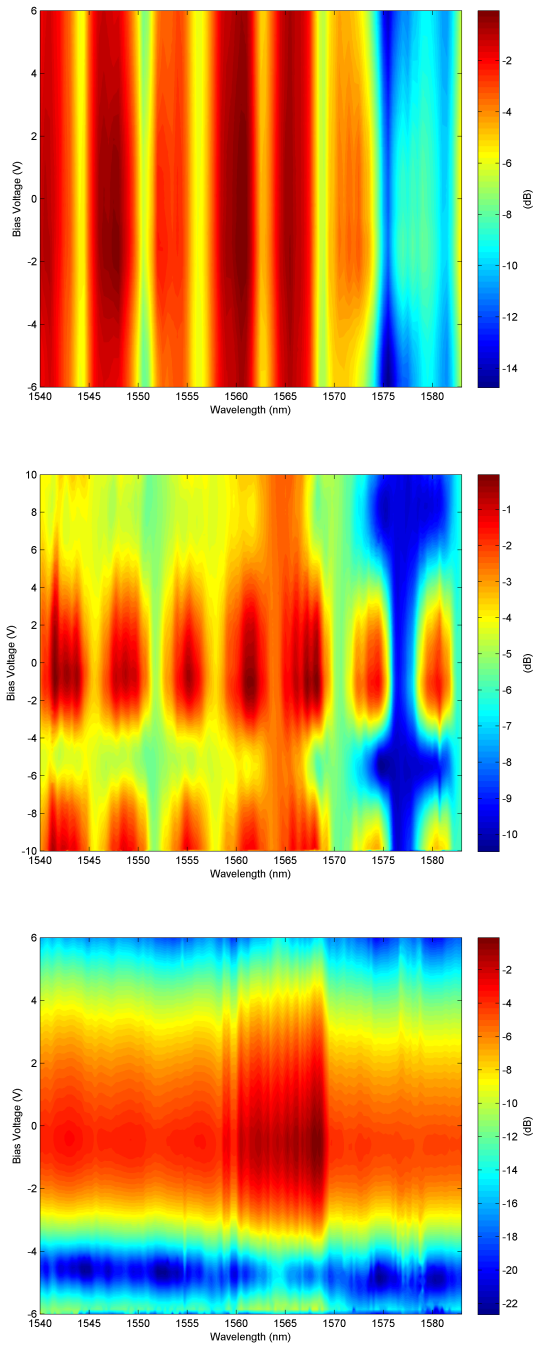


Figure 4.45: DC characterization contour plots for the 2000-μm PP MZM using a variety of input optical polarizations, which demonstrate the device's sensitivity to input optical polarization as a result of the adiabatic couplers. The optical polarization is closely aligned in the bottom plot.

4.6 Effect of Termination Resistance

Although not mentioned previously, three different TWE MZMs were fabricated with effective active region lengths of 1000-μm, that differed only in termination design, and thus termination resistance. The termination designs are denoted T1, T2, and T3, and the measured termination resistances can be found in Table 4.14. Note that the 1500-μm TWE MZM tested used termination design one. The S_{11} responses are presented in Figures 4.46, 4.48, and 4.50 for termination designs one, two and three, respectively. The S_{21} responses are presented in Figures 4.47, 4.49, and 4.51 for termination designs one, two and three, respectively.

As can be seen, although termination design one has the absolute largest minimum return loss, it actually has the smallest 3 dB bandwidth. Looking at termination design three, an improvement of 3.78 GHz is made in the mean 3 dB bandwidth merely by changing the termination design. Within only minor variations of the minimum return loss, there is no correlation between minimum return loss and 3 dB bandwidth. Instead, it is more important that there be more loss at low frequencies than strictly increasing the absolute minimum return loss. This is because PRBS test patterns have more power at low frequencies than at high frequencies (see Section 5.4), so attenuating the reflection of low frequency components is more important than attenuating the entire spectrum. Comparing Figures 4.46 and 4.50 verifies the previous statement, as termination design three has an attenuation of roughly 40 dB at DC, versus termination design one or two's attenuation of roughly 15 dB.

Table 4.14: Measured termination resistances for the different termination designs used in the 1000-µm TWE MZMs. The minimum return loss is taken from the S_{11} response, and the mean 3 dB bandwidth is the arithmetic mean of the 3 dB bandwidths for the straight and cross ports.

Termination	Resistance (Ω)	Min Return Loss (dB)	μ, 3 dB Bandwidth (GHz)
T1	61.50	9.06	8.95
T2	52.15	7.57	10.57
T3	40.82	8.67	12.73

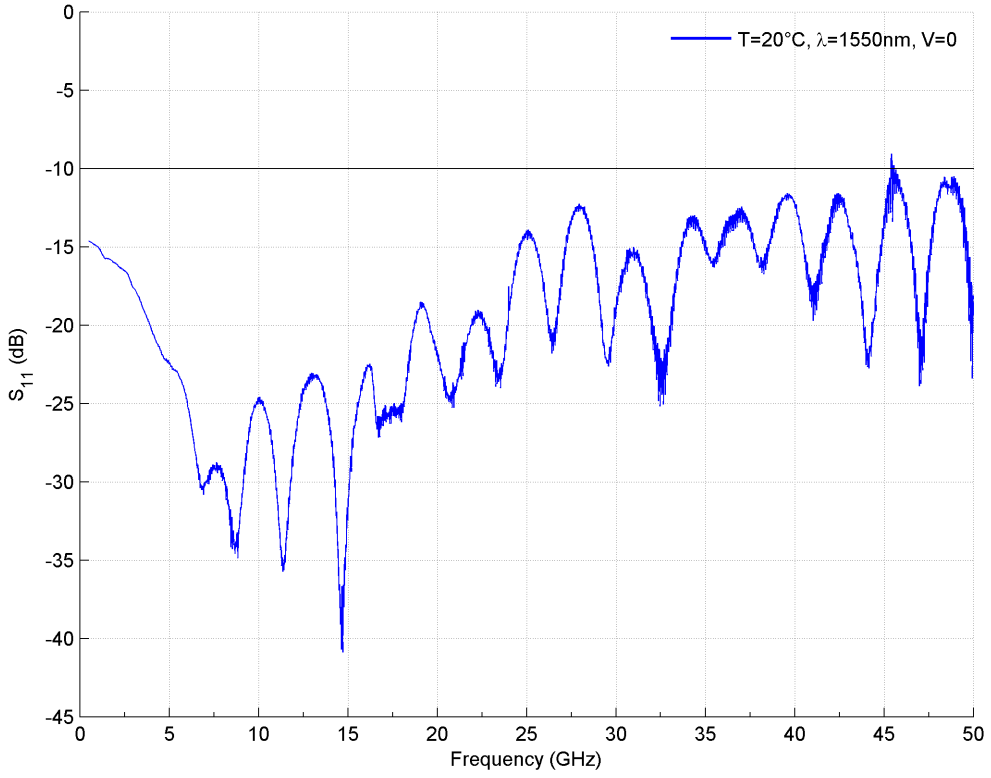
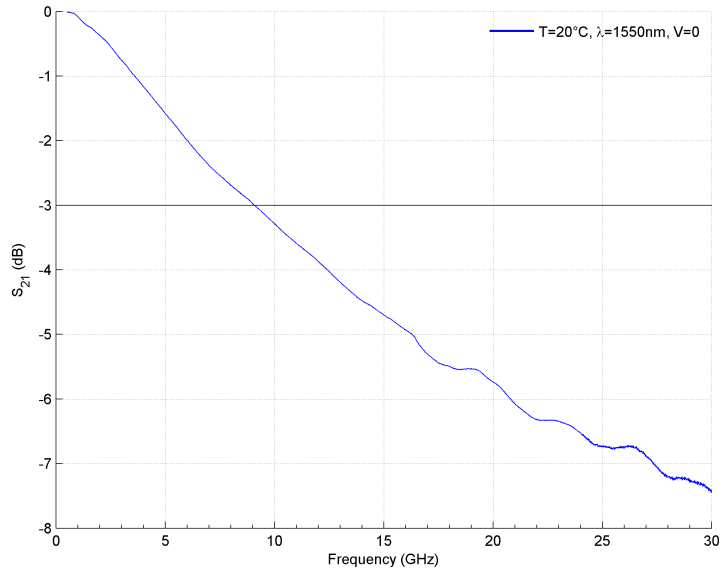
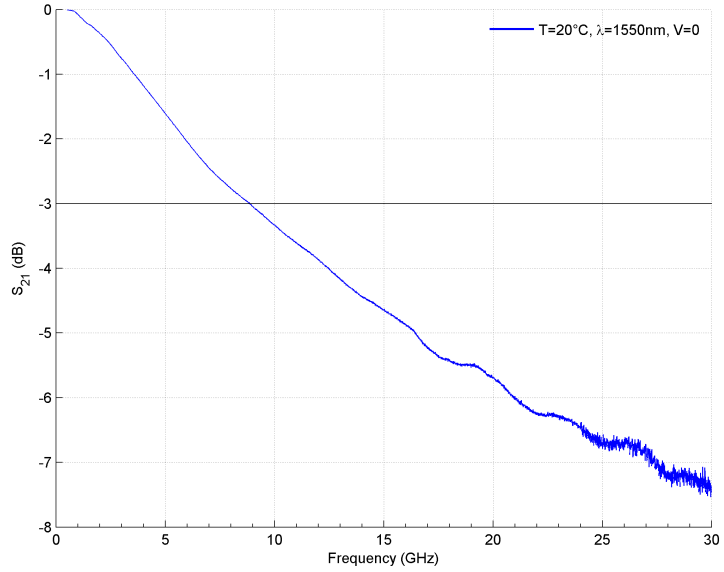


Figure 4.46: S_{11} for the 1000-µm TWE MZM, with termination design one, and results in a minimum return loss of 9.06 dB.



(a) Straight Port



(b) Cross Port

Figure 4.47: S_{21} for the 1000-µm TWE MZM, with termination design one, at 1550 nm. The device performance is $f_{3\,\mathrm{dB}}$ =9.07 GHz/8.82 GHz and $f_{6\,\mathrm{dB}}$ = 20.79 GHz/20.84 GHz (straight/cross ports).

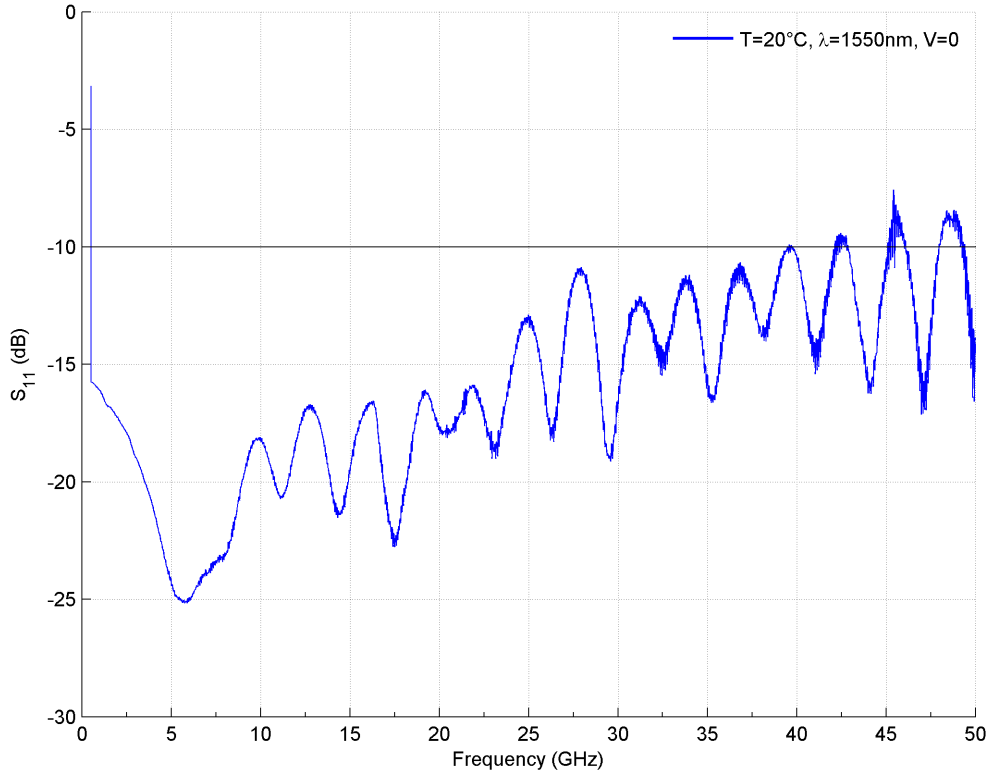
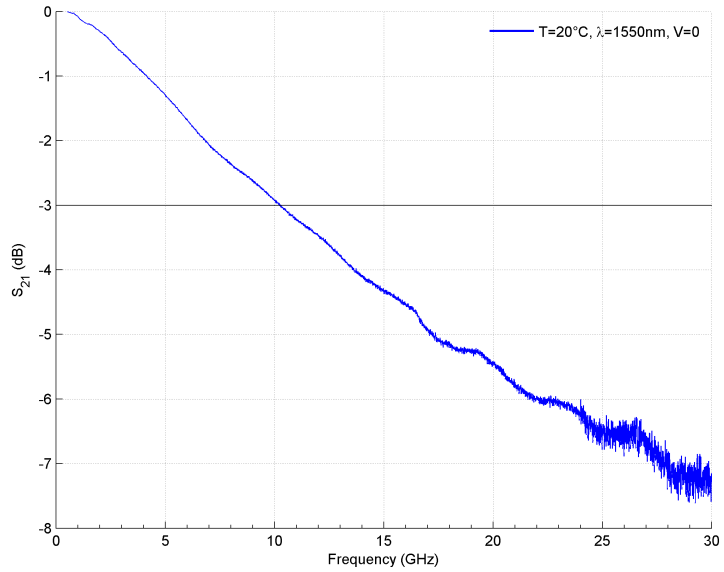
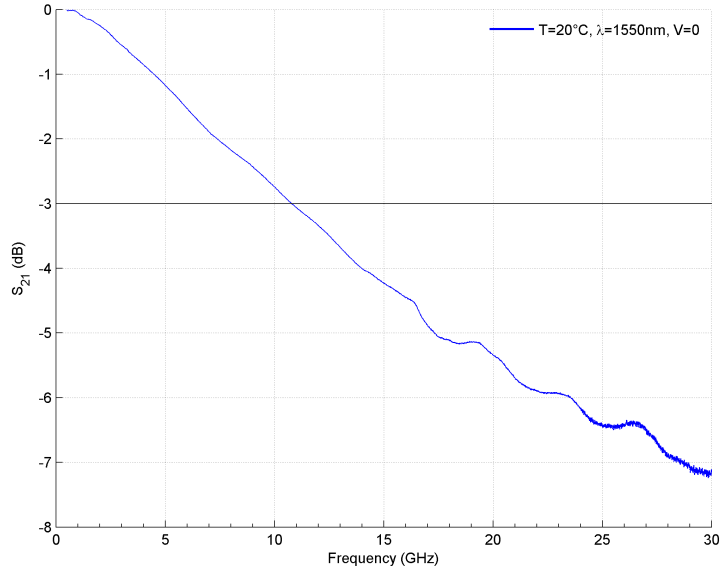


Figure 4.48: S_{11} for the 1000-μm TWE MZM, with termination design two, and results in a minimum return loss of 7.57 dB.



(a) Straight Port



(b) Cross Port

Figure 4.49: S_{21} for the 1000-µm TWE MZM, with termination design two, at 1550 nm. The device performance is $f_{3\,\mathrm{dB}}$ =10.37 GHz/10.78 GHz and $f_{6\,\mathrm{dB}}$ = 22.29 GHz/23.58 GHz (straight/cross ports).

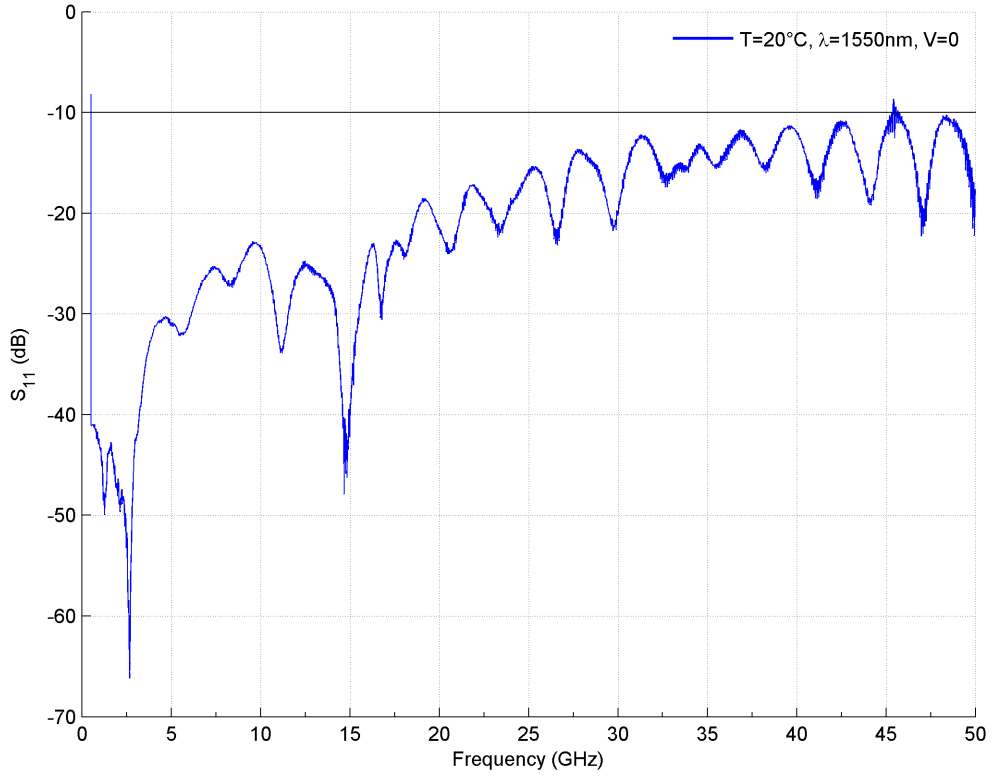
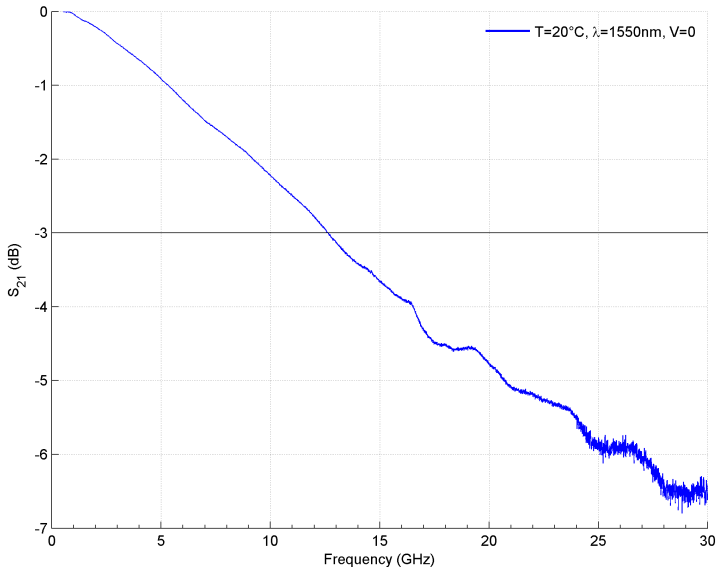
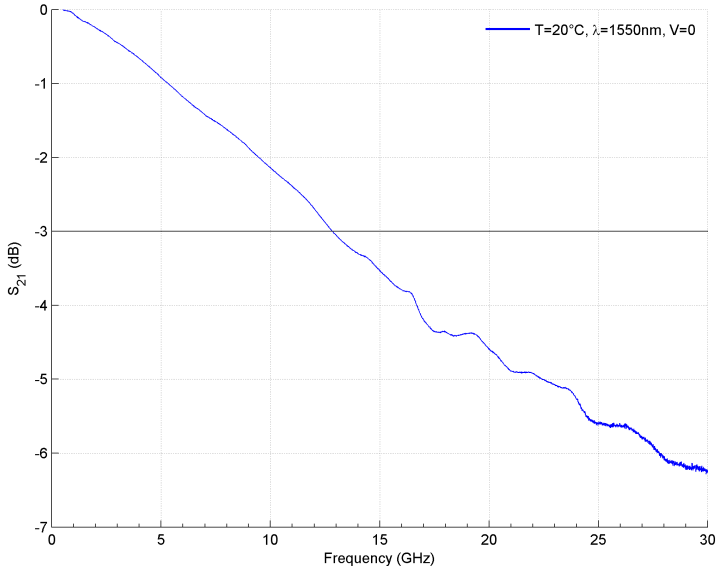


Figure 4.50: S_{11} for the 1000-μm TWE MZM, with termination design three, and results in a minimum return loss of 8.67 dB.



(a) Straight Port



(b) Cross Port

Figure 4.51: S_{21} for the 1000-µm TWE MZM, with termination design two, at 1550 nm. The device performance is $f_{3\,\mathrm{dB}}$ =12.62 GHz/12.85 GHz and $f_{6\,\mathrm{dB}}$ = 25.17 GHz/27.88 GHz (straight/cross ports).

V. Modeling

THERE are a number of simulations than can be run and models constructed. Typically, finite-element models are more precise than analytical models; however, analytical models are generally easier to produce and less time consuming, and advantageously, provide general insight into the salient parameters and characteristics that define the results and operation of the model. The computational and algorithmic complexity of finite-element models generally means that specialized, commercial software is used in the production of such models. For the most part, finite-element modeling has been left for future work, and analytical models/simulations for the device impedance, device index, bandwidth, PRBS patterns, and eye diagrams were produced during the course of this work – the exception being the optical waveguide modes.

5.1 Optical Waveguide Mode Modeling

Using the waveguide mode solver available in [64], the optical mode carried in the ridge waveguide is shown in Figures 5.1 and 5.2. The ridge waveguide is 560 nm wide and 240 nm high, with a lower cladding of silicon dioxide (SiO_2) at 3 μm high, a core of crystalline silicon (Si) at 240 nm high, and an upper cladding of tetraethyl orthosilicate ($Si(OC_2H_5)_4$, but commonly abbreviated TEOS) at roughly 750 nm high. At a wavelength of 1550 nm, the index of refraction for SiO_2 is 1.528, for Si is 3.478, and for TEOS is 1.382. As can be seen, the mode is largely confined to the waveguide. The importance of showing the optical mode is not only to demonstrate the confinement of the mode to ensure low propagation loss, but more importantly to show the overlap between the carrier concentration change and the optical mode. By simple inspection, it is evident that the choice to use a vertical p–n junction provides better overlap than the use of a horizontal p–n junction.

The waveguide mode solver also produces an effective index, n_{eff}, which equals 2.595. With that, and the knowledge that complete destructive interference is achieved at a relative phase shift of π radians between the modes in the two arms of the waveguide, the necessary change in effective index can be calculated. The necessary change in optical path length (OPL) is half the wavelength in the medium, so the necessary change in n_{eff} (if using single-arm modulation) is then expressed as

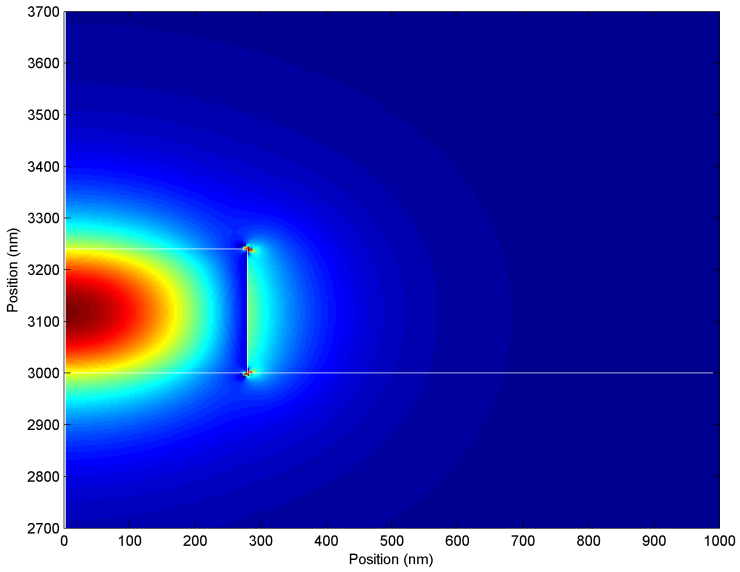
$$\Delta n_{\text{eff}} = \frac{\lambda_0}{2n_{\text{eff}}} \cdot L \tag{5.1}$$

where λ_0 is the free-space wavelength of the electromagnetic radiation, and L is the length of the active region of the MZM. Table 5.1 lists the results of this calculation for each MZM.

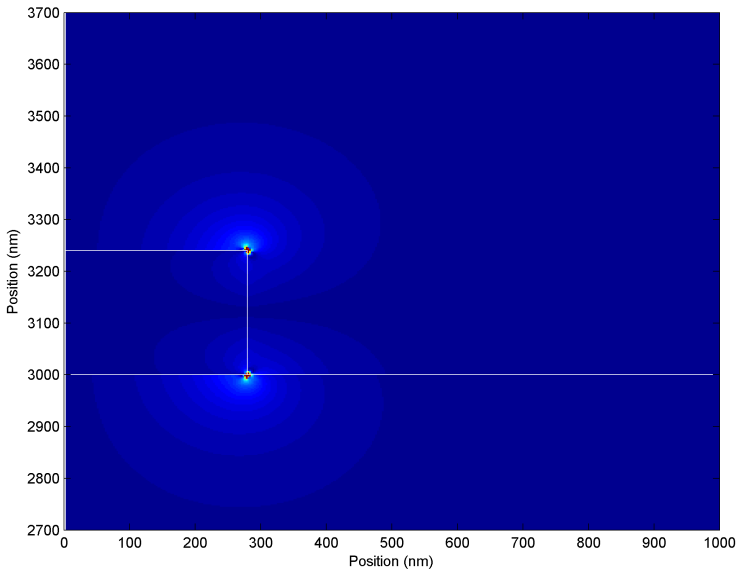
Table 5.1: The necessary change in effective index, Δn_{eff}, for each MZM in order to achieve a π radian phase shift.

Device	Δn_{eff}
1000 PP	2.987×10^{-4}
2000 PP	1.494×10^{-4}
1500 TWE	1.991×10^{-4}

It may be beneficial to make the dopant density asymmetric, or change the implantation energies/penetration dept,h such that the optical mode maximizes the modal overlap with the p-type material, as per Section 2.1.1. Doing so would result in lower values for V_π, and decrease overall FCA.

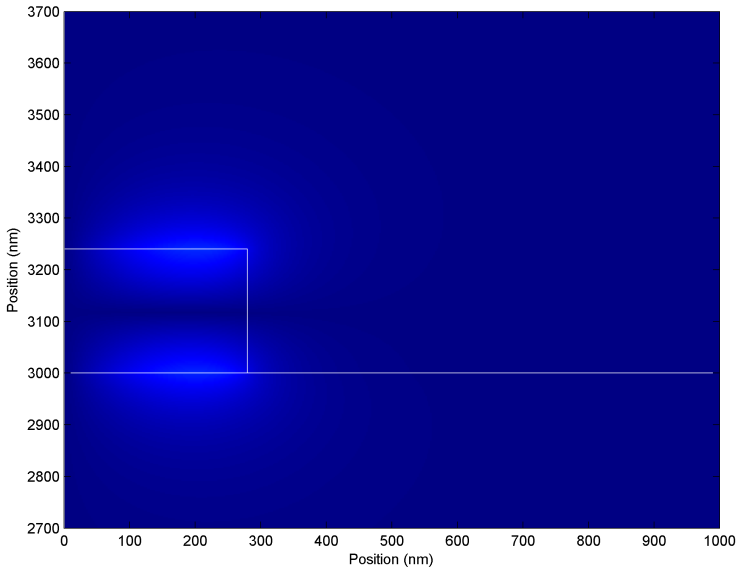


(a) E_x (Fundamental Mode)

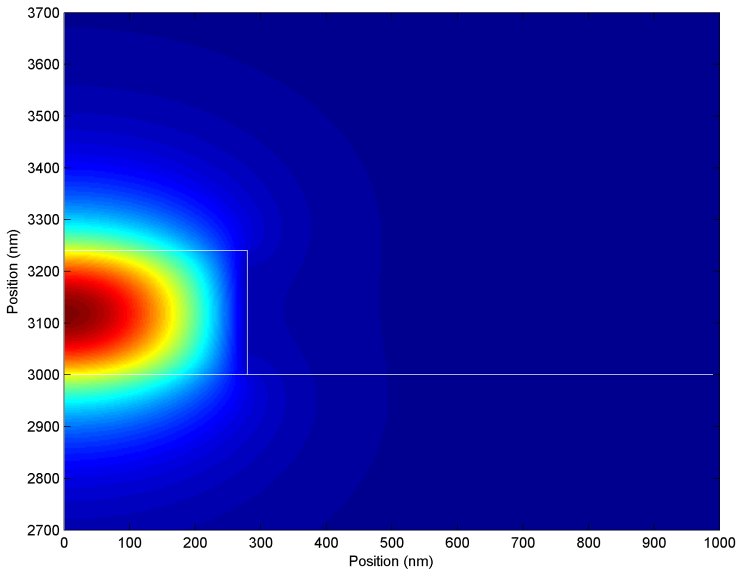


(b) E_y (Fundamental Mode)

Figure 5.1: Optical mode carried by the ridge waveguides used in the MZMs, with a width of 560 nm, and a height of 240 nm. The geometry of the ridge waveguide is shown by the white lines. Note that only half of the waveguide is shown, namely the right half of the waveguide, and that the distribution is symmetric about the y-axis.



(a) H_x (Fundamental Mode)



(b) H_y (Fundamental Mode)

Figure 5.2: Optical mode carried by the ridge waveguides used in the MZMs, with a width of 560 nm, and a height of 240 nm. The geometry of the ridge waveguide is shown by the white lines. Note that only half of the waveguide is shown, namely the right half of the waveguide, and that the distribution is symmetric about the y-axis.

5.2 Impedance and Microwave Index Modeling

The importance of the traveling-wave electrode is that it meets the impedance and velocity matching requirements established in Section 2.1.5. The PP device used lumped-element electrodes, and is thus unable to achieve the velocity matching requirement. As noted earlier, the traveling-wave electrode is designed as a segmented, capacitively-loaded traveling-wave electrode, where the periodic sections of the active region are connected to the transmission line via metallic bridges, and appear as a capacitive load. The unloaded transmission line is simply the coplanar strip waveguide (CPS), and is designed to have an impedance higher than 50 Ω, as the capacitive load decreases the impedance of the device. The microwave velocity of the unloaded transmission line is faster than the optical velocity, and the capacitive load decreases the microwave velocity (increases the microwave index).

The unloaded impedance and microwave index can be calculated using (2.46) and (2.47), respectively. Note that for these equations the capacitance and inductance are specified per unit length. Using the methods in [65], a rough value for capacitance and inductance per unit length can be calculated, as well as the unloaded impedance, either using a lossy or lossless model for the transmission line. For microwave structures, the assumption of a lossless transmission line is normally valid — the lossy and lossless impedances are generally very close. The methods used in [65] only produce an approximate value for capacitance, inductance, and impedance, because the microwave structure assumed in [65] is less complex than the one used in these MZMs. The structure assumed in [65] is a CPS waveguide configuration between slab dielectric waveguides, with two dielectric slabs below the waveguide, and one above. The structure used in this MZM has been shown previously (Figure 2.14), and rather than using dielectric slabs, uses a Ti/TiN/AlCu/TiN stack, a significantly more complex structure [54].

The unloaded impedance in [54] is noted as being 95 Ω, with a microwave index of 2.3. The methods in [65] result in an unloaded lossless impedance of 219.9005 Ω, an unloaded

lossy impedance of 219.9021 Ω, and a microwave velocity of 1.0557. The disagreement between values is in part due to the difference in structures, and also the inaccuracies with analytical modeling. As can be seen, the assumption of a lossless transmission line is valid. After calculating the values for the loaded transmission line using the methods in [66], the impedance (assumed lossless) decreases to 199.14 Ω, and the microwave index increases to 2.567. Despite the disparity in calculated and modeled impedances, the microwave index increases to a value very nearly identical to the optical effective index, which would indicate that the velocity matching condition has been met.

5.3 Bandwidth Modeling

The normalized frequency response of a segmented traveling-wave electrode can be calculated using (2.54), which uses a transmission matrix to calculate the voltages at each metal bridge, and thus the voltage across each periodic section of the active region. Fundamentally, the calculation also relies upon parameter values for resistance, inductance, and capacitance: R_a, L_a, and C_a. The normalized frequency response also assumes a lossless transmission line. Regrettably, the frequency response of the TWE MZM could not be reproduced using (2.54), although the frequency response in [52] could be reproduced. Even when varying R_a, L_a, and C_a, the S_{21} response of the TWE MZM could not be reproduced at any bias voltage; the reason is unknown.

5.4 PRBS Simulations

A PRBS pattern generator was also built using MATLAB, which conforms to the definition of a PRBS pattern as given in Section 3.1.2. A relatively simple method of generating a pseudorandom bit sequence would be to generate a pseudorandom number between one and zero, and then round; however, doing so does not produce a pattern that conforms to the definition of a PRBS pattern. Instead, every possible n-bit word, with the exception of the word with all zeros, is concatenated into a pattern, and then expanded in

time with multiple samples during each bit using a Kronecker tensor product. The resulting PRBS pattern can then be used for further simulations or models, with any specification for n, or at any clock frequency. Unfortunately, the algorithm takes both exponential time and exponential memory, neither of which is desirable in the the case of simulations (in Big O notation, $O(c^n), c > 1$). As such, simulations with $n \geq 11$ are intractable.

As noted earlier in Section 3.1.2, the envelope of the Power Spectral Density (PSD) for the PRBS pattern follows a sinc^2 function, with spectral nulls at integer multiples of the clock frequency. The PSD of a PRBS-7 pattern is shown in Figure 5.3, and verifies the previous statement. Note that the PSD is not normalized in Figure 5.3, and strictly speaking does not follow a sinc^2 envelope unless it is normalized. Additionally, the PSD only follows a sinc^2 envelope if there is no DC offset to the signal, that is, if the signal is centered around 0. The clock frequency does not change the PSD of the pattern in any way other than to scale the PSD in the frequency spectrum, and notably, moves the location of spectral nulls to integer multiples of the clock frequency. This can be verified by looking at Figure 5.4, which shows the PSD of a PRBS-7 pattern at multiple frequencies.

The PRBS pattern specification, namely n in the PRBS pattern, controls the spacing between successive peaks in the frequency spectrum, as can be seen in Figure 5.5. As n increases, the spacing between spectral peaks decreases.

During BER tests, there was evidence to suggest that the first transmission errors encountered (errors at low BER, i.e. errors at a BER of 2×10^{-12}) were the result of intersymbol interference; that is to say, errors were encountered immediately after the transmission of a long sequence of ones or zeros. The presumed cause for this behavior was that during long sequences of ones and zeros, the signal's instantaneous spectral content shifted from high to low frequency content. The PRBS signal had to pass through a bias tee, which has a low frequency cutoff of 60 kHz, and it was assumed that the bias tee then

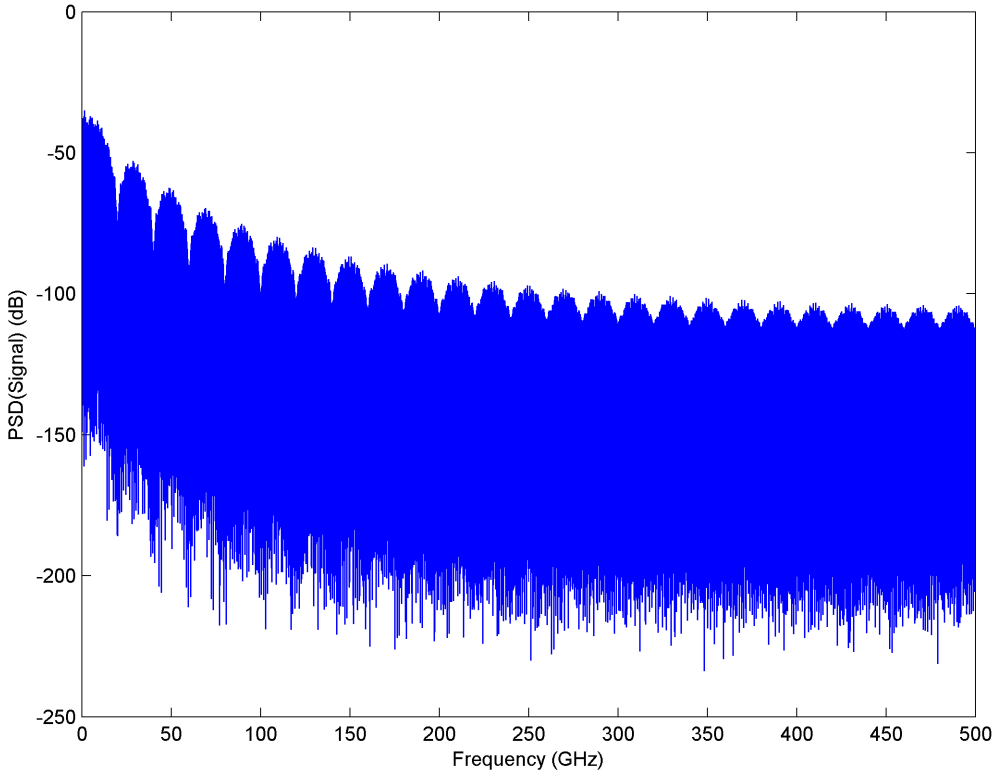


Figure 5.3: PSD for PRBS-7 patterns at a clock frequency of 20 GHz, showing the envelope of the PSD.

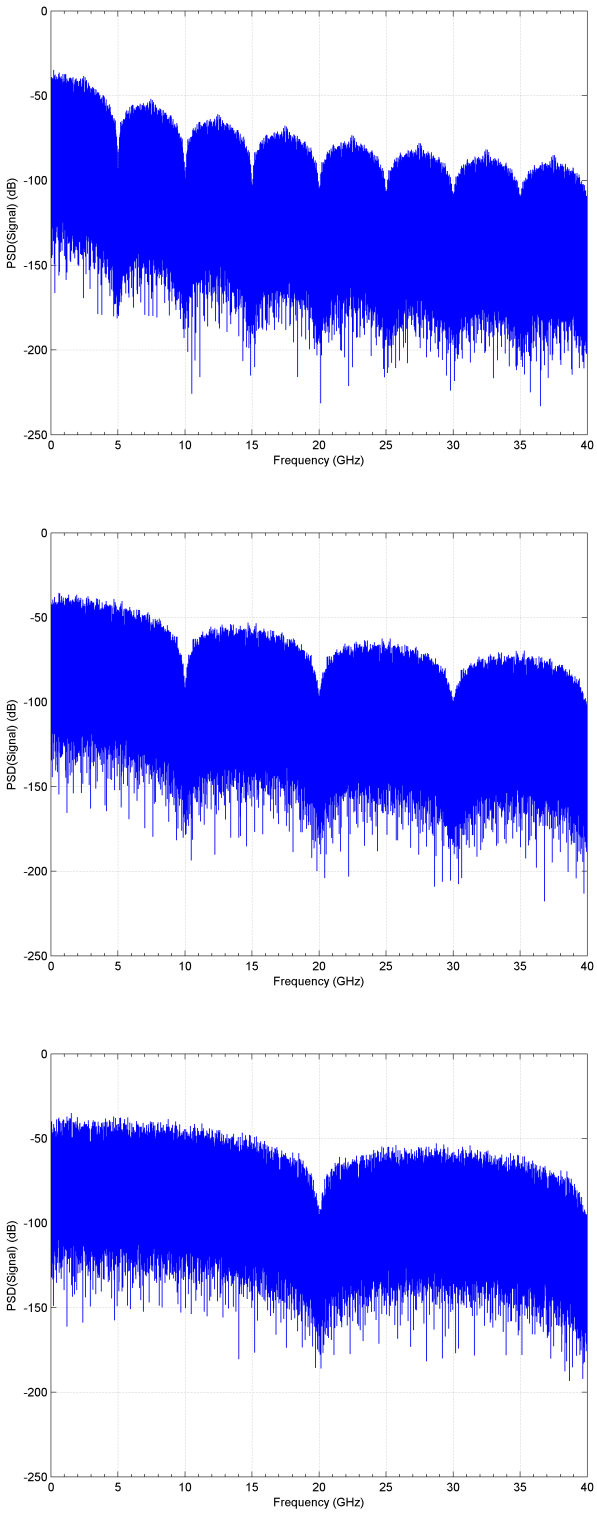


Figure 5.4: PSD for PRBS-7 patterns at clock frequencies of 5 GHz, 10 GHz and 20 GHz, from top to bottom. Note that only a small frequency subset of the PSD is shown, which is done to emphasize the location of the spectral nulls.

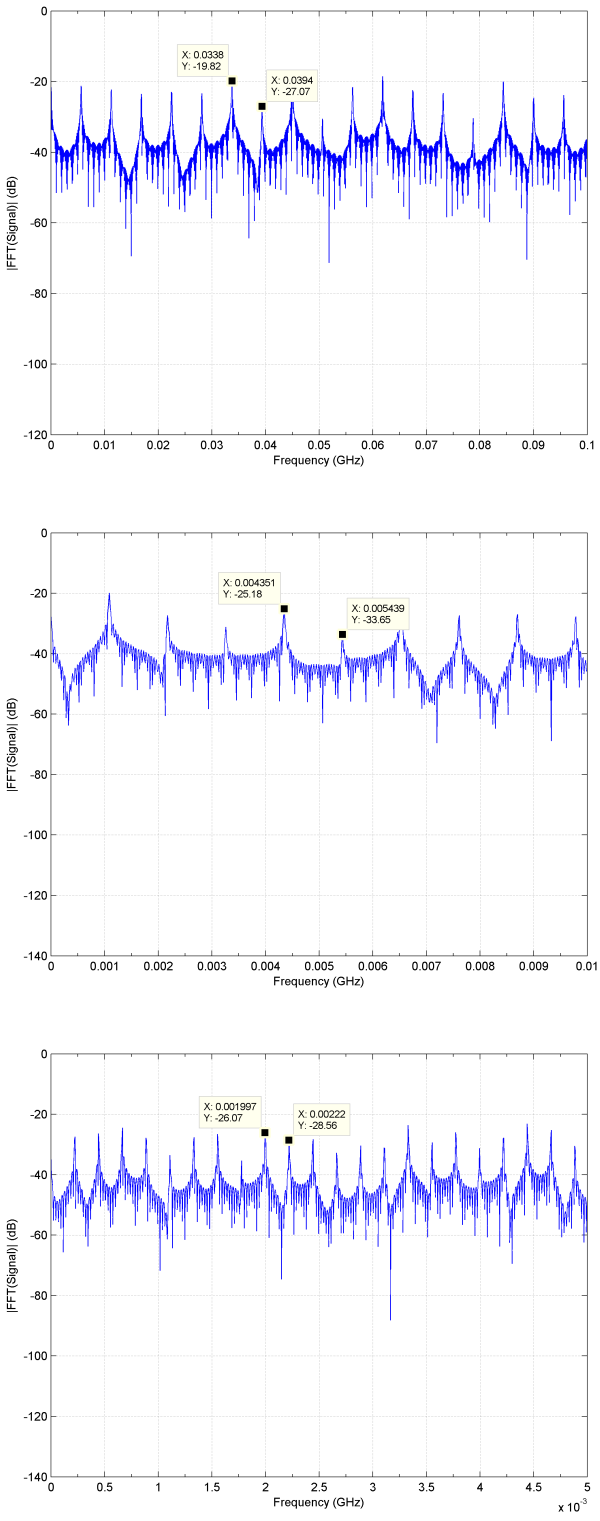


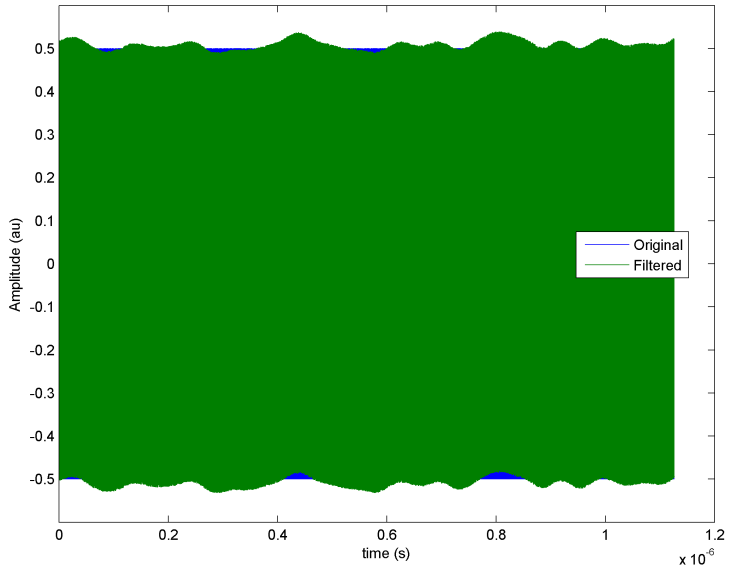
Figure 5.5: Fast Fourier Transform for PRBS-7, PRBS-9, and PRBS-11 patterns, at a clock frequency of 5 GHz, showing the spacing between successive peaks in the frequency spectrum. Spacings are 5.60 MHz, 1.09 MHz and 0.22 MHz, respectively.

blocked the low frequency content of the signal, which in turn caused the transmission error.

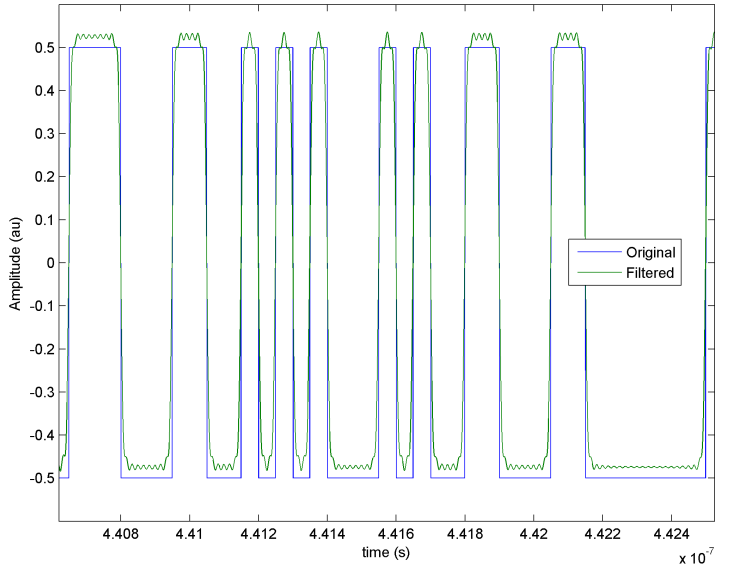
The PRBS signal simulator was then used to either confirm or deny that hypothesis. Figure 5.6 shows the effects of the low frequency cutoff of the bias tee on the signal. Figure 5.6a shows the entire PRBS-11 signal, the original in blue, and the signal after having traveled through the bias tee in green. The bias tee is simply implemented as a bandpass filter, with a stop band attenuation of at least 3 dB. Figure 5.6 is zoomed out such that it does not show individual transitions so that the effects on the signal can be seen. During consecutive periods where there are not a large number of transitions, the signal does drift as a result of an increase in the low frequency content and the filtering effects of the bias tee. Despite the fact that drift does occur, it is not significant, as can be seen in 5.6b, and should not negatively affect the performance of the MZM.

Even if the signal amplitude was not affected from a macroscopic perspective, there is still the possibility that the transition times can be affected. This is because a large portion of the high-frequency content in the signal is localized to the areas immediately surrounding the transitions. Thus, a long sequence of either zeros or ones could either drift up or down prior to the expected transition — a phenomena which would not be visible when the signal is zoomed out as it is in 5.6. As a result, a small time period of the entire signal is presented in 5.7. The ringing is a result of Gibb's phenomena caused by the high-frequency cutoff of the bias-tee, and is damped by a moving average, which is used to slow transitions and reflects the non-ideal nature of test equipment (lest the transitions have a step-like nature without the presence of the bias-tee). Irrespective of the ringing, it can be seen that despite the successive transmission of nine bits, all with a value of one, there is no drift in the signal.

This may not necessarily be the case with longer length patterns, say for example with a PRBS-31 pattern, but the exponential time and memory requirements make simulation of



(a) Overall signal, which displays some drift during periods where there is an increased amount of low-frequency content in the signal.



(b) Small subset of the overall signal where there is drift in the signal.

Figure 5.6: PRBS-11 pattern at a clock frequency of 20 GHz, demonstrating the effects of the bias tee's bandwidth on the signal, implemented as a bandpass filter with a passband of 60 kHz - 60 GHz.

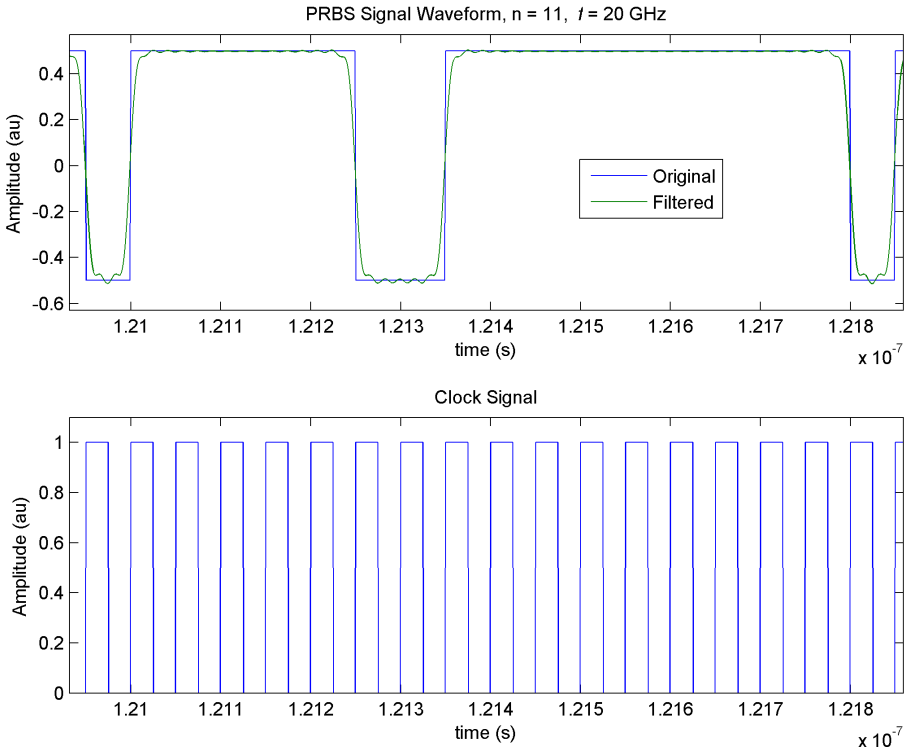


Figure 5.7: Small subset of the overall signal from Figure 5.6 where there is no drift, but a long sequence of ones, which would result in an increase in the low-frequency content of the signal. The top plot shows the PRBS signal, and the bottom plot the clock signal.

a PRBS-31 pattern infeasible. Additionally, the ISI was seen during experimentation with a PRBS-7 pattern. As such, the bias-tee can be ruled out as the cause of ISI. It seems that the most likely source of ISI in this case is thermal effects due to localized junction heating, which causes a slow transition after a long sequence of ones or zeros.

5.5 Eye Diagram Simulations

Linear Time-Invariant (LTI) system theory can be used to generate eye diagram simulations from the S_{21} response in conjunction with the output of the PRBS pattern generator. There are, however, some noted problems with using LTI system theory with these MZMs: primarily, one must assume that the modulator satisfies the requirements of an LTI system. While the performance of the modulator is time-invariant, the performance of the modulator is not linear. This is easily verified by looking at any number of figures in the DC characterization section. The transmission of the modulator is roughly linear over a small subset of voltages. Beyond that small subset though, the behavior is non-linear. Additionally, device physics not reflected in a simple DC characterization, would also drive device performance in a non-linear manner during high-speed experimentation. As such, the assumption of homogeneity can only be made if operating within a small subset of voltages, and the assumption of superposition can only be made if the amplitudes of the input signals are sufficiently small.

Despite this, the eye diagram simulations are implemented using LTI system theory by taking the S_{21} response (frequency response), Fourier transforming that response to get the impulse response of the MZM, and then convolving the impulse response of the MZM with the PRBS signal to get the resultant transmission of the MZM. In practice though, the convolution theorem was used in lieu of direct convolution, which is more computationally expensive than simply Fourier transforming the PRBS signal, and then inverse Fourier transforming the product of the transformed PRBS signal and the S_{21} response of the MZM. Note that the S_{21} was first fit with a polynomial to both smooth the response, and to generate a frequency response with sufficient resolution for the convolution. The phase information from the S_{21} response was also incorporated in the calculations, rather than solely using the magnitude information. The phase response for the modulator was linear. By using a proper functional form for the magnitude response, rather than a polynomial, it would

be possible to characterize device performance beyond the largest frequency measured, 50 GHz.

As a result of the non-linear transmission behavior of the MZM, statistics such as ER and eye width or height, lose meaning during the course of these simulations. If the ER were low during simulation, one could simply increase the PRBS signal amplitude, and thereby increase ER. In practice, increasing the signal amplitude beyond V_π actually deteriorates performance. The eye diagrams produced here are only of value in the sense that they may provide some insight as to when the eye can be expected to begin closing during experimentation.

Figures 5.8, 5.9, and 5.10 show simulations for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs, respectively. Looking solely at these simulations, one might be tempted to think that within the assumptions mentioned above, these eye diagram simulations can be effectively used to estimate device performance. From Section 4.2, the 6 dB bandwidth of the 1000-μm PP MZM is roughly 20 GHz, and the 6 dB bandwidth of the 1500-μm TWE MZM is roughly 45 GHz. The first irregularity is that there is perhaps no difference between the performance of the 1000-μm PP MZM and the 2000-μm PP MZM, even at 20 GHz, where the 2000-μm PP MZM should be severely bandwidth limited.

Figure 5.11 makes it clear that these simulations are not accurate, as the simulations show eye diagrams that are wide open for both the 1000-μm and 2000-μm PP MZMs even at 50 GHz. In fact, the figures are not altogether different from the simulation of the 1500-μm TWE MZM at 50 GHz.

It is, however, interesting that the crossing time is not in the center of the eye, as that behavior is controlled by the bias point of the device and the non-linearity of the device transmission. These simulations assume homogeneity, and as such, one would expect the crossings to be vertically centered on the eye. This may be the result of incorporating the phase information in the eye diagram simulations. The same rationale can be given for

the rise and fall times, which are also affected by device physics not captured in the S_{21} response.

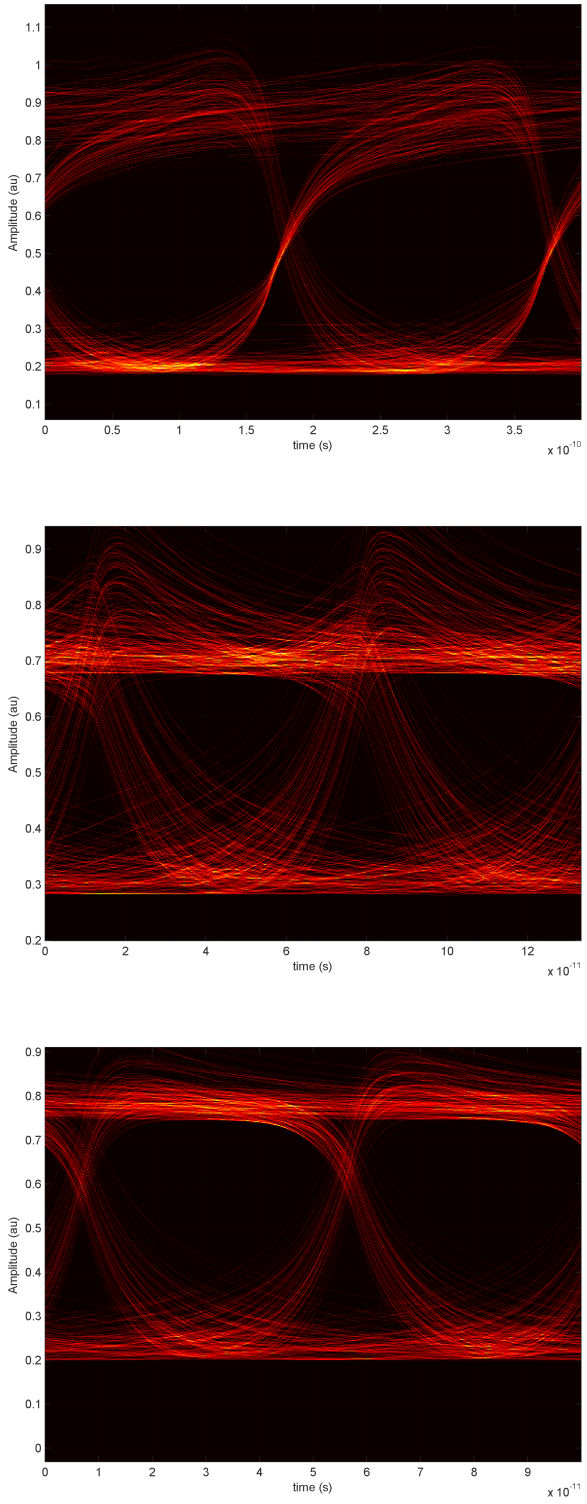


Figure 5.8: Simulated eye diagrams for the 1000-μm PP MZM at clock frequencies of 5 GHz, 15 GHz and 20 GHz, from top to bottom, with a PRBS-7 pattern at a bias of −2.5 V.

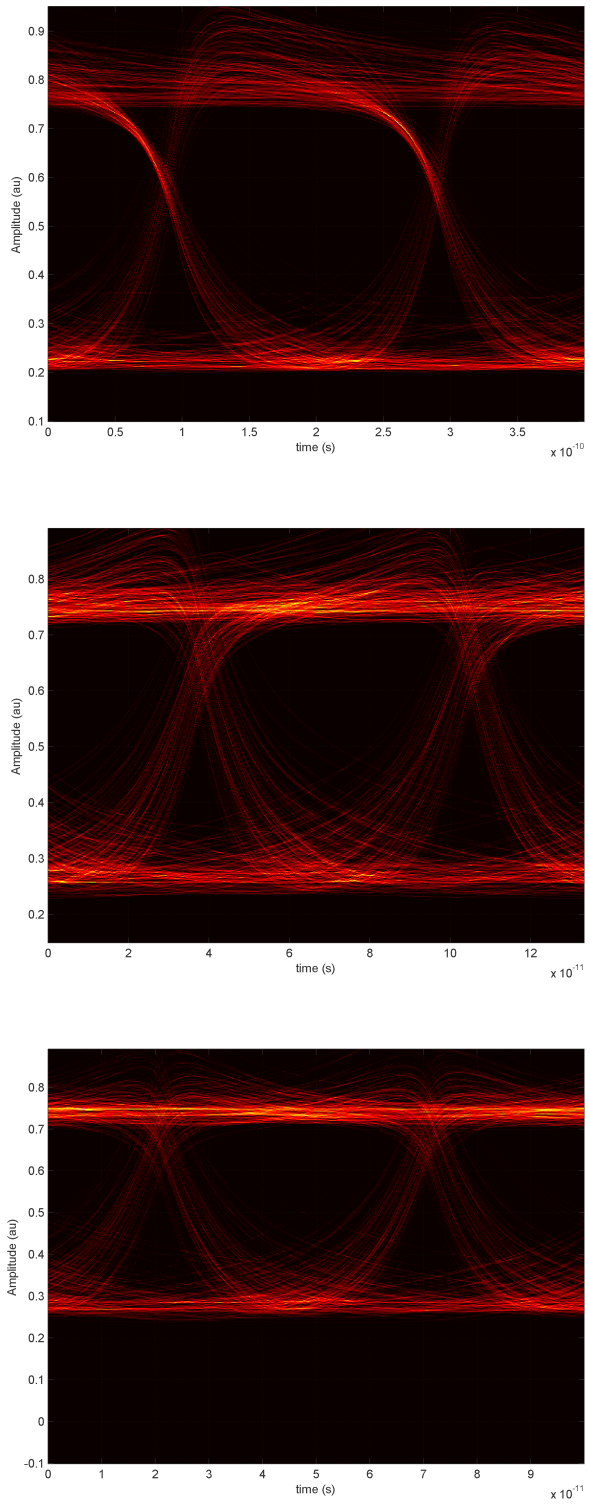


Figure 5.9: Simulated eye diagrams for the 2000-μm PP MZM at clock frequencies of 5 GHz, 15 GHz and 20 GHz, from top to bottom, with a PRBS-7 pattern at a bias of −2.5 V.

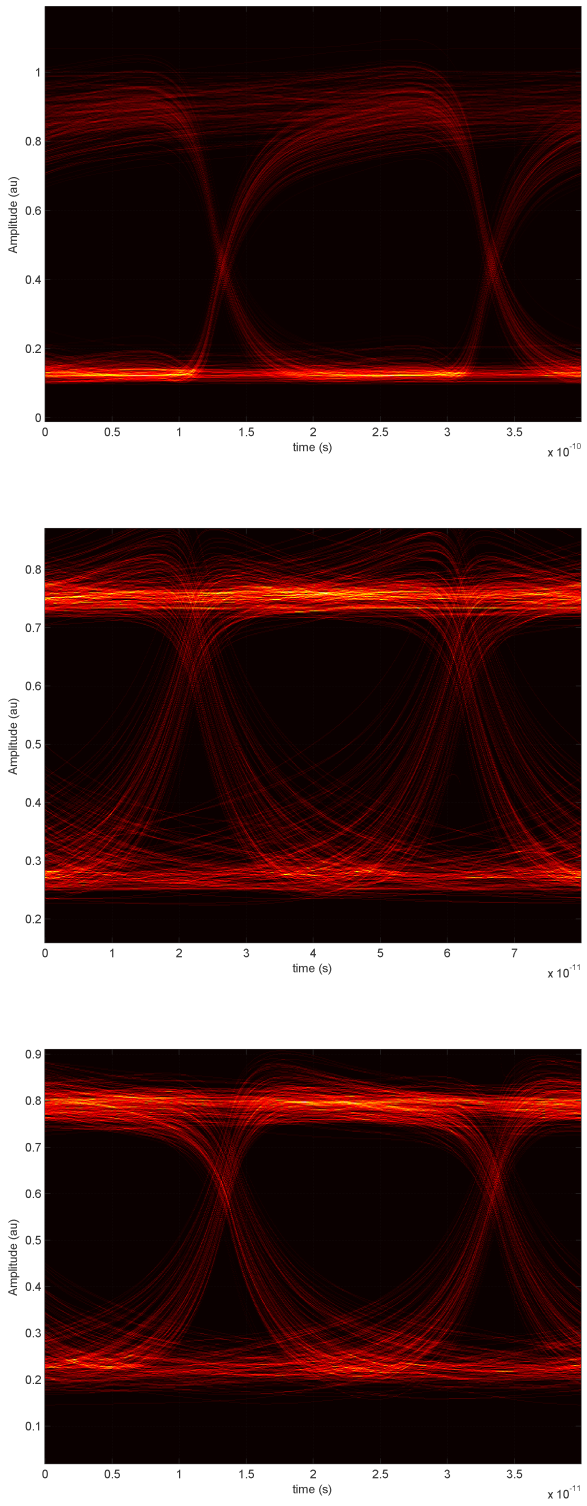
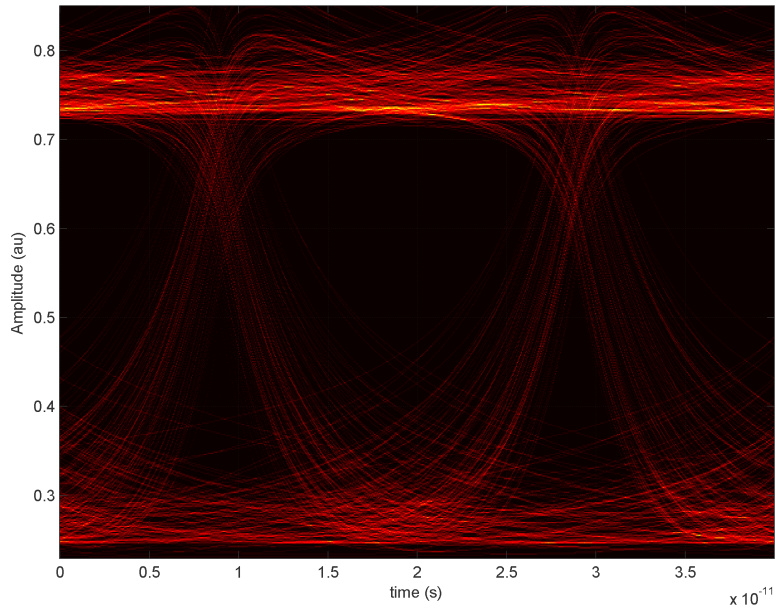
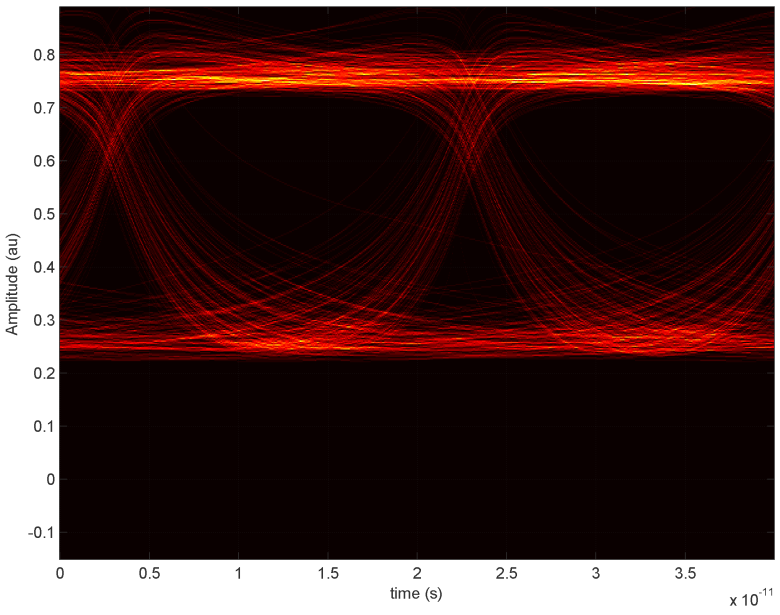


Figure 5.10: Simulated eye diagrams for the 1500-μm TWE MZM at clock frequencies of 5 GHz, 25 GHz and 50 GHz, from top to bottom, with a PRBS-7 pattern at a bias of 4 V.



(a) 1000-µm PP MZM



(b) 2000-µm PP MZM

Figure 5.11: Simulated eye diagrams for the 1000-µm and 2000-µm PP MZMs at a clock frequency of 50 GHz, with a PRBS-7 pattern at a bias of −2.5 V.

VI. Conclusions and Recommendations

6.1 Summary

To the best of my knowledge this is the first published work to compare the performance of PP and TWE MZMs that are similar in terms of waveguide structure and coupler design. It was initially predicted that the TWE MZM would outperform the PP MZMs with regard to bandwidth, simply due to the velocity and impedance matching achieved by using a segmented traveling-wave electrode rather than a lumped-element design for the electrode. What was not predicted was the fact that the TWE MZM would outperform both of the PP MZMs during the course of power-penalty measurements — even at low data rates. As the power-penalty measurements emulate the requirements levied on the modulator by the receiver, power-penalty serves as an important measurement when modulators are employed in a real-world application. The fact that the TWE MZM does not need to deliver as much optical power as the PP MZMs to achieve accurate transmission at a given BER, underscores the importance of matching optical and microwave velocities, as well as impedance requirements. The lower optical power requirements in turn translate to power savings at the source laser, or the ability to tolerate additional optical propagation losses — both important points if the MZM is to be employed in an intrachip or interchip communications network.

The results and data collected here also enable increases in device performance for future iterations, with successful modification of the MZM design. Specifically, with the wealth of data amassed on these devices — in excess of 134 million individual data points collected over 1000 hours of experimentation — device parameters may be extracted, and the design modified such that increases in bandwidth or ER are possible. The significant effort placed into the design of the experimental fixture provided the stability necessary for long-term experimentation, and the effort placed into the design of AutoLab enabled the

collection of vast amounts of data. The experimental fixture included custom mounts to hold and cool the silicon chips, as well as custom mounts to hold the lensed fiber. Typically when data is collected by hand, large step sizes and interpolation are used to speed up the process, resulting in data that is jagged rather than smooth. One such example is the surface and contour plots in the DC characterization experiments, where a resolution of 250 pm was used for wavelength, and 100 mV used for bias voltage. Additionally, five replicates were performed for each point, and the data then averaged to reduce the effect of noise. Without the use of AutoLab, which continued to collect data even when unattended, such unprecedented detail and resolution would not have been achieved in the data sets.

An unintended point of emphasis brought about by experimentation is the necessity of aligning the polarization state of the optical mode with the principal axes of the waveguide. Initial data collection was marred by the presence of a ripple that appeared to be a function of wavelength — a ripple which should not have existed given the wide optical bandwidth of all the components of the MZM. Practically, the ripple had severe effects on the ER achieved at any given wavelength. With the polarization improperly aligned with the axes of the silicon waveguide, a difference of a few nanometers of operating wavelength could easily sacrifice 5 dB in ER.

6.2 Future Work

An element that was not explored during the course of this thesis work was the effect of temperature on the performance of the MZM. All literature indicated that the thermo-optic coefficient of Si would only affect the performance of the MZM if there was a length mismatch between the two arms of the MZM. An asymmetric MZM would result in a device with a finite free-spectral range, and the location of the null points could then be wavelength tuned by changing the temperature of the MZM. It is possible that increasing the temperature of the MZM significantly could lead to lower modulation depth due to an increase in FCA via thermal excitation, and it would therefore be interesting to explore the

effects of temperature on the high-speed performance of the MZM. AutoLab was initially designed with this intent in mind, as it has the capability to change the temperature of the silicon chip to anywhere in the range of 10 °C to 120 °C, and can also use a feedback loop to maintain optical coupling in spite of the chip's movement with temperature changes.

Additionally, while the use of a tunable laser diode facilitated characterization at multiple wavelengths, the performance characterization with respect to wavelength ended with the DC characterization. Thereafter, all tests were run at a wavelength of 1550 nm. As noted in Chapter 2, the linewidth of the laser affects performance insofar as each wavelength requires a slightly different driving voltage to achieve a π radian phase shift. Therefore, a laser with a broad linewidth cannot be extinguished to the level that a laser with a narrow linewidth can, all other factors being the same. Taking this into consideration, it would be useful to accomplish experimentation using a DFB, or other narrow-linewidth laser. Additionally, changing the laser source from a tunable laser diode might result in the use of a laser with a higher maximum output power, eliminating the need for an EDFA on the input side of the MZM, thereby increasing the OSNR and increasing the measured device performance.

Yet another topic of interest would be to see if/how the device performance changes across fabrication runs. AFIT took possession of MZMs with an identical design, that came from a different fabrication run at Sandia National Laboratories. Variation in fabrication runs might introduce unintentional variations in the device, despite the fabrication tolerances at Sandia. Although in a commercial environment the fabrication process would be fine-tuned to reduce variability between batches, it would nonetheless be interesting to determine how tolerant the devices are to changes in fabrication. Ideally, the devices would be largely tolerant to variations in fabrication, as a salient feature of the adiabatic couplers was the fact that they need only meet a minimum length specification, and the splitting/coupling performance would then be assured. Issues such as junction

alignment within the waveguide, surface roughness, doping levels and doping profile would then drive device variation and performance.

Yet another series of experiments would be to fabricate each component of the MZM — the ridge waveguide, the adiabatic coupler, the active region — independently, to gain insight on how the performance of individual components affects the MZM as a whole. Factors such as wavelength dependence, polarization dependence, and optical losses can then be traced back to individual components of the MZM, and the individual components then optimized.

Finally, it would be useful to investigate the degree to which equipment limitations in the experimental setup hamper performance of the MZM. For example, mention was made throughout Section 4.3 that a driving voltage of $\leq$ 2-V_{pp} limited the ER of the MZM, and thus the maximum achievable error-free bit rate. Equipment used after the conclusion of experimentation suggested that much higher error-free speeds were possible on the MZMs tested, had they been driven with the voltage V_π. Figure 6.1 shows the performance of the 1000-μm PP MZM at 25 Gbps, and bearing in mind the fact that the 1000-μm PP MZM had a 3-dB bandwidth that was less than half of the 1500-μm TWE MZM, the performance of the 1500-μm TWE MZM looks promising.

Although not pictured earlier in this thesis, a number of silicon chips were delivered to AFIT by Sandia National Laboratories, each with large quantities of silicon photonics devices, to include a number of additional MZMs, ring resonators, and highly unique SiGe high-speed photodetectors. The identical geometries of the chips, and the fact that light is coupled onto and off of the chip via similar ridge waveguides, mean that the current experimental fixture and AutoLab can continue to be used in future work and on other devices.

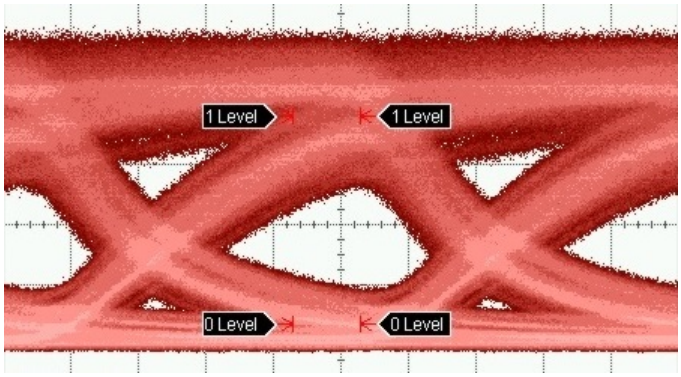


Figure 6.1: Eye diagram for the 1000-μm PP MZM at 25 Gbps, and operating at a wavelength of 1550 nm using the straight port. The bias voltage was set to −4.25 V, and the driving voltage to 6-V_{pp}. The signal displays an ER of 8.33 dB with 60 mV per division. Although a BER test was not accomplished using this experimental setup, the size of the eye opening looks to be promising, and indicates that the MZM performance may have been limited by the experimental setup used during the course of this thesis work.

Appendix A: Additional Data

This section contains additional data not presented in the main body of the text, in order to improve document readability and reduce the length of the main portion of the text.

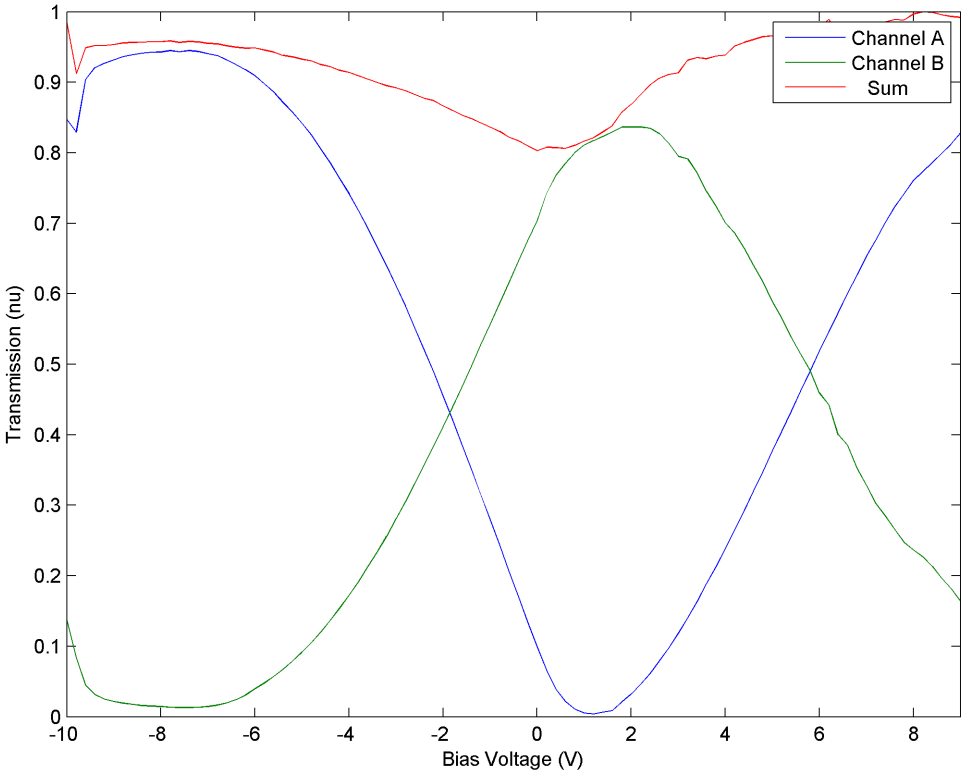


Figure A.1: The output of the 1000-μm PP MZM, showing the complementary nature of the top and bottom output arms. The output transmission is normalized, and given on a linear scale at λ=1550 nm.

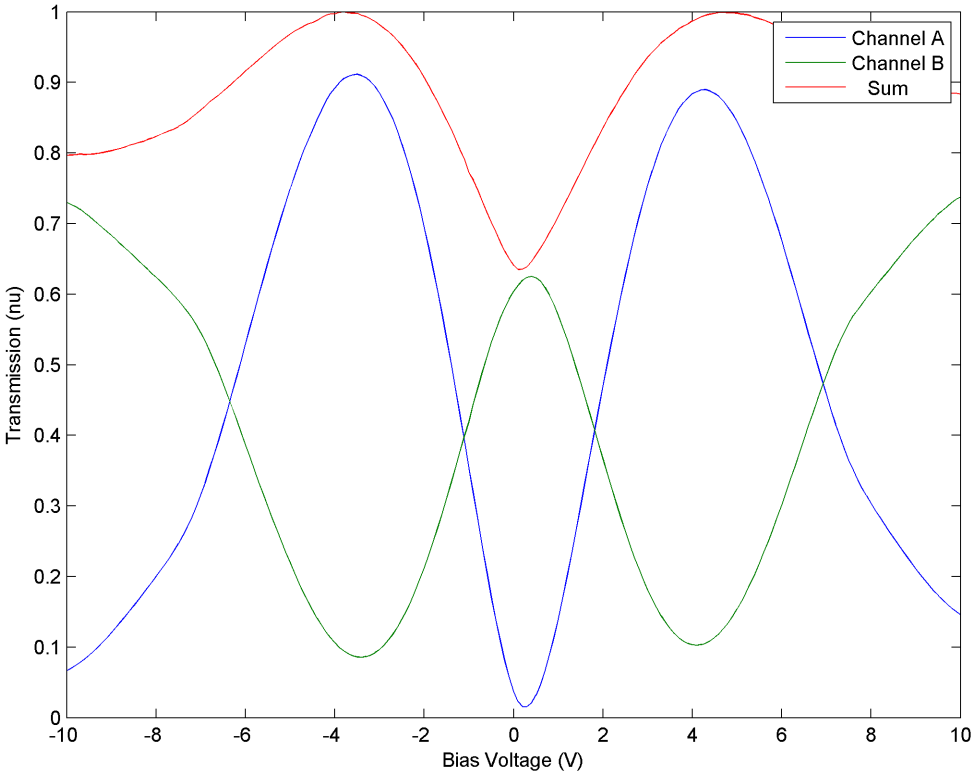


Figure A.2: The output of the 2000-μm PP MZM, showing the complementary nature of the top and bottom output arms. The output transmission is normalized, and given on a linear scale at λ=1550 nm.

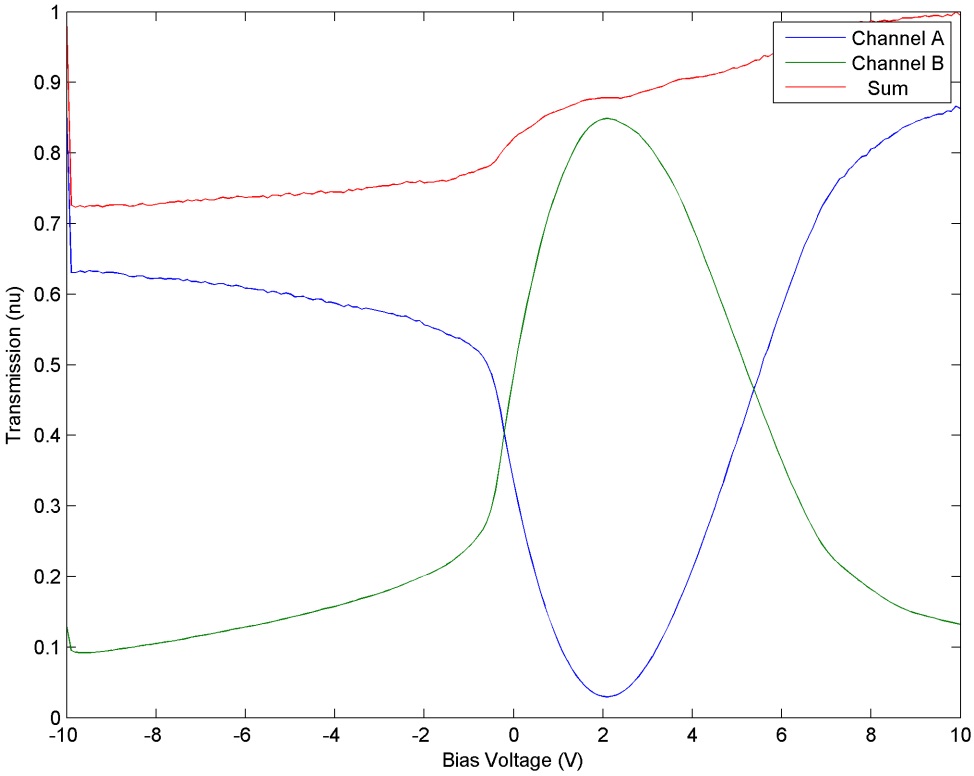
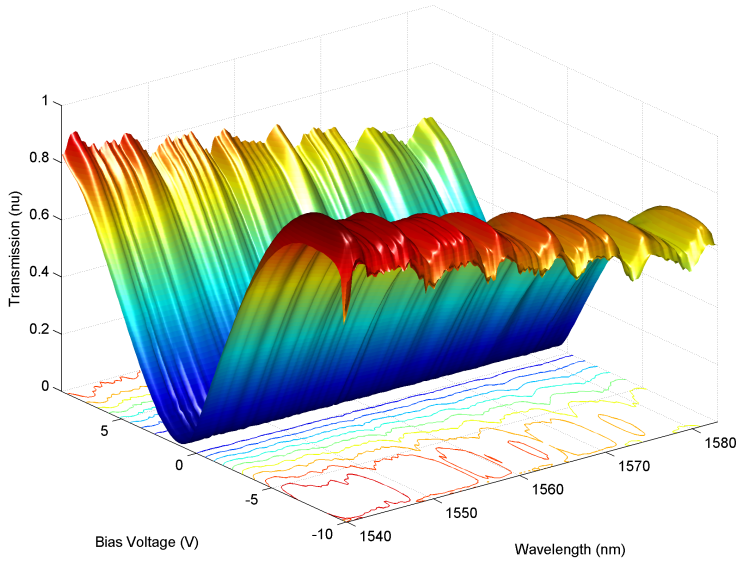
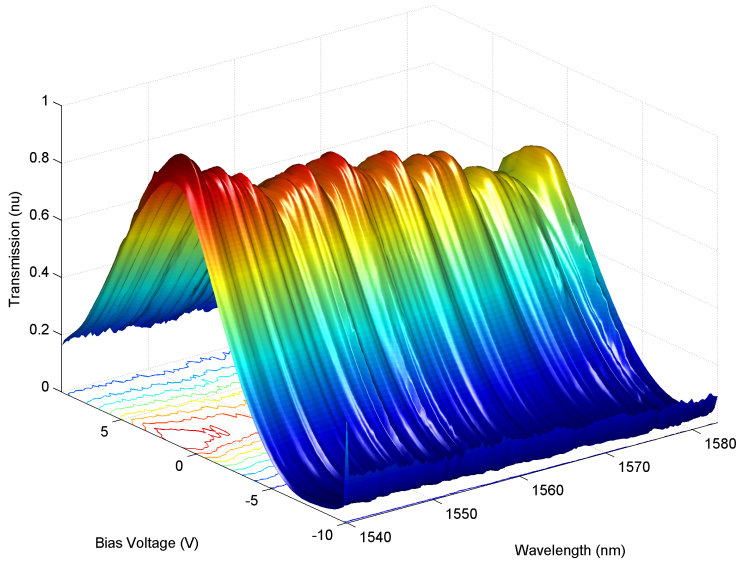


Figure A.3: The output of the 1500-μm TWE MZM, showing the complementary nature of the top and bottom output arms. The output transmission is normalized, and given on a linear scale at λ=1550 nm.

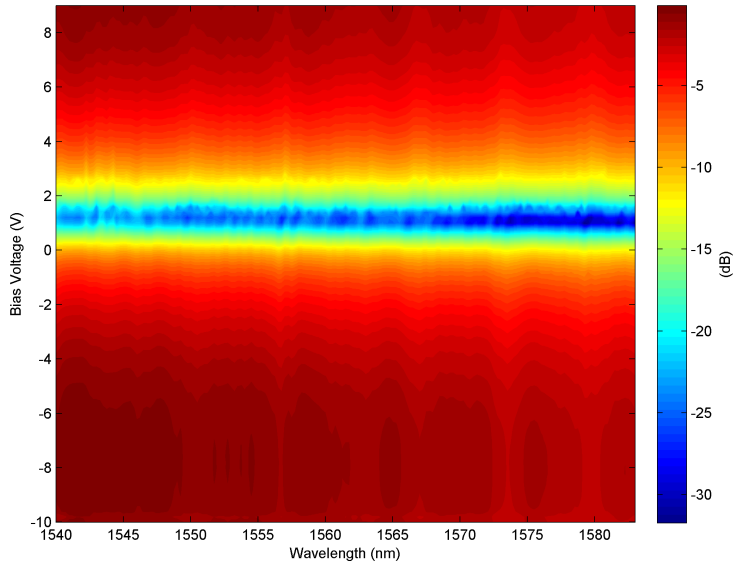


(a) Straight Port

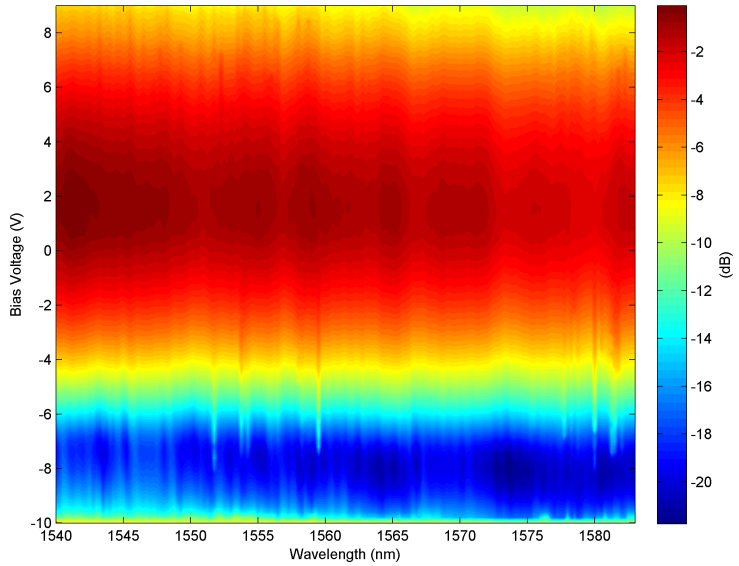


(b) Cross Port

Figure A.4: The transmission of the 1000-μm PP MZM as a function of wavelength and bias, using a linear scale. The output transmission is normalized to the largest transmission value in each plot.

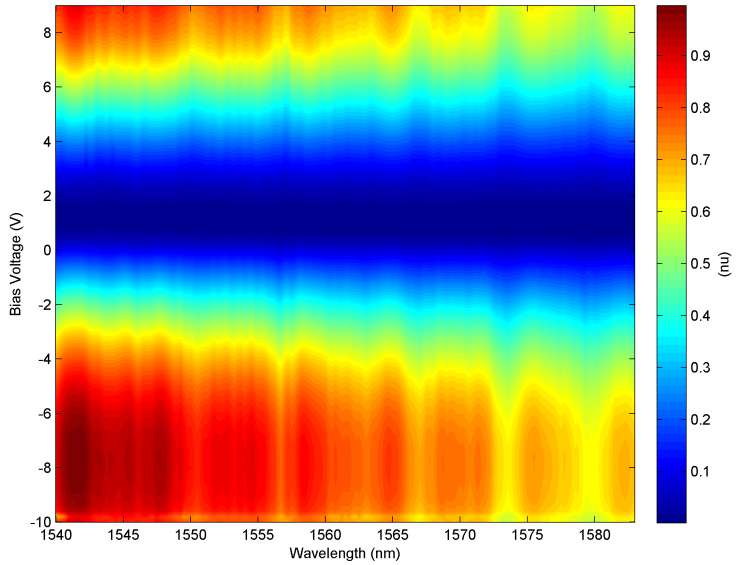


(a) Straight Port

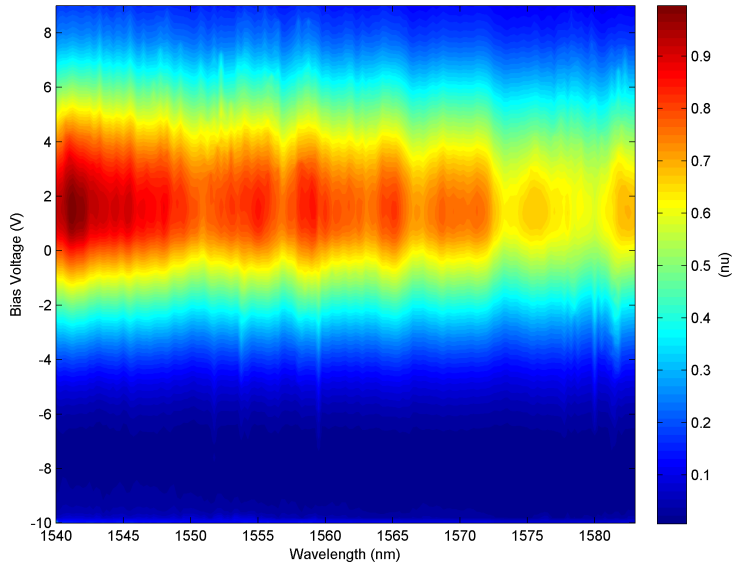


(b) Cross Port

Figure A.5: The transmission of the 1000-μm PP MZM as a function of wavelength and bias, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.

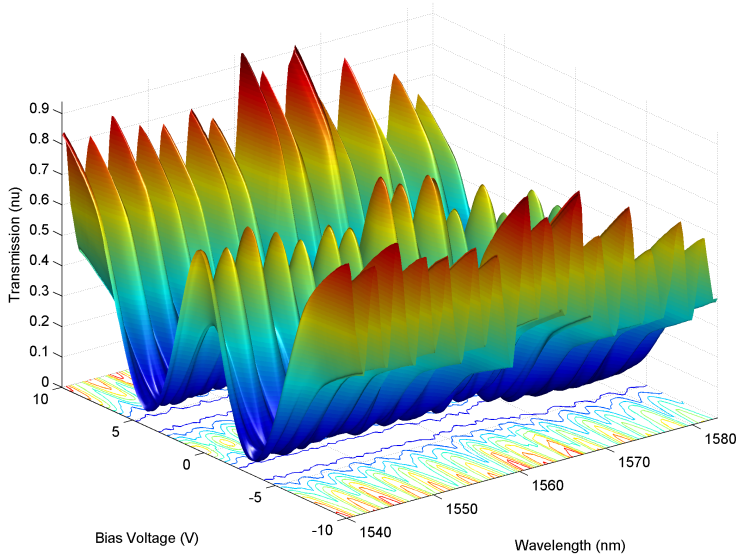


(a) Straight Port

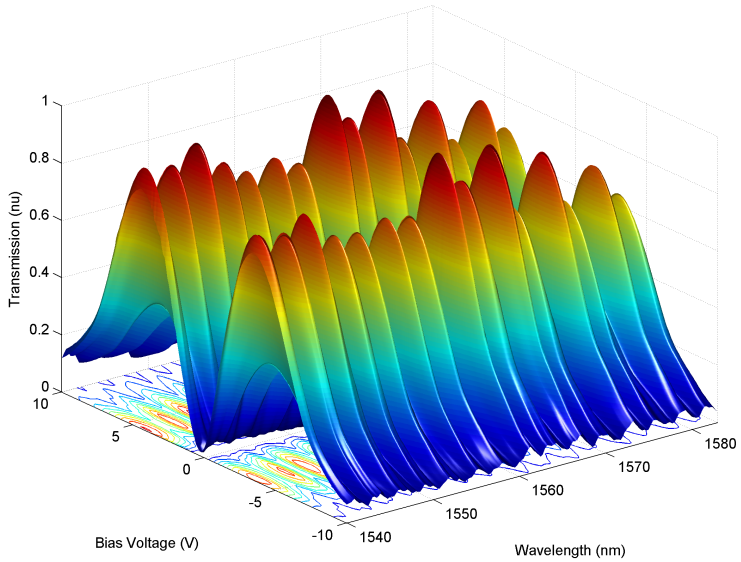


(b) Cross Port

Figure A.6: The transmission of the 1000-μm PP MZM as a function of wavelength and bias, using a linear scale. The output transmission is normalized to the largest transmission value in each plot.

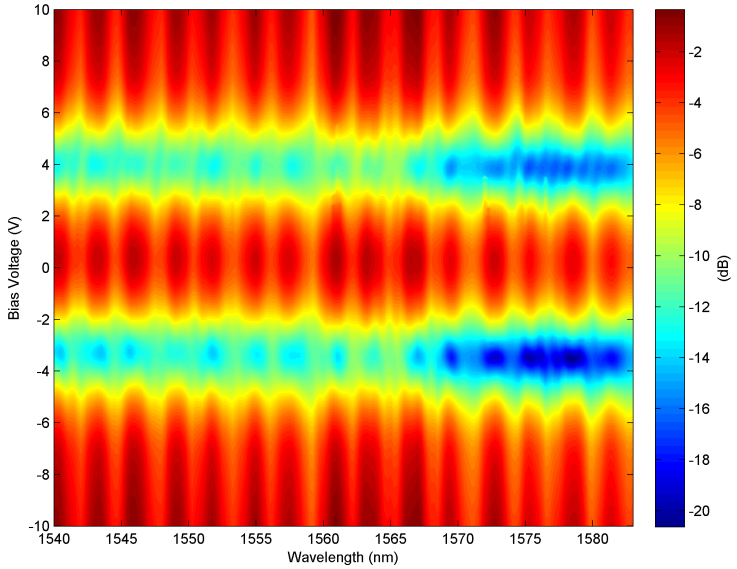


(a) Straight Port

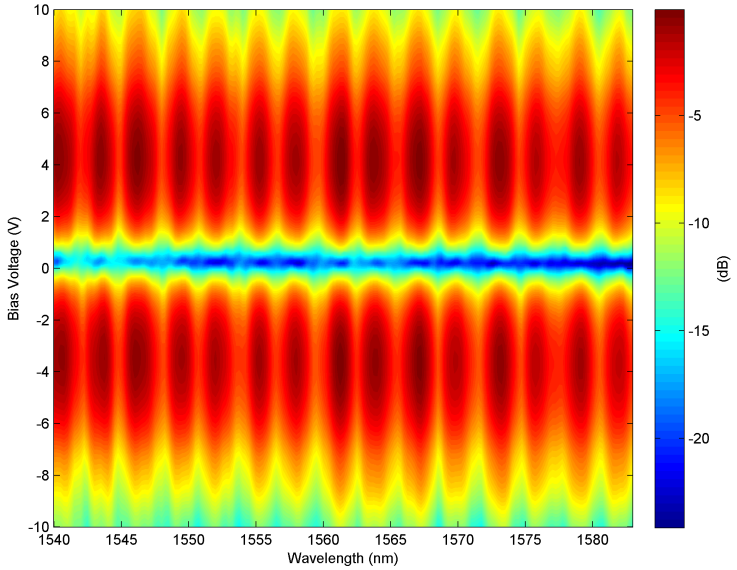


(b) Cross Port

Figure A.7: The transmission of the 2000-μm PP MZM as a function of wavelength and bias, using a linear scale. The output transmission is normalized to the largest transmission value in each plot.

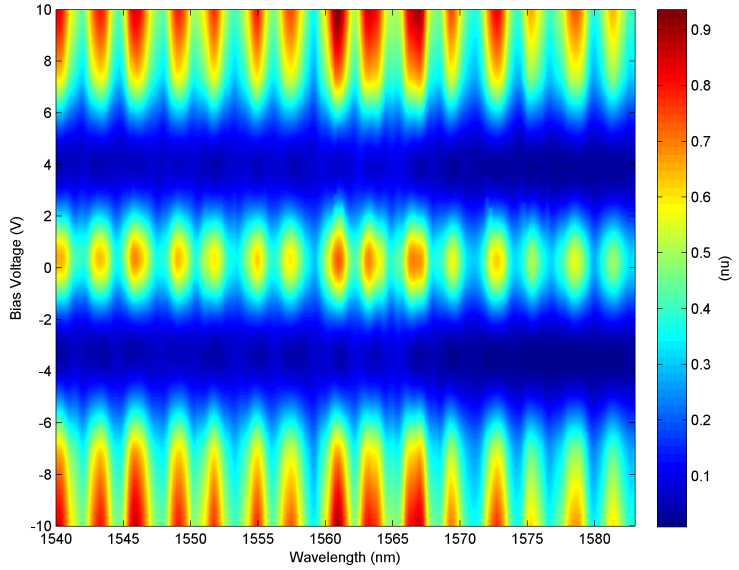


(a) Straight Port

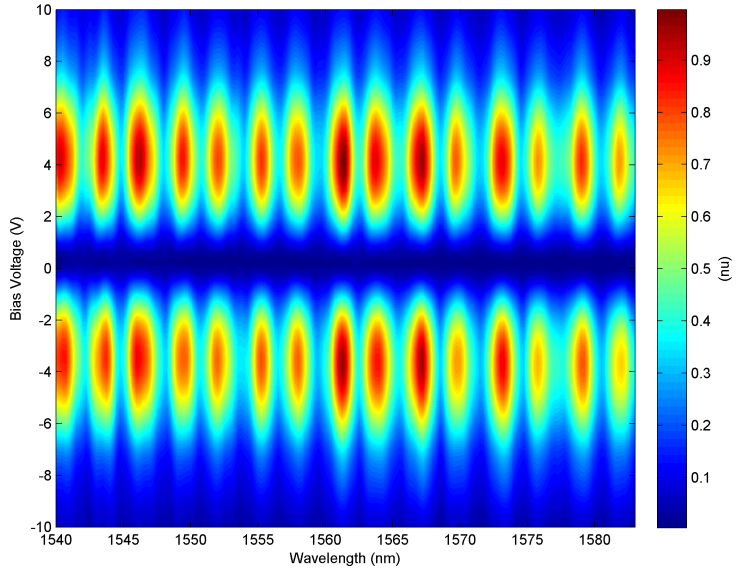


(b) Cross Port

Figure A.8: The transmission of the 2000-μm PP MZM as a function of wavelength and bias, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.

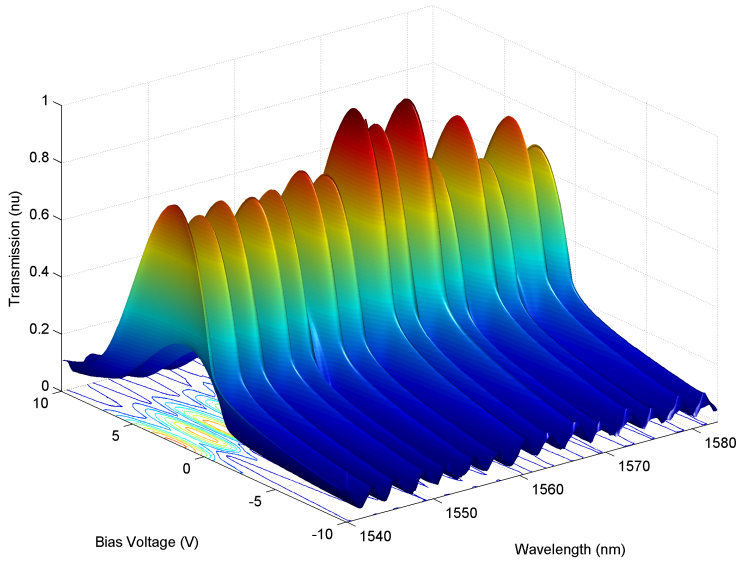


(a) Straight Port

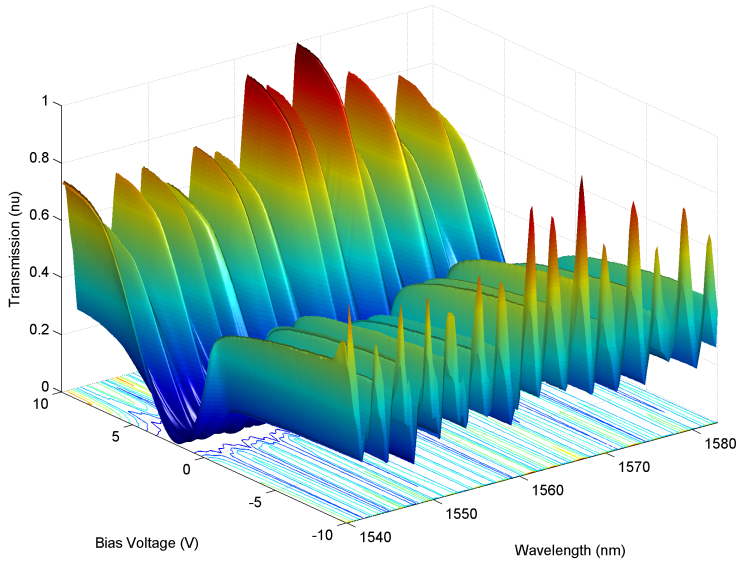


(b) Cross Port

Figure A.9: The transmission of the 2000-μm PP MZM as a function of wavelength and bias, using a linear scale. The output transmission is normalized to the largest transmission value in each plot.

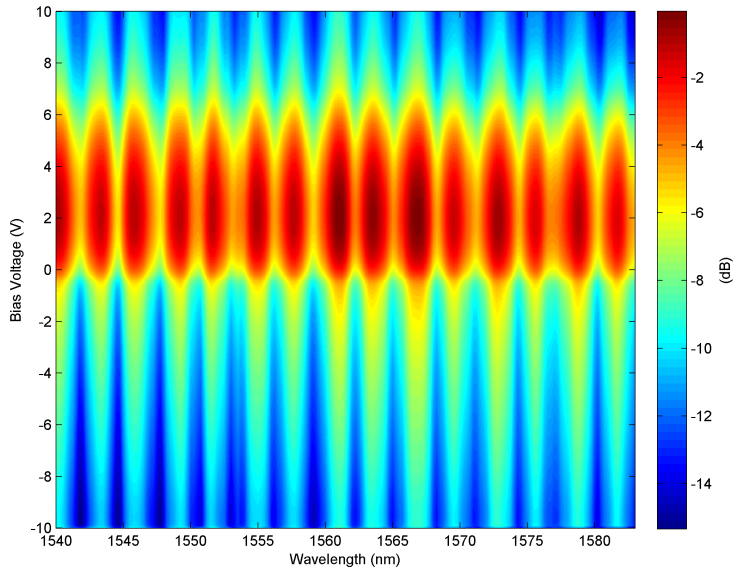


(a) Straight Port

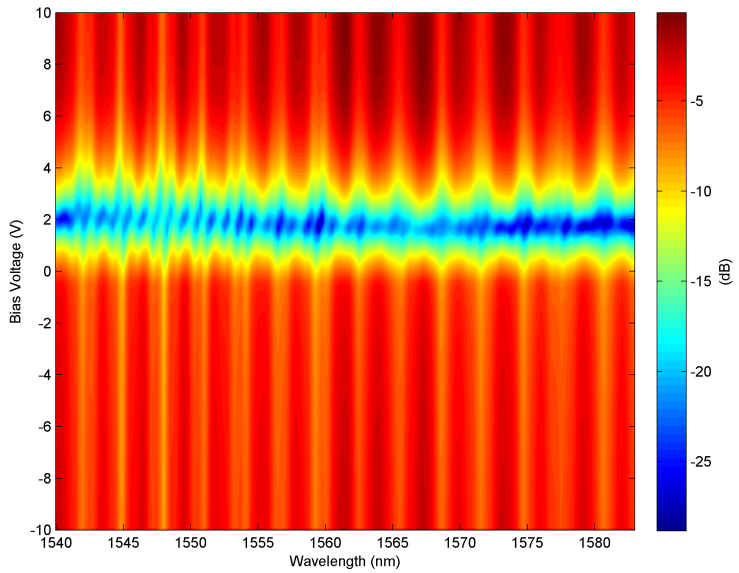


(b) Cross Port

Figure A.10: The transmission of the 1500-μm TWE MZM as a function of wavelength and bias, using a linear scale. The output transmission is normalized to the largest transmission value in each plot.

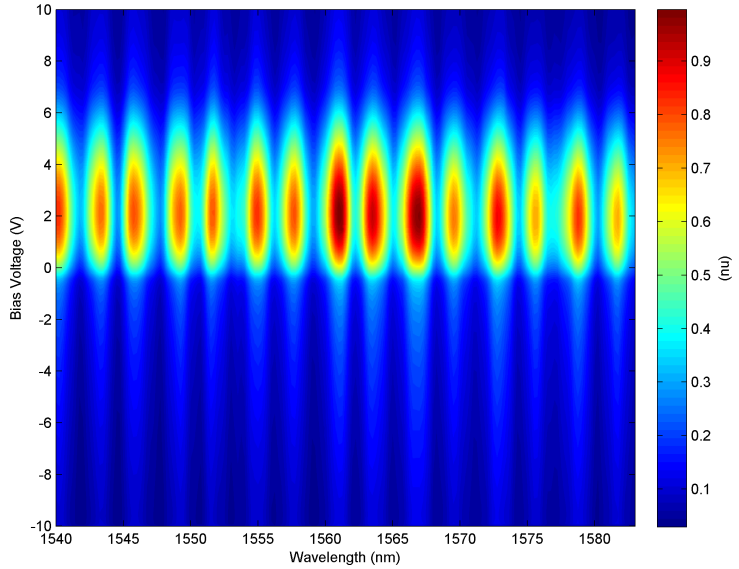


(a) Straight Port

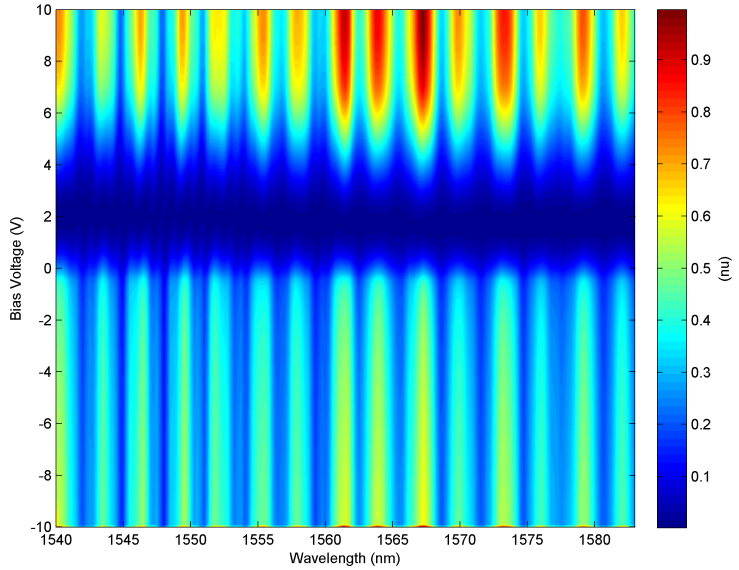


(b) Cross Port

Figure A.11: The transmission of the 1500-μm TWE MZM as a function of wavelength and bias, using a logarithmic scale. The output transmission is normalized to the largest transmission value in each plot.



(a) Straight Port



(b) Cross Port

Figure A.12: The transmission of the 1500-μm TWE MZM as a function of wavelength and bias, using a linear scale. The output transmission is normalized to the largest transmission value in each plot.

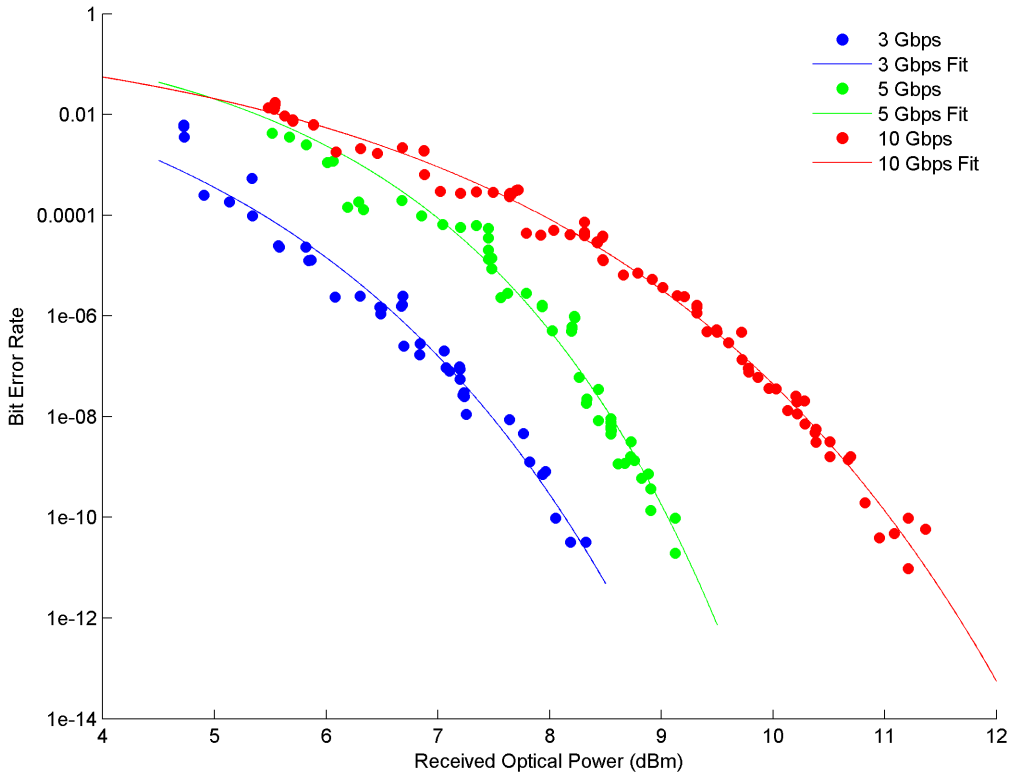


Figure A.13: Power-penalty curves for the 1000-µm PP MZM at 3 Gbps, 5 Gbps and 10 Gbps using an end-to-end connection, taken from the cross port at 1550 nm.

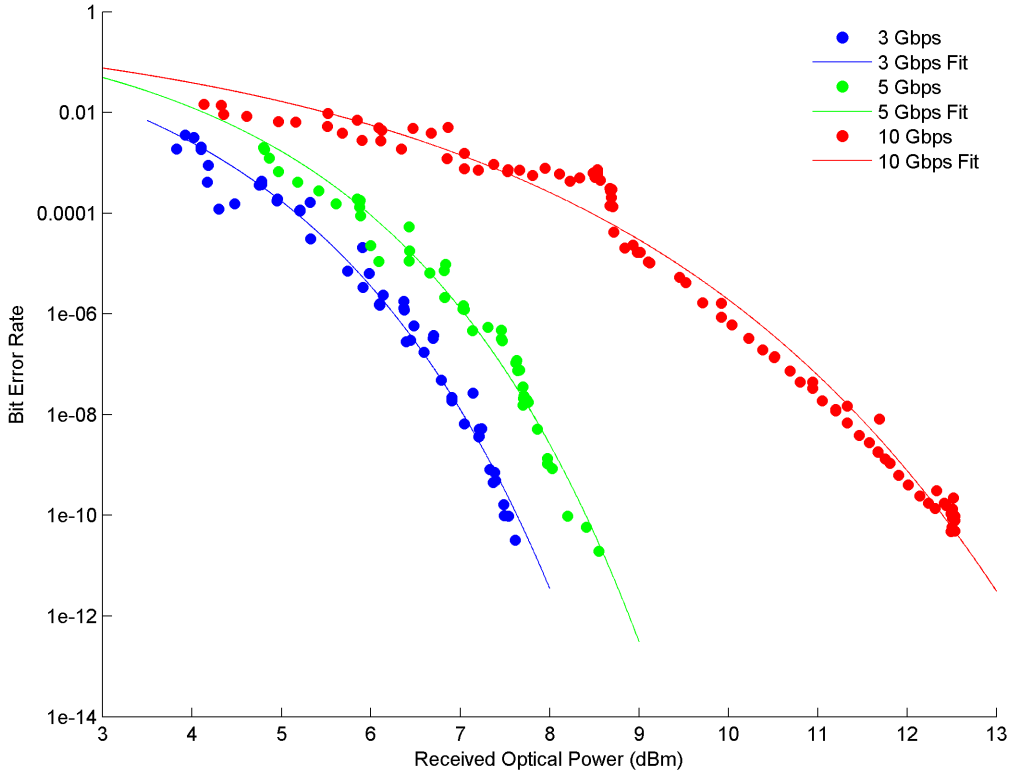


Figure A.14: Power-penalty curves for the 2000-µm PP MZM at 3 Gbps, 5 Gbps and 10 Gbps using an end-to-end connection, taken from the straight port at 1550 nm.

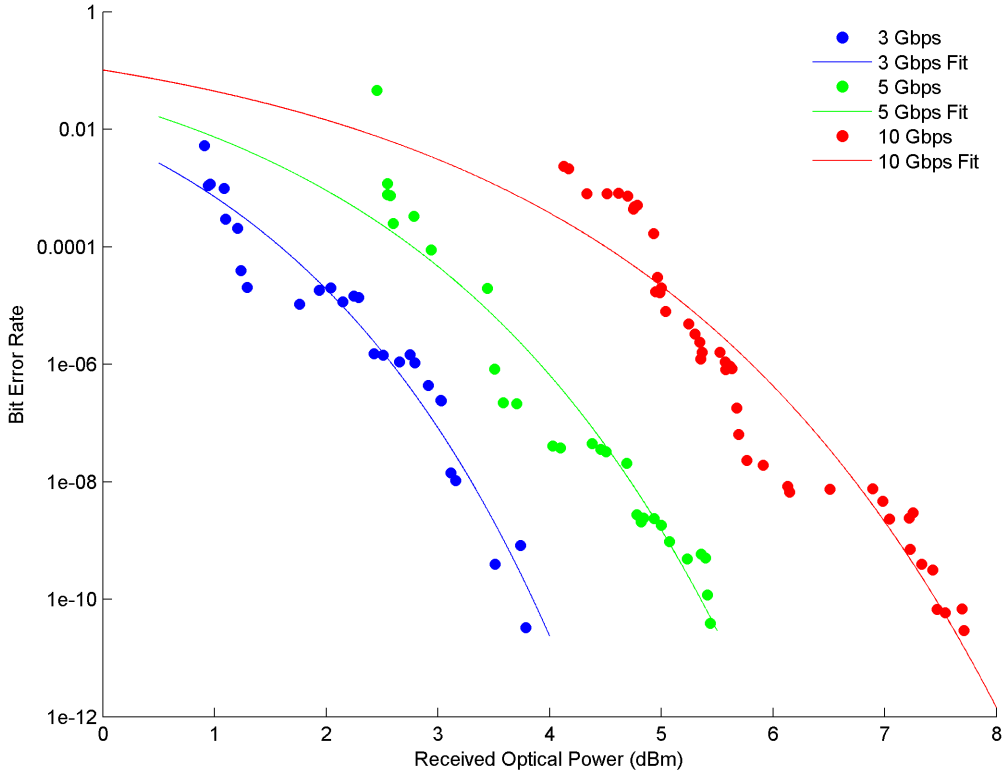


Figure A.15: Power-penalty measurements for the 1500-µm TWE MZM at 3 Gbps, 5 Gbps and 10 Gbps using an end-to-end connection, taken from the straight port at 1550 nm.

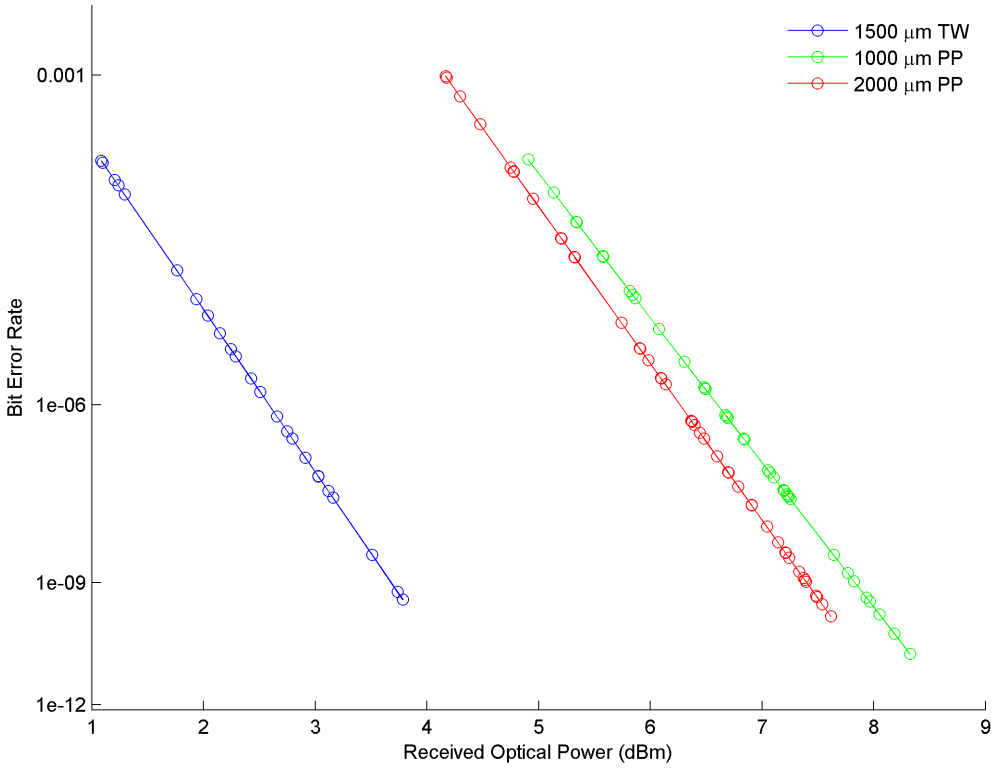


Figure A.16: Power-penalty curves for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs at 3 Gbps using an end-to-end connection at 1550 nm.

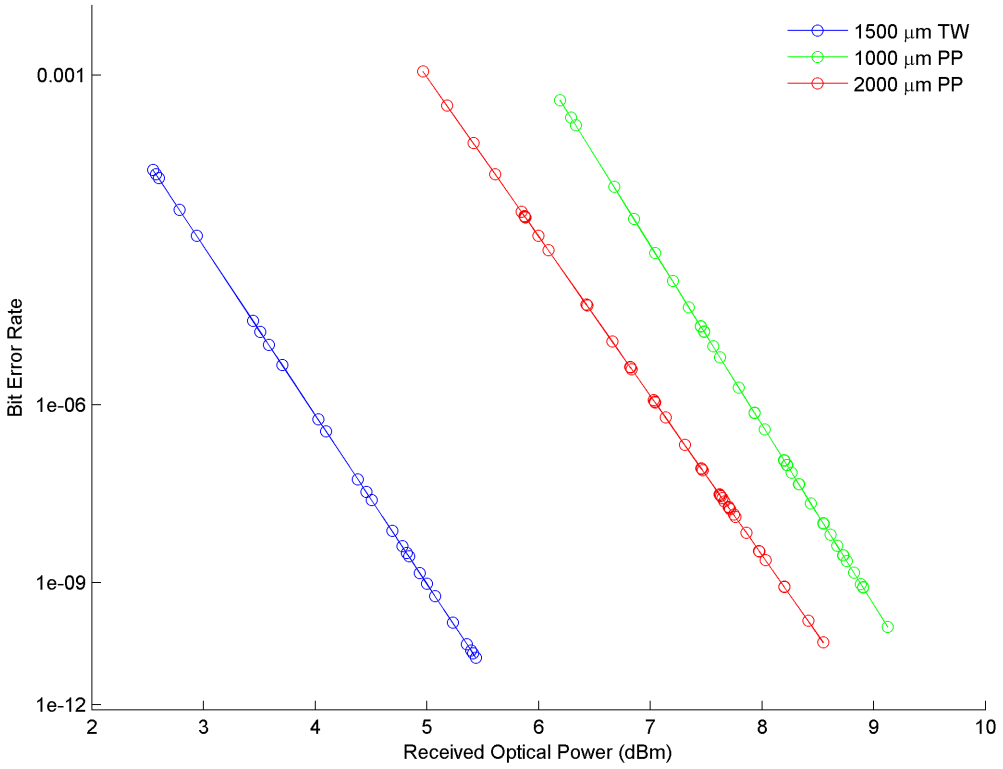


Figure A.17: Power-penalty curves for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs at 5 Gbps using an end-to-end connection at 1550 nm.

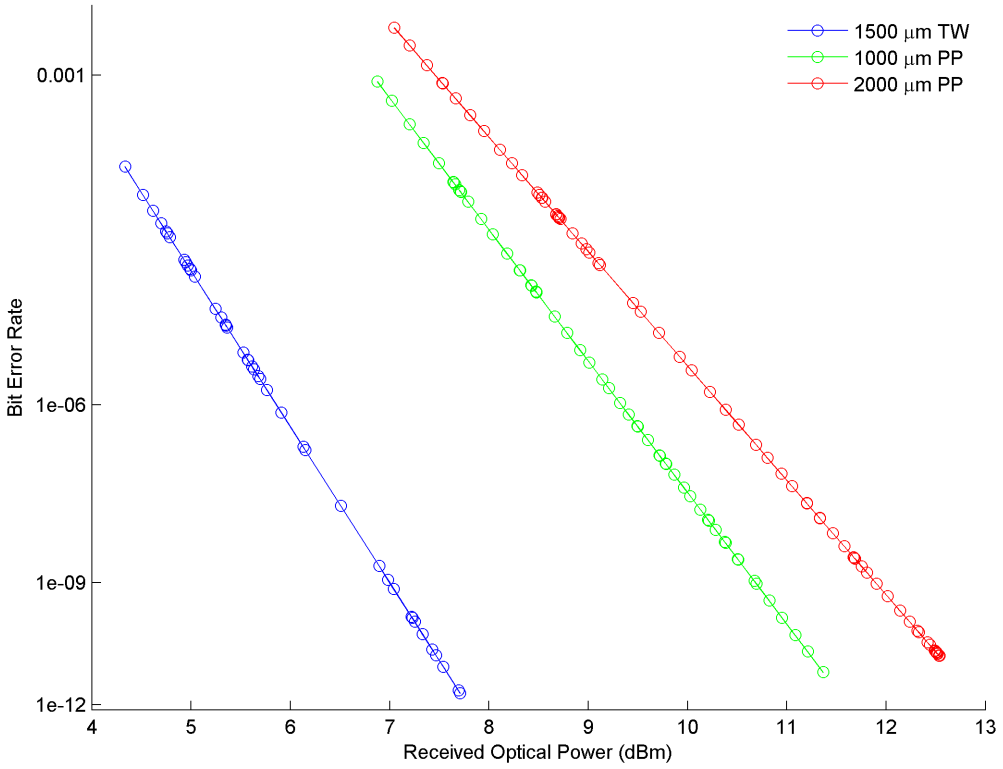


Figure A.18: Power-penalty curves for the 1000-μm PP, 2000-μm PP, and 1500-μm TWE MZMs at 10 Gbps using an end-to-end connection at 1550 nm.

Appendix B: Derivation of the Intensity Upon the Interference of Two Waves

The following derivation is for the transmission intensity of the MZM, taking into account the use of a push-pull modulation scheme. First suppose that the optical mode's electric fields take the following form:

$$
\begin{aligned}
\$ \quad \vec{E_1} &= E_{01}\, e^{-\alpha_0 L} e^{-\Delta\alpha_1 L} \cos\left(\vec{k_1}\cdot\vec{r_1} - \omega t + \phi_1\right) \\
&= E_{01}\, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \cos\left(\vec{k_1}\cdot\vec{r_1} - \omega t + \phi_1\right) \\
\vec{E_2} &= E_{02}\, e^{-\alpha_0 L} e^{-\Delta\alpha_2 L} \cos\left(\vec{k_2}\cdot\vec{r_2} - \omega t + \phi_2\right) \\
&= E_{02}\, e^{-(\alpha_0 L + \Delta\alpha_2 L)} \cos\left(\vec{k_2}\cdot\vec{r_2} - \omega t + \phi_2\right)
\end{aligned}
$$

The time-average of the electric field can then be computed as follows, where the subscript x is meant to designate a specific electric field, for example $\vec{E_1}$ or $\vec{E_2}$. This result will be used in later computations.

$$
\begin{aligned}
\left\langle \vec{E_x}^{\,2} \right\rangle_T &= \frac{1}{T} \int_{t_0}^{t_0+T} \vec{E_x}^{\,2} \, \mathrm{d}t, \, T \gg 2\pi/\omega, \, T \gg t_0 \\
&= \frac{1}{T} \int_{t_0}^{t_0+T} \left[E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \cos\left(\vec{k_1} \cdot \vec{r_1} - \omega t + \phi_1 \right) \right]^2 \, \mathrm{d}t \\
&= \frac{1}{T} \int_{t_0}^{t_0+T} \left(E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \right)^2 \cos^2\left(\vec{k_1} \cdot \vec{r_1} - \omega t + \phi_1 \right) \, \mathrm{d}t \\
&= \frac{1}{T} \left(E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \right)^2 \int_{t_0}^{t_0+T} \cos^2\left(\vec{k_1} \cdot \vec{r_1} - \omega t + \phi_1 \right) \, \mathrm{d}t \\
&= \frac{1}{T} \left(E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \right)^2 \int_{t_0}^{t_0+T} \frac{1}{2} \left[1 + \cos\left(2\vec{k_1} \cdot \vec{r_1} - 2\omega t + 2\phi_1 \right) \right] \, \mathrm{d}t \\
&= \frac{1}{T} \left(E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \right)^2 \frac{T}{2} + \frac{1}{T} \int_{t_0}^{t_0+T} \frac{1}{2} \cos\left(2\vec{k_1} \cdot \vec{r_1} - 2\omega t + 2\phi_1 \right) \, \mathrm{d}t \\
&= \frac{1}{2} \left(E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \right)^2 + \frac{1}{T} \int_{t_0}^{t_0+T} \frac{1}{2} \cos\left(2\vec{k_1} \cdot \vec{r_1} - 2\omega t + 2\phi_1 \right) \, \mathrm{d}t \\
&= \frac{1}{2} \left(E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \right)^2 + \frac{1}{2T} \sin\left(2\vec{k_1} \cdot \vec{r_1} - 2\omega t + 2\phi_1 \right) \Big|_{t=t_0}^{t_0+T} \\
&= \frac{1}{2} \left(E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} \right)^2 + \frac{1}{2T} \left[\sin\left(2\vec{k_1} \cdot \vec{r_1} - 2\omega(t_0 + T) + 2\phi_1 \right) \right. \\
&\quad \left. - \sin\left(2\vec{k_1} \cdot \vec{r_1} - 2\omega t_0 + 2\phi_1 \right) \right] \\
&= \frac{1}{2} \left(E_{0x} \, e^{-(\alpha_0 L + \Delta\alpha_x L)} \right)^2 \\
&= \frac{1}{2} E_{0x}^2 \, e^{-2(\alpha_0 L + \Delta\alpha_x L)}
\end{aligned}
$$

The transmission intensity is then the product of the phase velocity, dielectric constant, and time average of the electric field squared — the electric field being a superposition of the two input electric fields.

$$
\begin{aligned}
T &= \nu\varepsilon \left\langle E^2 \right\rangle_T, \nu = \frac{c}{n}, \varepsilon = \varepsilon_r\varepsilon_0 \\
&= \frac{c\varepsilon_r\varepsilon_0}{n} \left\langle E^2 \right\rangle_T, E = E_1 + E_2 \\
&= \frac{c\varepsilon_r\varepsilon_0}{n} \left\langle (E_1 + E_2)^2 \right\rangle_T \\
&= \frac{c\varepsilon_r\varepsilon_0}{n} \left\langle E_1^2 + E_2^2 + 2E_1E_2 \right\rangle_T \\
&= \frac{c\varepsilon_r\varepsilon_0}{n} \left(\left\langle E_1^2 \right\rangle_T + \left\langle E_2^2 \right\rangle_T + \langle 2E_1E_2 \rangle_T \right) \\
&= \frac{c\varepsilon_r\varepsilon_0}{n} \left(\left\langle E_1^2 \right\rangle_T + \left\langle E_2^2 \right\rangle_T + 2\langle E_1E_2 \rangle_T \right) \\
&= \frac{c\varepsilon_r\varepsilon_0}{n} \left(\frac{1}{2}E_{01}^2 \, e^{-2(\alpha_0 L + \Delta\alpha_1 L)} + \frac{1}{2}E_{02}^2 \, e^{-2(\alpha_0 L + \Delta\alpha_2 L)} + 2\langle E_1E_2 \rangle_T \right)
\end{aligned}
$$

It is then necessary to compute the time-average of the product of two electric fields, as is required by the above equation. The product of two electric fields is first calculated.

$$
\begin{aligned}
\vec{E_1}\vec{E_2} &= E_{01} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} E_{02} \, e^{-(\alpha_0 L + \Delta\alpha_2 L)} \cos\left(\vec{k_1} \cdot \vec{r_1} - \omega t + \phi_1\right) \cos\left(\vec{k_2} \cdot \vec{r_2} - \omega t + \phi_2\right) \\
&= E_{01} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} E_{02} \, e^{-(\alpha_0 L + \Delta\alpha_2 L)} \Big[\Big(\cos\left(\vec{k_1} \cdot \vec{r_1} + \phi_1\right) \cos(\omega t) \\
&\quad + \sin\left(\vec{k_1} \cdot \vec{r_1} + \phi_1\right) \sin(\omega t) \Big) \Big(\cos\left(\vec{k_2} \cdot \vec{r_2} + \phi_2\right) \cos(\omega t) + \sin\left(\vec{k_2} \cdot \vec{r_2} + \phi_2\right) \sin(\omega t) \Big) \Big] \\
&= E_{01} \, e^{-(\alpha_0 L + \Delta\alpha_1 L)} E_{02} \, e^{-(\alpha_0 L + \Delta\alpha_2 L)} \Big[\cos\left(\vec{k_1} \cdot \vec{r_1} + \phi_1\right) \cos\left(\vec{k_2} \cdot \vec{r_2} + \phi_2\right) \cos^2(\omega t) \\
&\quad + \cos\left(\vec{k_2} \cdot \vec{r_2} + \phi_2\right) \cos(\omega t) \sin\left(\vec{k_1} \cdot \vec{r_1} + \phi_1\right) \sin(\omega t) \\
&\quad + \cos\left(\vec{k_1} \cdot \vec{r_1} + \phi_1\right) \cos(\omega t) \sin\left(\vec{k_2} \cdot \vec{r_2} + \phi_2\right) \sin(\omega t) \\
&\quad + \sin\left(\vec{k_1} \cdot \vec{r_1} + \phi_1\right) \sin\left(\vec{k_2} \cdot \vec{r_2} + \phi_2\right) \sin^2(\omega t) \Big]
\end{aligned}
$$

$$
\left\langle \vec{E_1}\vec{E_2} \right\rangle_T = \frac{1}{T} \int_{t_0}^{t_0+T} \vec{E_1}\vec{E_2} \, \mathrm{d}t, T \gg 2\pi/\omega, T \gg t_0
$$

It is then necessary to compute the time average of several sinusoids, for use in the previous equation.

$$
\begin{aligned}
\left\langle \cos^2(\omega t) \right\rangle_T &= \frac{1}{T} \int_{t_0}^{t_0+T} \cos^2(\omega t)\, \mathrm{d}t,\ T \gg 2\pi/\omega,\ T \gg t_0 \\
&= \frac{1}{T} \int_{t_0}^{t_0+T} \frac{1}{2} \left(1 + \cos(2\omega t)\right)\ \mathrm{d}t \\
&= \frac{1}{T} \left(\frac{1}{2} T + \int_{t_0}^{t_0+T} \frac{1}{2} \cos(2\omega t)\, \mathrm{d}t \right) \\
&= \frac{1}{T} \left(\frac{1}{2} T + \frac{1}{2} \sin(2\omega t) \right) \Bigg|_{t=t_0}^{t_0+T} \\
&= \frac{1}{2} + \frac{1}{2T} \sin(2\omega t) \Big|_{t=t_0}^{t_0+T} \\
&= \frac{1}{2} + \left(\frac{1}{2T} \sin(2\omega(t_0+T)) - \frac{1}{2T} \sin(2\omega t_0) \right) \\
&= \frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
\left\langle \sin^2(\omega t) \right\rangle_T &= \frac{1}{T} \int_{t_0}^{t_0+T} \sin^2(\omega t)\, \mathrm{d}t,\ T \gg 2\pi/\omega,\ T \gg t_0 \\
&= \frac{1}{T} \int_{t_0}^{t_0+T} \frac{1}{2} \left(1 + \sin(2\omega t)\right)\ \mathrm{d}t \\
&= \frac{1}{T} \left(\frac{1}{2} T + \int_{t_0}^{t_0+T} \frac{1}{2} \sin(2\omega t)\, \mathrm{d}t \right) \\
&= \frac{1}{T} \left(\frac{1}{2} T - \frac{1}{2} \cos(2\omega t) \right) \Bigg|_{t=t_0}^{t_0+T} \\
&= \frac{1}{2} - \frac{1}{2T} \cos(2\omega t) \Big|_{t=t_0}^{t_0+T} \\
&= \frac{1}{2} - \left(\frac{1}{2T} \cos(2\omega(t_0+T)) - \frac{1}{2T} \cos(2\omega t_0) \right) \\
&= \frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
\langle \sin(\omega t)\cos(\omega t)\rangle_T &= \frac{1}{T}\int_{t_0}^{t_0+T} \sin(\omega t)\cos(\omega t)\,\mathrm{d}t,\ T \gg 2\pi/\omega,\ T \gg t_0,\ u=\omega t,\ \mathrm{d}u=\omega\,\mathrm{d}t \\
&= \frac{1}{T}\int_{t_0}^{t_0+T} \sin(u)\cos(u)\frac{\mathrm{d}u}{\omega} \\
&= \frac{1}{\omega T}\int_{t_0}^{t_0+T} \sin(u)\cos(u)\,\mathrm{d}u,\ v=\cos(u),\ \mathrm{d}v=-\sin(u)\,\mathrm{d}u \\
&= \frac{1}{\omega T}\int_{t_0}^{t_0+T} -v\,\mathrm{d}v \\
&= \frac{-1}{\omega T}\int_{t_0}^{t_0+T} v\,\mathrm{d}v \\
&= \frac{-1}{\omega T}\left.\frac{u^2}{2}\right|_{t_0}^{t_0+T} \\
&= \frac{-1}{2\omega T}\left.u^2\right|_{t_0}^{t_0+T} \\
&= \frac{-1}{2\omega T}\left.\cos^2(u)\right|_{t_0}^{t_0+T} \\
&= \frac{-1}{2\omega T}\left.\cos^2(\omega t)\right|_{t_0}^{t_0+T} \\
&= \frac{-1}{2\omega T}\left[\cos^2(\omega(t_0+T)) - \cos^2(\omega t_0)\right] \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\therefore \vec{E}_1\vec{E}_2 &= E_{01}\,e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\,e^{-(\alpha_0 L+\Delta\alpha_2 L)}\left[\frac{1}{2}\cos\left(\vec{k_1}\cdot\vec{r_1}+\phi_1\right)\cos\left(\vec{k_2}\cdot\vec{r_2}+\phi_2\right)\right. \\
&\quad \left.+\frac{1}{2}\sin\left(\vec{k_1}\cdot\vec{r_1}+\phi_1\right)\sin\left(\vec{k_2}\cdot\vec{r_2}+\phi_2\right)\right] \\
&= \frac{1}{2}E_{01}\,e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\,e^{-(\alpha_0 L+\Delta\alpha_2 L)}\left[\cos\left(\vec{k_1}\cdot\vec{r_1}+\phi_1\right)\cos\left(\vec{k_2}\cdot\vec{r_2}+\phi_2\right)\right. \\
&\quad \left.+\sin\left(\vec{k_1}\cdot\vec{r_1}+\phi_1\right)\sin\left(\vec{k_2}\cdot\vec{r_2}+\phi_2\right)\right] \\
&= \frac{1}{2}E_{01}\,e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\,e^{-(\alpha_0 L+\Delta\alpha_2 L)}\left[\cos\left(\vec{k_1}\cdot\vec{r_1}+\phi_1-\left(\vec{k_2}\cdot\vec{r_2}+\phi_2\right)\right)\right] \\
&= \frac{1}{2}E_{01}\,e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\,e^{-(\alpha_0 L+\Delta\alpha_2 L)}\left[\cos\left(\vec{k_1}\cdot\vec{r_1}+\phi_1-\vec{k_2}\cdot\vec{r_2}-\phi_2\right)\right] \\
&= \frac{1}{2}E_{01}\,e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\,e^{-(\alpha_0 L+\Delta\alpha_2 L)}\left[\cos\left(\vec{k_1}\cdot\vec{r_1}-\vec{k_2}\cdot\vec{r_2}+\phi_1-\phi_2\right)\right]
\end{aligned}
$$

Because silicon is a linear, isotropic medium,

$$\left|\vec{k_x}\right| = \frac{2\pi}{\lambda}$$
$$\left|\vec{r_x}\right| = L$$
$$\vec{k_x} \parallel \vec{r_x}$$
$$\vec{k_x} \cdot \vec{r_x} = \frac{2\pi L}{\lambda}, \lambda = \frac{\lambda_0}{n_{\text{eff}x}(V)}$$
$$= \frac{2\pi L n_{\text{eff}x}(V)}{\lambda_0}$$

The wavevector can then be computed using the above relations, and their definition inserted in the equation for the product of two electric fields, to include an explicit dependence on applied voltage.

$$\begin{aligned}
\therefore \vec{E_1}\vec{E_2} &= \frac{1}{2}E_{01}\, e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\, e^{-(\alpha_0 L+\Delta\alpha_2 L)}\cos\left[\frac{2\pi L n_{\text{eff}1}(V)}{\lambda_0} - \frac{2\pi L n_{\text{eff}2}(V)}{\lambda_0} + \phi_1 - \phi_2\right] \\
&= \frac{1}{2}E_{01}\, e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\, e^{-(\alpha_0 L+\Delta\alpha_2 L)}\cos\left[\frac{2\pi L}{\lambda_0}\left(n_{\text{eff}1}(V) - n_{\text{eff}2}(V)\right)\right. \\
&\quad \left. + \phi_1 - \phi_2\right], \phi_x = \frac{2\pi n_{\text{eff}x}\Delta n_{\text{eff}x}(V)L}{\lambda_0} \\
&= \frac{1}{2}E_{01}\, e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\, e^{-(\alpha_0 L+\Delta\alpha_2 L)}\cos\left[\frac{2\pi L}{\lambda_0}\left(n_{\text{eff}1}(V) - n_{\text{eff}2}(V)\right)\right. \\
&\quad \left. + \frac{2\pi n_{\text{eff}1}\Delta n_{\text{eff}1}(V)L}{\lambda_0} - \frac{2\pi n_{\text{eff}2}\Delta n_{\text{eff}2}(V)L}{\lambda_0}\right] \\
&= \frac{1}{2}E_{01}\, e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\, e^{-(\alpha_0 L+\Delta\alpha_2 L)}\cos\left[\frac{2\pi L}{\lambda_0}\left(n_{\text{eff}1}(V) - n_{\text{eff}2}(V)\right.\right. \\
&\quad \left.\left. + n_{\text{eff}1}(V)\Delta n_{\text{eff}1}(V)L - n_{\text{eff}2}(V)\Delta n_{\text{eff}2}(V)L\right)\right] \\
&= \frac{1}{2}E_{01}\, e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\, e^{-(\alpha_0 L+\Delta\alpha_2 L)}\cos\left[\frac{2\pi L}{\lambda_0}\left(n_{\text{eff}1}(V)\left(1 + \Delta n_{\text{eff}1}(V)L\right)\right.\right. \\
&\quad \left.\left. - n_{\text{eff}2}(V)\left(1 + \Delta n_{\text{eff}2}(V)L\right)\right)\right]
\end{aligned}$$

The final transmission can then be calculated as follows, assuming equal input electric field amplitudes, applied voltages, and FCA.

$$\therefore T = \frac{c\varepsilon_r\varepsilon_0}{n}\left[\frac{1}{2}E_{01}^2\, e^{-2(\alpha_0 L+\Delta\alpha_1 L)} + \frac{1}{2}E_{02}^2\, e^{-2(\alpha_0 L+\Delta\alpha_2 L)}\right.$$
$$+E_{01}\, e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_{02}\, e^{-(\alpha_0 L+\Delta\alpha_2 L)}\cos\left(\frac{2\pi L}{\lambda_0}\left(n_{\text{eff1}}(V)\,(1+\Delta n_{\text{eff1}}(V)L)\right.\right.$$
$$\left.\left.\left.- n_{\text{eff2}}(V)\,(1+\Delta n_{\text{eff2}}(V)L)\right)\right)\right]$$

\$ $E_{01} = E_{02} = E_0,\ n_{\text{eff1}}(V) = n_{\text{eff2}}(V) = n_{\text{eff}}(V)$

$$T = \frac{c\varepsilon_r\varepsilon_0}{n}\left[\frac{1}{2}E_0^2\, e^{-2(\alpha_0 L+\Delta\alpha_1 L)} + \frac{1}{2}E_0^2\, e^{-2(\alpha_0 L+\Delta\alpha_2 L)}\right.$$
$$\left.+E_0\, e^{-(\alpha_0 L+\Delta\alpha_1 L)}E_0\, e^{-(\alpha_0 L+\Delta\alpha_2 L)}\cos\left(\frac{2\pi n_{\text{eff}}(V)L}{\lambda_0}\left(\Delta n_{\text{eff1}}(V) - \Delta n_{\text{eff2}}(V)\right)\right)\right]$$
$$= \frac{c\varepsilon_r\varepsilon_0}{n}\left[\frac{1}{2}E_0^2\, e^{-2(\alpha_0 L+\Delta\alpha_1 L)} + \frac{1}{2}E_0^2\, e^{-2(\alpha_0 L+\Delta\alpha_2 L)}\right.$$
$$\left.+E_0^2\, e^{-(2\alpha_0 L+\Delta\alpha_1 L+\Delta\alpha_2 L)}\cos\left(\frac{2\pi n_{\text{eff}}(V)L}{\lambda_0}\left(\Delta n_{\text{eff1}}(V) - \Delta n_{\text{eff2}}(V)\right)\right)\right]$$

\$ $\Delta\alpha_1 = \Delta\alpha_2 = \Delta\alpha$

$$= \frac{c\varepsilon_r\varepsilon_0}{n}\left[E_0^2\, e^{-2(\alpha_0 L+\Delta\alpha L)} + E_0^2\, e^{-2(\alpha_0 L+\Delta\alpha L)}\cos\left(\frac{2\pi n_{\text{eff}}(V)L}{\lambda_0}\left(\Delta n_{\text{eff1}}(V) - \Delta n_{\text{eff2}}(V)\right)\right)\right]$$

Appendix C: Derivation of Impedance and Velocity Matching Relations

The following are derivations of Equations (2.50 - 2.53) from Equations (2.46 - 2.49), which relate the impedance and velocity matching requirements to the capacitance and inductance of the device's transmission line.

$$
\begin{aligned}
Z_L n_L &= \sqrt{\frac{L_U}{C_U + C_L}} c \sqrt{L_U (C_U + C_L)} \\
&= c \sqrt{\frac{L_U L_U (C_U + C_L)}{C_U + C_L}} \\
&= c \sqrt{\frac{L_U^2 (C_U + C_L)}{C_U + C_L}} \\
&= c \sqrt{L_U^2} \\
&= c L_U \\
\therefore L_U &= \frac{Z_L n_L}{c}
\end{aligned}
$$

$$
\begin{aligned}
Z_U n_U &= \sqrt{\frac{L_U}{C_U}} c \sqrt{L_U C_U} \\
&= c \sqrt{\frac{L_U L_U C_U}{C_U}} \\
&= c \sqrt{\frac{L_U^2 C_U}{C_U}} \\
&= c \sqrt{L_U^2} \\
&= c L_U, c L_U = Z_L n_L \\
\therefore Z_U n_U &= Z_L n_L
\end{aligned}
$$

$$\frac{n_L}{Z_L} = \frac{c\sqrt{L_U\left(C_U + C_L\right)}}{\sqrt{L_U / \left(C_U + C_L\right)}}$$

$$\frac{n_L}{cZ_L} = \frac{\sqrt{L_U\left(C_U + C_L\right)}}{\sqrt{L_U / \left(C_U + C_L\right)}}$$

$$= \sqrt{\frac{L_U\left(C_U + C_L\right)}{L_U / \left(C_U + C_L\right)}}$$

$$= \sqrt{\frac{L_U}{L_U}\left(C_U + C_L\right)^2}$$

$$= \sqrt{\left(C_U + C_L\right)^2}$$

$$= C_U + C_L$$

$$n_U = c\sqrt{L_U C_U}$$

$$n_U^2 = c^2 L_U C_U$$

$$\therefore C_U = \frac{n_U^2}{c^2 L_U}$$

$$C_U + C_L = \frac{n_L}{cZ_L}$$

$$C_L = \frac{n_L}{cZ_L} - C_U, C_U = \frac{n_U^2}{c^2 L_U}$$

$$= \frac{n_L}{cZ_L} - \frac{n_U^2}{c^2 L_U}, L_U = \frac{Z_L n_L}{c}$$

$$= \frac{n_L}{cZ_L} - \frac{n_U^2 c}{c^2 Z_L n_L}$$

$$= \frac{n_L}{cZ_L} - \frac{n_U^2}{cZ_L n_L}$$

$$= \frac{n_L}{cZ_L} * \frac{n_L}{n_L} - \frac{n_U^2}{cZ_L n_L}$$

$$= \frac{n_L^2}{cZ_L n_L} - \frac{n_U^2}{cZ_L n_L}$$

$$= \frac{n_L^2 - n_U^2}{cZ_L n_L}$$

Bibliography

[1] J. L. Hennessy and D. A. Patterson, *Computer Architecture: A Quantitative Approach*, 5th ed., K. Asanovic, Ed. Waltham, MA: Morgan Kaufmann, 2011.

[2] M. R. Watts, W. A. Zortman, D. C. Trotter, R. W. Young, and A. L. Lentine, "Vertical junction silicon microdisk modulators and switches," *Optics Communications*, vol. 19, no. 22, pp. 21 989–22 003, 2011. [Online]. Available: http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-19-22-21989&origin=search

[3] B. Jalali and S. Fathpour, "Silicon Photonics," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4600–4615, Dec. 2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4063407

[4] R. Soref, "The Past, Present, and Future of Silicon Photonics," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 6, pp. 1678–1687, Nov. 2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4026582

[5] N. Izhaky, M. T. Morse, S. Koehl, O. Cohen, D. Rubin, A. Barkai, G. Sarid, R. Cohen, and M. J. Paniccia, "Development of CMOS-Compatible Integrated Silicon Photonics Devices," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 6, pp. 1688–1698, Nov. 2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4026544

[6] M. R. Watts, "Silicon photonic components and networks," in *Optical Fiber Communication Conference and Exposition (OFC/NFOEC) and the National Fiber Optic Engineers Conference*. Los Angeles, California, USA: Institute of Electrical and Electronics Engineers (IEEE), 2011, p. OThM4. [Online]. Available: http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5875447&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D5875447

[7] D. Marris-Morini, L. Vivien, G. Rasigade, J.-M. Fedeli, E. Cassan, X. Le Roux, P. Crozat, S. Maine, A. Lupu, P. Lyan, P. Rivallin, M. Halbwax, and S. Laval, "Recent Progress in High-Speed Silicon-Based Optical Modulators," *Proceedings of the IEEE*, vol. 97, no. 7, pp. 1199–1215, Jul. 2009. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5075749

[8] O. Boyraz and B. Jalali, "Demonstration of directly modulated silicon Raman laser," *Optics Express*, vol. 13, no. 3, pp. 796–800, 2005. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=OPEX-13-3-796

[9] P. Dong, L. Chen, C. Xie, L. L. Buhl, and Y.-k. Chen, "50-Gb/s silicon quadrature phase-shift keying modulator," *Optics Express*,

vol. 20, no. 19, pp. 21 181–21 186, Aug. 2012. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=oe-20-19-21181

[10] Y. Silberberg, P. Perlmutter, and J. Baran, "Digital optical switch," *Applied Physics Letters*, vol. 51, no. 16, pp. 1230–1232, 1987. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4856064

[11] H.-W. Chen, Y.-h. Kuo, and J. E. Bowers, "High speed silicon modulators," in *OptoeElectronics and Communications Conference (OECC)*, no. 805. Hong Kong: Institute of Electrical and Electronics Engineers (IEEE), Jul. 2009, pp. 1–2. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5219965

[12] T. Baba, S. Akiyama, M. Imai, N. Hirayama, H. Takahashi, Y. Noguchi, T. Horikawa, and T. Usuki, "Efficient 50-Gb/s silicon microring modulator based on forward-biased pin diodes," in *IEEE International Conference on Group IV Photonics*. IEEE, Aug. 2013, pp. 69–70. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6644430

[13] B. Jalali, M. Paniccia, and G. Reed, "Silicon Photonics," *IEEE Microwave Magazine*, vol. 7, no. 3, pp. 58–68, Jun. 2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1638290

[14] B. Qi, P. Yu, Y. Li, Y. Hao, Q. Zhou, X. Jiang, and J. Yang, "Ultracompact Electrooptic Silicon Modulator With Horizontal Photonic Crystal Slotted Slab," *IEEE Photonics Technology Letters*, vol. 22, no. 10, pp. 724–726, May 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5432967

[15] M. Lipson, "Guiding, modulating, and emitting light on Silicon-challenges and opportunities," *Journal of Lightwave Technology*, vol. 23, no. 12, pp. 4222–4238, Dec. 2005. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1566750

[16] H. C. Nguyen, Y. Sakai, M. Shinkawa, N. Ishikura, and T. Baba, "Photonic Crystal Silicon Optical Modulators: Carrier-Injection and Depletion at 10 Gb/s," *IEEE Journal of Quantum Electronics*, vol. 48, no. 2, pp. 210–220, Feb. 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6068218

[17] P. E. Powers, *Fundamentals of Nonlinear Optics*, 1st ed. CRC Press, 2011.

[18] R. Soref and B. Bennett, "Electrooptical effects in silicon," *IEEE Journal of Quantum Electronics*, vol. 23, no. 1, pp. 123–129, Jan. 1987. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1073206

[19] D. Sun, S. Abdul-Majid, Z. Hu, R. Vandusen, I. Hasan, T. G. Tarr, and T. J. Hall, "An intrinsic limitation to silicon-on-insulator waveguide Mach-Zehnder interference-based electro-optic devices," in *Optoelectronic Devices and Integration III*, X. Zhang, H. Ming, and A. X. Wang, Eds., vol. 7847.

Beijing, China: SPIE, Nov. 2010, pp. 78 470P–78 470P–6. [Online]. Available: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1348074

[20] L. Gu, W. Jiang, X. Chen, and R. T. Chen, "Physical Mechanism of p-i-n-Diode-Based Photonic Crystal Silicon Electrooptic Modulators for Gigahertz Operation," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 14, no. 4, pp. 1132–1139, 2008. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4582372

[21] M. Watts, W. Zortman, D. Trotter, R. Young, and A. Lentine, "Low-Voltage, Compact, Depletion-Mode, Silicon MachZehnder Modulator," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 1, pp. 159–164, 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5395612

[22] S. Spector, C. Sorace, M. Geis, M. Grein, J. Yoon, T. Lyszczarz, E. Ippen, and F. Kartner, "Operation and Optimization of Silicon-Diode-Based Optical Modulators," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 1, pp. 165–172, 2010. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5406284

[23] D. J. Thomson, F. Y. Gardes, J.-M. Fedeli, S. Zlatanovic, Y. Hu, B. P. P. Kuo, E. Myslivets, N. Alic, S. Radic, G. Z. Mashanovich, and G. T. Reed, "50-Gb/s Silicon Optical Modulator," *IEEE Photonics Technology Letters*, vol. 24, no. 4, pp. 234–236, Feb. 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6086568

[24] J. Fujikata, S. Takahashi, M. Takahashi, and T. Horikawa, "High speed and highly efficient Si optical modulator with MOS junction for 1.55 m and 1.3 m wavelengths," in *IEEE International Conference on Group IV Photonics*. IEEE, Aug. 2013, pp. 65–66. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6644428

[25] R. S. Muller, T. I. Kamins, and M. Chan, *Device Electronics for Integrated Circuits*, 3rd ed., B. Zobrist, Ed. Wiley, 2002.

[26] M. Ziebell, D. Marris-Morini, G. Rasigade, P. Crozat, J.-M. Fedeli, E. Cassan, and L. Vivien, "Low-loss and high extinction ratio 40 Gbit/s optical modulator with self-aligned fabrication process," in *IEEE International Conference on Group IV Photonics*, vol. 12. San Diego, California, USA: Institute of Electrical and Electronics Engineers (IEEE), Aug. 2012, pp. 273–275. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6324157

[27] R. Ding, T. Baehr-Jones, Y. Liu, A. Ayazi, T. Pinguet, N. Harris, M. Streshinsky, P. Lee, Y. Zhang, A. E.-j. Lim, T.-y. Liow, S. H.-g. Teo, G.-q. Lo, and M. Hochberg, "A 25 Gb/s 400 fJ/bit silicon traveling-wave modulator," in *Optical Interconnects Conference*, vol. 4. Santa Fe, New Mexico, USA: Institute of

Electrical and Electronics Engineers (IEEE), May 2012, pp. 131–132. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6224411

[28] Oclaro Inc., "PowerLog AM-20; AM-40 Datasheet," 2009. [Online]. Available: http://dev.oclaro.com/datasheets/Oclaro_PowerLog_AM2040_v1_1.pdf

[29] D. Samara-Rubio, U. Keil, T. Franck, D. Hodge, D. Rubin, and R. Cohen, "Customized drive electronics to extend silicon optical modulators to 4 gb/s," *Journal of Lightwave Technology*, vol. 23, no. 12, pp. 4305–4314, Dec. 2005. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1566759

[30] J. D. Love, "Application of a low-loss criterion to optical waveguides and devices," *Optoelectronics, IEE Proceedings J*, vol. 136, no. 4, pp. 225–228, 1989. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=34703

[31] A. Biberman, M. J. Shaw, E. Timurdogan, J. B. Wright, and M. R. Watts, "Ultralow-loss silicon ring resonators," *Optics Letters*, vol. 37, no. 20, pp. 4236–4238, Oct. 2012. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=ol-37-20-4236

[32] A. Snyder and J. D. Love, *Optical Waveguide Theory*, 1st ed. Springer, 1983.

[33] J.-i. Takahashi, T. Tsuchizawa, T. Watanabe, and S.-i. Itabashi, "Oxidation-induced improvement in the sidewall morphology and cross-sectional profile of silicon wire waveguides," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 22, no. 5, p. 2522, 2004. [Online]. Available: http://link.aip.org/link/JVTBD9/v22/i5/p2522/s1&Agg=doi

[34] A. Khilo, M. a. Popović, M. Araghchini, and F. X. Kärtner, "Efficient planar fiber-to-chip coupler based on two-stage adiabatic evolution," *Optics Express*, vol. 18, no. 15, pp. 15 790–15 806, Jul. 2010. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=oe-18-15-15790

[35] C.-W. Liao, Y.-T. Yang, S.-W. Huang, and M.-C. M. Lee, "Fiber-Core-Matched Three-Dimensional Adiabatic Tapered Couplers for Integrated Photonic Devices," *Journal of Lightwave Technology*, vol. 29, no. 5, pp. 770–774, Mar. 2011. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5672761

[36] L. Cao, A. Elshaari, A. Aboketaf, and S. Preble, "Adiabatic Couplers in SOI Waveguides," in *Conference on Lasers and Electro-Optics (CLEO)*. San Jose, California, USA: Optical Society of America (OSA), 2010, p. CThAA2. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=CLEO-2010-CThAA2

[37] Y. Shani, C. Henry, R. Kistler, R. Kazarinov, and K. Orlowsky, "Integrated optic adiabatic devices on silicon," *IEEE Journal of Quantum Electronics*, vol. 27, no. 3, pp. 556–566, Mar. 1991. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=81363

[38] M. Mohamed, Z. Li, E. Dudley, X. Chen, L. Shang, W. Park, and A. Mickelson, "Adiabatic Couplers for Linear Power Division," in *Advanced Photonics*. Toronto, Ontario, Canada: Optical Society of America (OSA), 2011, p. IMD6. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=IPRSN-2011-IMD6

[39] T. Ramadan, R. Scarmozzino, and R. Osgood, "Adiabatic couplers: design rules and optimization," *Journal of Lightwave Technology*, vol. 16, no. 2, pp. 277–283, 1998. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=661021

[40] S. J. Spector, M. W. Geis, G.-R. Zhou, M. E. Grein, F. Gan, M. a. Popovic, J. U. Yoon, D. M. Lennon, E. P. Ippen, F. Z. Kärtner, and T. M. Lyszczarz, "CMOS-compatible dual-output silicon modulator for analog signal processing," *Optics Express*, vol. 16, no. 15, pp. 11 027–11 031, Jul. 2008. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=oe-16-15-11027

[41] G. Heise and R. Narevich, "Simple model for polarization sensitivity of silica waveguide Mach-Zehnder interferometer," *IEEE Photonics Technology Letters*, vol. 17, no. 10, pp. 2116–2118, Oct. 2005. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1512291

[42] G. V. Treyz, P. G. May, and J.-m. Halbout, "Silicon MachZehnder waveguide interferometers based on the plasma dispersion effect," *Applied Physics Letters*, vol. 59, no. 7, pp. 771–773, 1991. [Online]. Available: http://link.aip.org/link/APPLAB/v59/i7/p771/s1&Agg=doi

[43] Y. Wei, Y. Zhao, J. Yang, M. Wang, and X. Jiang, "Chirp characteristics of silicon Mach-Zehnder modulators under small-signal modulation," *Journal of Lightwave Technology*, vol. 29, no. 7, pp. 1011–1017, Apr. 2011. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5708158

[44] G. P. Agrawal, *Fiber-Optic Communication Systems*, 4th ed. Wiley, 2010.

[45] H.-W. Chen, J. D. Peters, and J. E. Bowers, "Forty Gb/s hybrid silicon Mach-Zehnder modulator with low chirp," *Optics Express*, vol. 19, no. 2, pp. 1455–1460, Jan. 2011. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=oe-19-2-1455

[46] D. Sun, Z. Hu, S. Abdul-Majid, R. Vandusen, Q. Zheng, I. Hasan, N. G. Tarr, S. Bidnyk, and T. J. Hall, "Limitation Factor Analysis for Silicon-on-Insulator Waveguide MachZehnder Interference-Based Electro-Optic Switch," *Journal of Lightwave Technology*, vol. 29, no. 17, pp. 2592–2600, Sep. 2011. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5948312

[47] H. Park, M. N. Sysak, H.-w. Chen, A. W. Fang, D. Liang, L. Liao, B. R. Koch, J. Bovington, Y. Tang, K. Wong, M. Jacob-Mitos, R. Jones, and J. E. Bowers, "Device and Integration Technology for

Silicon Photonic Transmitters," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 3, pp. 671–688, May 2011. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5722968

[48] R. Walker, "High-speed III-V semiconductor intensity modulators," *IEEE Journal of Quantum Electronics*, vol. 27, no. 3, pp. 654–667, Mar. 1991. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=81374

[49] J. Wang, H. Li, L. Li, A. Pang, Z. Sbeng, S. Zou, and F. Gan, "High-speed silicon Mach-Zehnder modulator operating beyond 40 Gb/s," in *Asia Communications and Photonics Conference*. Guangzhou, China: Institute of Electrical and Electronics Engineers (IEEE), 2012, pp. 3–5. [Online]. Available: http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6510934&url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel7%2F6504590%2F6510568%2F06510934.pdf%3Farnumber%3D6510934

[50] 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," *Optics Express*, vol. 19, no. 12, pp. 11 507–11 516, May 2011. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=oe-19-12-11507

[51] 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," *Optics Express*, vol. 20, no. 3, pp. 3209–3218, Jan. 2012. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=oe-20-3-3209

[52] G. Li, T. Mason, and P. Yu, "Analysis of Segmented Traveling-Wave Optical Modulators," *Journal of Lightwave Technology*, vol. 22, no. 7, pp. 1789–1796, Jul. 2004. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1310429

[53] H.-W. Chen, Y.-h. Kuo, and J. E. Bowers, "A high speed Mach-Zehnder silicon evanescent modulator using capacitively loaded traveling wave electrode," in *IEEE International Conference on Group IV Photonics*, vol. 6. San Francisco, California, USA: Institute of Electrical and Electronics Engineers (IEEE), Sep. 2009, pp. 247–249. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5338370

[54] C. T. DeRose, D. C. Trotter, W. A. Zortman, and M. R. Watts, "High speed travelling wave carrier depletion silicon Mach-Zehnder modulator," in *Optical Interconnects Conference*, vol. 6. Santa Fe, New Mexico, USA: Institute of Electrical and Electronics Engineers (IEEE), May 2012, pp. 135–136. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6224486

[55] D. C. Montgomery, *Design and Analysis of Experiments*, 8th ed. Hoboken, NJ: Wiley, 2012.

[56] Agilent Technologies, "Understanding the Fundamental Principles of Vector Network Analysis," USA, 2012.

[57] W. A. Zortman, A. L. Lentine, D. C. Trotter, and M. R. Watts, "Long-Distance Demonstration and Modeling of Low-Power Silicon Microdisk Modulators," *IEEE Photonics Technology Letters*, vol. 23, no. 12, pp. 819–821, Jun. 2011. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5742765

[58] W. Freude, R. Schmogrow, B. Nebendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, and J. Leuthold, "Quality metrics for optical signals: Eye diagram, Q-factor, OSNR, EVM and BER," in *International Conference on Transparent Optical Networks*, vol. 1. Coventry, England, UK: Institute of Electrical and Electronics Engineers (IEEE), 2012, pp. 1–4. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6254380

[59] P. Dong, L. Chen, and Y.-k. Chen, "High-speed low-voltage single-drive push-pull silicon Mach-Zehnder modulators," *Optics Express*, vol. 20, no. 6, pp. 6163–6169, Feb. 2012. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=oe-20-6-6163

[60] K. Solehmainen, M. Kapulainen, M. Harjanne, and T. Aalto, "Adiabatic and Multimode Interference Couplers on Silicon-on-Insulator," *IEEE Photonics Technology Letters*, vol. 18, no. 21, pp. 2287–2289, Nov. 2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1715403

[61] W. a. Zortman, M. R. Watts, D. C. Trotter, R. W. Young, and A. L. Lentine, "Low-power high-speed silicon microdisk modulators," in *Conference on Lasers and Electro-Optics (CLEO) and Quantum Electronics and Laser Science Conference (QELS)*. San Jose, California, USA: Institute of Electrical and Electronics Engineers (IEEE), 2010, pp. 1–2. [Online]. Available: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5500977

[62] M. P. Li, *Jitter, Noise, and Signal Integrity at High-Speed*, 1st ed. Upper Saddle River, NJ: Prentice Hall, 2008.

[63] D.-X. Xu, W. N. Ye, A. Bogdanov, D. Dalacu, A. Delage, P. Cheben, S. Janz, B. Lamontagne, and M.-J. Picard, "Design of polarization-insensitive components using geometrical and stress-induced birefringence in SOI waveguides," in *Optoelectronic Integration on Silicon II*, J. A. Kubby and G. E. Jabbour, Eds., vol. 5730, Mar. 2005, pp. 158–172. [Online]. Available: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1324698

[64] A. B. Fallahkhair, K. S. Li, and T. E. Murphy, "Waveguide Mode Solver," 2011. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/12734-waveguide-mode-solver

[65] S. Gevorgian, H. Berg, H. Jacobsson, and T. Lewin, "Basic parameters of coplanar-strip waveguides on multilayer dielectric/semiconductor substrates, part 1: high permittivity superstrates," *IEEE Microwave Magazine*, vol. 4, no. 2, pp. 60–70, Jun. 2003. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1201599

[66] N. Jaeger and Z. Lee, "Slow-wave electrode for use in compound semiconductor electrooptic modulators," *IEEE Journal of Quantum Electronics*, vol. 28, no. 8, pp. 1778–1784, 1992. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=142575

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
27–03–2014	Master's Thesis	Oct 2013–Mar 2014

4. TITLE AND SUBTITLE

Characterization and Performance Comparison of Low-Voltage, High-Speed, Push-Pull and traveling-wave Silicon Mach-Zehnder Modulators

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Latchu, Tristan A., Second Lieutenant, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB, OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENG-14-M-48

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Intentionally Left Blank

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

This work is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

The well-known power and memory walls are recognized as the current bottlenecks in computing performance, and with the increasing computational load of commonly run applications, it is necessary to find ways to alleviate the issues presented by the aforementioned bottlenecks. It is therefore necessary to not focus solely on extracting performance improvement by way of changes to the processing architecture, but rather by holistically improving the computing platform, namely the communications backbone. This work focuses on the characterization and performance comparison of two families of optical data modulators, both fundamentally Mach–Zehnder modulators (MZMs); namely, a modulator with a Push-Pull (PP) modulation scheme, and another modulator with a Traveling-Wave Electrode (TWE) design, prioritizing the effects on high-speed performance.

A number of operating conditions — temperature, wavelength, bias voltage, and bit rate — were varied to determine effects of modulator performance, measured in terms of DC performance characterization, bit error rate, electrical bandwidth, and power-penalty. Custom fixtures were designed and fabricated to achieve long-term experimental stability, and software was written to accomplish long term experimentation; the confluence of the two resulted in a wealth of data for use in performance comparison.

Despite the use of a push-pull modulation scheme, the devices using a traveling-wave electrode outperformed the push-pull modulators in almost all metrics, even at what was assumed to be a relatively low bit rate of 3 Gbps. This work then accentuates the importance of velocity and impedance matching, even at presumably low data rates, in spite of increased device fabrication complexity.

15. SUBJECT TERMS

Add four or five key words/phrases for indexing

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Maj Michael C. Pochet (ENG)
U	U	U	UU	256	19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636 x4393 michael.pochet@afit.edu