UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

New Computational Method to Estimate The Effect of Roughness On Scattering Loss and Its Implementation In a Hybrid Heterojunction Optical Modulator

by

Ehsan Jaberansary

Thesis for the degree of Doctor of Philosophy

May 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES
Electronics and Computer Science

Doctor of Philosophy

NEW COMPUTATIONAL METHOD TO ESTIMATE THE EFFECT OF ROUGHNESS ON SCATTERING LOSS AND ITS IMPLEMENTATION IN A HYBRID HETEROJUNCTION OPTICAL MODULATOR

by Ehsan Jaberansary

In the past year, there has been an advancement in the development of optical active and passive Silicon-On-Insulator for Photonic Integrated Circuits (PICs) applications. Following the continuous miniaturisation trend in this technology the study of the loss performance of these devices also become an attractive subject. Among the many components silicon based optical modulators are specifically important for low loss and high bandwidth short reach interconnects. The general aim of this thesis is to design and computationally investigate the performance of a multilayer hybrid silicon modulator. The design involves a heterojunction structure that operates based on plasma dispersion effect taking the effect of surface roughness and scattering loss into account. For this purpose a novel numerical approach is developed to estimate scattering loss due to the roughness in general waveguide structures and also the proposed silicon based optical waveguide modulator in this work. The applicability can be however extended towards wider range of optical waveguide based devices including multilayer configurations. The method is based on 2D Fourier transform technique that is widely used in Magneto Resonance Imaging technique.

Firstly, the effect of three forms of roughness is investigated in a general strip waveguide structure; isotropic, anisotropic and the mixture of isotropic and anisotropic. In each case the generated sidewall roughness is implemented in various SOI high contrast refractive index waveguides. The waveguide dimensions have been chosen to cover a large variety of waveguide sizes to evaluate the accuracy of the modelling technique. Two SOI waveguide samples have been fabricated. The first sample is 100 nm, 1200 nm and 1400 nm in width and 500 nm in height which are single mode in Mid Infra Red (MIR) region. The second sample contains waveguide with 220 nm height and 330 nm widths that is single mode in Near Infra Read (NIR) wavelength. Other dimensions are chosen from several published works. The calculated losses using FDTD show good agreement with all measured fabricated waveguide and the referred experimental works in the literature. The three dimensional model successfully explains the scattering loss dependence on the width of a high aspect ratio waveguide when the result is compared with a published work. While the measurement shows the loss reduces from 32 dB/cm to 0.8 dB/cm, the simulation results varies from 44 dB/cm to 1 dB/cm. The interdependence of scattering loss is also investigated against other theoretical approaches when the correlation length varies from 0 to 1000 nm. The relatively low aspect ratio waveguide is chosen to have a fixed dimension of 220 nm by 330 nm and 5 nm value.

In order to study the scattering loss caused by roughness the roughness model is applied in the proposed modulator structure. The design involves a multimode strip like p doped silicon material wrapped by transparent ZnO as a naturally n doped active material. This forms a pn heterojunction that is implemented in one arm of a Mach Zehnder interferometer. The switch is designed to operate at 1.55 μm and in depletion mode to avoid minority carrier life time effect in switching speed. The calculated capacitance switching speed of a pulse were less than 1 pf/cm and 90 ps respectively. The resistivity

is higher compared to a general form of heterojunction due to the relativity larger ZnO/Si contact area of the device. The phase shifter is implemented in Mach Zehnder structure to change from phase modulation to amplitude modulation using a MMI structure. The calculated extinction ratio was as high as 23.7 dB with the insertion loss of 2.5 dB. Further simulation results shows that the 100 *nm* change in the ZnO thickness can alter the effective index of refraction and loss performance of the devices. In an ideal situation, as the thickness increases from 50 nm to 150 *nm* the loss changes from 2 to 8 dB/cm for TE mode. The involvement of sidewall roughness results higher insertion loss by at least by 0.2 dB when the rms of the sidewall roughness increases by 7 nm. As the ZnO coating thickness increases, the roughness effect is counterbalanced by almost 50 %.

Contents

List of Tables

Declaration of Authorship

I, Ehsan Jaberansary , declare that the thesis entitled *New Computational Method to Estimate The Effect of Roughness On Scattering Loss and Its Implementation In a Hybrid Heterojunction Optical Modulator* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission

Signed:.....

Date:.....

Acknowledgements

I would like to express gratitude towards Dr. Harold Chong for his supervision and companionship through the years of my PhD. and all his generous supports for his continued encouragements and helpful advices.

Very important thanks go to Dr. Martin Charlton, Prof D. Bagnall, Prof. P. Ashburn amd Dr. G. Z. Mashanovich whom their constructive comments helped me towords finishing my PhD.

I am also very thankful to the Nano Research Group at Southampton for providing me with a generous studentship for my PhD. and to my colleagues at the group who were always encouraging during the years of my PhD.

Thank you Taha. B. Masaud, M. M. Milosevic, M. Nedeljkovic, Nima Kalhor, Hamed Shahidi, Adel Ghiass, Mehdi Banakar, and Mohammad Reza Sarshogh.

It is more than appropriate that I express most heartfelt gratitude towards my family for their truly unwavering support in all senses of the matter.

Chapter 1

Introduction

1.1 Photonic Integrated Circuits

Within the last decade, there has been a continuing strong interest in Photonic Integrated Circuits (PICs) research. This is due to the need for low power consumption, reliability and high density integration. It is also believed by many researchers, that the next technological step towards high bandwidth optical communication systems can be achieved through successful miniaturisation of optical integrated device.

Photonic integration can be either hybrid or monolithic. In monolithic integration many optical devices are consolidated into a single photonic material. The devices are built into a common substrate. However, all-optical technologies at the current stage are more expensive, more complex, or inefficient in comparison with electronic IC. Unlike electronics where transistors are defined as the elementary component, there is not such a dominant device at photonic level. Hence, the compatibility of optical networks with electronic components that are purely relegated to the edges of network can maximise system functionality and service flexibility. This is normally realised by considering hybrid integration where multiple optical devices are assembled into a single package with associated electronic devices that are interconnected to each other by electronic and/or optical links.

For instance, among the rapid technological development the need for high-bandwidth interconnects is a highlighted topic to enhance computing capability. Interconnects in current microelectronic industries are conventionally fabricated from copper [1]. As the demand for high density integration increases, the physical limitation of the low cost electrical interconnects starts to emerge. In particular for higher bandwidth where large amount of data is transferred, current electrical interconnects experience a high cross-talk effect [2]. The cross talk in interconnects happened when the transmitted signal in one part of the network creates an undesired effect in another part. The problem

is more significant in high density integrated circuits. Hence other alternatives have to be utilized. Optical interconnect is a potential candidate as it can provide dispersion free communications and minimum cross talk effects within enormous bandwidth. The key component of such optical interconnect is the optical modulator. The performance of integrated optical modulators has been improved dramatically in recent years; in particular the bandwidth has increased from MHz to the multi GHz regime within half a decade. However the demand of optical interconnect is significant, and there are yet many unanswered questions. Minimising some of the performance metrics such as the energy per bit and device footprint are as important as maximising some of the others such as bandwidth and modulation depth. All these have to be achieved having an acceptable thermal stability, optical spectral width and using CMOS compatible fabrication processes. Cost effective technologies will therefore be required in order to adopt the integration of optical modulators in electronic industry. To meet future technology targets, there are however other fundamental concerns in the integration of optical interconnects and relevant components such as modulators in a chip either with or without electronic devices which can be categorised as follows.

1.2 Major Concerns

1.2.1 Power Consumption

High power consumption is still a problem (97 fJ by 2015) for electrical interconnects. Nevertheless it is argued that the high power consumption can be compensated by many advantages that optical interconnects and other photonic devices may offer [2]. In addition, the power consumption of current electrical interconnects are already higher by at least more than 5 times, which can justify the applicability of optical interconnects in current technologies.

1.2.2 Size Reduction

Optical structures have been conventionally used as the discrete building blocks of long distance telecommunications. Clearly using such discrete components are not feasible for high density integrated chips. As well as the continuous shrinking of electronic devices in a chip, the photonic devices should also be scaled down towards submicron region to provide comparable density requirements. Good level of integrations have already been demonstrated using silica on silicon and silicon oxy-nitride technology [3]. Further improvement in scaling photonic devices is possible with silicon-on-insulator (SOI) technology, where high contrast waveguides can be shrunk down to a submicron cross sections. Following the miniaturization trend, there would be limitations arisen by

physical waveguide parameters which can not be negligible any more and can be very significant at nanoscales.

One of the most fundamental restrictions facing photonic devices is optical propagating loss, as the size is reduced with respect to wavelength. The increased propagation losses are due to waveguide roughness and insertion losses. The finite resolution of the lithography may result additional coupling issues between waveguide gaps [1]. The loss penalty would be even higher for the high contrast waveguide.

Among of all, there have been several theoretical and experimental studies showing that surface roughness is the main source for high propagation losses. Furthermore, surface roughness is also responsible for high bending losses that is measured to be 1 dB per 90 degree bend which itself limits the applicability for high density PICs [4–6].

1.2.3 Compatibility With CMOS Technology

CMOS has been known as the most successful technology in electronics industry. Any technology that can be considered in highly compact hybrid or monolithic photonic circuit integration must be compatible with CMOS technology.

1.3 Silicon Photonics: A Promising Solution

The leading candidate technology for optical modulators is silicon photonics at relatively wide wavelengths from 1300 nm to 1630 nm. Silicon photonics is a relatively new platform in optoelectronic industry that attracts many attentions by offering optimum solutions for the mentioned concerns. The technology can perform functions which were previously preferred to be done by employing expensive materials based on III-V semiconductors. There are several factors that make silicon more attractive in comparison with traditional III-V materials. Due to the relatively high refractive index of silicon at NIR and MIR wavelength range, light confinement can be achieved in submicron dimensions. The low cost nature of silicon is another advantage; Silicon has a dominant presence in microelectronic industry and CMOS technology which brings its potential in electronics and photonics integration into attention.

However, Si is inherently considered as an optically passive material. It can only be used for active application if it is doped. The penalty would be more absorption loss caused by implanted free carriers. In order to resolve this issue and extend the functionality of Si waveguide devices complimentary materials should be involved to reform new material systems that are able to work appropriately with Si. The employment of these cost effective and technologically compatible materials such as ZnO and GaAs with silicon in the structure of optical modulators can fulfil the functionality of interconnects. ZnO

as a Si friendly material can be integrated into silicon photonics using different methods such as Atomic Layer Deposition [1]. The refractive index is 1.9 which is similar to that provided by SiO_2 and ZnO is transparent at NIR wavelengths. Furthermore, it is naturally n-doped and hence does not need a complex ion implantation.

Si also shows a low absorptive behaviour at NIR wavelengths. The roughness of the Si surface is however very much fabrication dependent. Recent technological advances allow to avoid the top surface irregularities in certain waveguides and geometries, but not yet appropriately on sidewall surfaces [7]. Several successful studies show different oxidation process of rough Si surfaces can be used to smooth the sidewalls [8–10]. However, this method effectively reduced the Si waveguide dimensions, which is inappropriate for PIC design. Laser annealing also applied for certain materials such as ZnO waveguides. Nevertheless the applicability of this method is very limited and prohibitive. Hence the roughness is still considered as an important topic to be further studied in photonic devices and IPCs.

The development of various low loss silicon based optical interconnect components such as laser light source, waveguides, modulator and optical detector is on-going. In particular, there is still much to explore regarding loss performance caused by scattering with roughness in devices such as fast optical switches and multiplexers. The optimum goal is to achieve compact integration of these building blocks either in a hybrid form or monolithically in a single chip by shrinking down the size of these components towards nanoscale dimensions.

1.4 Objective of Research

The aim of this thesis is to develop and computationally investigate novel silicon based optical modulator design with the contribution of ZnO as an active complementary material. A comprehensive modelling technique is also introduced to accurately estimate roughness effect in the performance of the proposed optical modulator and more generally in wide range of other structures, from simple waveguides to more complex devices. The adoption of this technique will also lead to study of:

1. The loss performance of the proposed modulator structure, any general waveguide and other waveguide based photonic devices.
2. The validity of possible solutions which have been or may be later offered to reduce the scattering loss in the modulator and possibly other waveguide based devices.
3. The geometrical effect of roughness on polarization issues.

This approach is developed based on a technique used in Magneto Resonance Imaging which applies two dimension Fourier transform to model roughness from the provided

statistical data. In MRI, the magnetic resonance signals contain the information about the amount of the contribution of the frequencies within a certain range that corresponds to the unknown object. The inspiration of using this approach in the case of roughness is arising from the fact that in MRI the actual detailed image of unknown and invisible object can be figured out from the frequency properties of the MR signals. Similarly the roughness as a random and unknown property can be derived if there is enough information about the frequencies correspond to the roughness. The roughness model is then applied to the optical modulator that is a heterojunction Si based diode utilizing ZnO as a transparent metal oxide material. This provides the possibility of computationally investigating roughness effect on the loss performance of the device.

1.5 Outline of This Report

The outline of this thesis is managed in the following fashion.

Chapter 2 delivers a historical literature review of the main optical modulation methods. The applicability of these approaches in silicon photonic integration will be discussed. Furthermore recent advances in silicon based photonic modulators are also referred, in particular for free carrier plasma dispersion based silicon modulators.

Chapter 3 focuses on the term loss in optical waveguides and provide essential background which was used by the author to follow his research.

Chapter 4 explains how the roughness model is developed and evaluated in comparison with experimental and theoretical results and also a number of other published works. The method enables researchers to extend their knowledge of loss performance for optical integrated devices.

Chapter 5 a proposed optical modulator design will be introduced and will be analytically discussed using different simulation software. The effect of scattering loss in the design performance will be also evaluated at the end.

Chapter 6 a conclusion of the research covered in this report is given with a summary of achievements. Ideas for future work will also be discussed.

Chapter 2

Optical Modulation Overview

2.1 Terminology

Modulators have a wide variety of applications include wideband analog optical communication systems, signal processing, beam deflection, frequency stabilization, Q-switching of lasers, and other integrated optical systems [11]. More importantly optical modulation is one of the fundamental functionalities required in a photonic circuit for any interconnect solution. Optical modulation in a waveguide is a process that the fundamental characteristics of light(amplitude, phase or polarization) are modulated by external or internal actuation of waveguide. Implementing semiconductor nanoscales waveguides in the modulator structure can boost the efficiency, leads to higher stability, speed response, and highly compact system [11].

In the system level, there are three fundamental schemes of optical modulations. It can be continuous that is normally called intensity modulation in optics and is similar to analogue modulation, with the only difference is in the nature of the optical carrier signal. The other alternatives are digital (discrete) and pulse width (discrete). The modulated signal is composed of a series of pulses with constant amplitude. While in pulse width modulation, the generated pulses have different pulse width in order to emphasize the modulation, in discrete modulation all pulses have the same width and switching between the two sates represents the modulation. The study of these schemes in system level is however beyond the scope of this thesis. Modulation is most notably attractive in silicon photonics, as the leading candidate because of the unique combination of potential low cost, performance enhancements due to the prospect of both photonic, and electronic/photonic integration, as well as compatibility with the world's most successful electronic technology, CMOS. However, whilst various silicon modulators design continue to be developed, other solutions include hybrid combinations of different materials to fulfil the functions that silicon on its own is, as yet, unable to deliver.

The general performance metrics of a modulator device are generally assessed based on several specifications as summarized in the next section.

2.2 Modulation Specifications

Table 2.1 shows the key performance metrics of typical optical modulator or switch. There are clear trade-offs between these specifications. For example in Mach Zehnder configuration (discussed in 5), reducing the effective length of a modulator also reduces its power consumption and on chip losses, but it also reduces the depth of the modulation as the two waves are not fully anti-phase. Therefore in order to achieve the most appropriate outcome that optimistically serves certain application, it is important to carefully optimize the modulator performance.

Specification	Definition	State of the art requirement
Modulation speed	The rate at which the data can be carried by a modulator	High>10 GHz
Modulation Bandwidth	The frequency at which the modulation is reduce to 50% of its maximum value.	Wide
Depth of Modulation Extinction ratio	The intensity ratio of the maximum transmission state to the minimum transmission state.	Large>7 dB
Optical Bandwidth	The useful operational wavelength range of the modulator.	Wide>9 nm
Loss	The optical power that is lost when the modulator is added to a photonic circuit, consisting reflection, absorption and scattering.	Low<4 dB
Length Efficiency (footprint)	The required effective length of the modulator for a complete transition between a maximum and minimum of the optical transmission.	Small<2000 μm
Power Consumption	The energy expended in producing each bit of data (energy/power per bit).	Low< 10 fJ/bit

Table 2.1: Modulator General Specifications

2.3 Modulation Methods in Silicon Photonics

The easiest way to obtain modulation of a light beam is direct modulation which the intensity of light is modulated through varying the current driving the light source itself. This causes the refractive index of the cavity to change. Because the lasing wavelength is determined from the feedback condition in the lasing cavity, the change of the refractive index also results broadening the emitting frequency (chirping effect). Although most optical detectors respond to this sort of modulation, it is normally avoided for laser

diodes where narrow linewidth is required due to a high bandwidth chirping effect [12]. In contrast, external light modulators become more attractive since they are universally applicable to a large variety of materials. In addition modulation speeds can be higher in external modulations and direct phase modulation is also possible. Furthermore, individual modulators can be used to feed the generated light by a single light source into multiple channels and hence the total power budget of the system can be reduced [13]. Also optical isolation and wavelength stabilization (which is required to overcome wavelength drift due to the temperature changes) can be only performed once for the whole system [13].

External optical modulators can be realized by varying one or more fundamental properties of an optical wave. These include power, irradiance, amplitude, frequency (wavelength), phase and polarization. At optical frequencies however, not all of these possibilities are equally suitable and their applicability strongly depends on the detection procedure [12].

External modulators can be classified in two major groups according to the change in a parameter of the optical beam being modulated. These are electro-absorptive modulators and electro-refractive modulators. The alterations may be performed by inducing certain changes in the properties of the participating materials that are used in the waveguide structure. In absorptive modulators absorption coefficient of the material is changed while in refractive modulators, manipulation of the light beam is based on the change that occurs in refractive index of the material, normally in accordance of an external electric field. Absorptive modulator techniques can be realized using Franz-Keldysh [14] or Quantum Confined Stark effect [15].

Franz-Keldysh effect usually requires massive applied voltage which is limiting its application in conventional electronics [16]. The QCSE has stronger changes in absorption coefficient compared to Franz Keldysh effect by shifting the absorption edge to lower photon energies ($> 0.1eV$). However, using these effects in specifically silicon platform devices requires the introduction of a new material(s), which may complicate device designs. In the literature, III-V semiconductor materials such as GaAs with AlGaAs barriers or InGaAs with InP barriers are among the most being studied combinations so far [15].

In electro-refractive modulators the refractive index change is normally produced through the change in applied external electric field (voltage). This is a preferred way of implementing modulation because it has little or no current flow (and hence low power) and a fast response time. Changes as high as...in refractive index are reported in the literature. Unlike electro-absorptive modulators which tend to change the light amplitude, the immediate effect of all refractive based modulator operation is change of the phase of a light beam. This can be converted into amplitude modulation using different mechanism. The most common approach is to use a Mach Zehnder Interferometer (MZI) and

destructive superposition of the light beams progressing through each wing. The MZI can be configured in symmetric or asymmetric arrangement depending on the required wavelength and integration conditions (Figure 2.1). Other alternative approach is to convert phase modulation to intensity modulation through using a resonance structure, such as ring resonator.

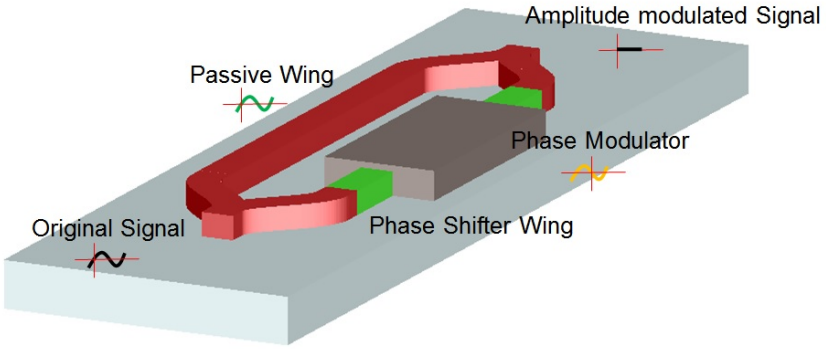

Figure 2.1: Basic layout of a symmetric Mach Zehnder interferometer structure.

ω_n Typically an effective index change of at least in the order of 10^{-4} is required to make the device viable at reasonable dimensions [17]. Refractive modulation techniques can be achieved by taking any of electro-optic [11], magneto optic [11], acousto-optic [18] and thermo-optic [17] effects into account.

Although electro-optical modulation has been demonstrated using silicon in inappropriately large structures (micro or mili scales), modulation is inefficient due to the centro-symmetric crystal structure and poor electro-optical features of silicon [19, 20]. The magneto-optical modulators also have considerably limited applications, since electric field is easier to generate [11]. In the literature, silicon has been highlighted as a material with the highest thermo-optic characteristics among the other common thermo-optical materials [21]. Nevertheless the application of this effect in communications and electronics is also restricted due to the slow response time and narrow provided bandwidth. Commonly the application is investigated in in-line optical variable attenuator, where the required response time is in the order of kilohertz, as well as in sensors technology [21]. Several piezoelectric materials such as lead zirconium titanate (PZT), zinc oxide (ZnO), polyalkene/polyvinylidene fluoride(PVF2), and vinylidene difluoridetrifluoroethylene (VDF-RFE) have been investigated to realize acousto optic modulation in various silicon based configurations (cylinder, clip-on, and resonators) [17, 22]. These modulators could reach the bandwidth as high as 1GHz which is still far away from the current state of the art plasma dispersion based modulators.

The plasma dispersion effect (free carrier concentration changes) is a common technique for both refractive and absorptive modulation. The method is one of the widely used techniques in particular for materials such as silicon. Silicon has a weak electro-optic

effect because of its centro-symmetric crystal structure. Therefore free carrier plasma dispersion effect is a preferred method to use in Si based modulators. However, in silicon, the optical absorption due to the plasma dispersion effect is relatively weak [23]. Consequently, electro-absorptive modulators, constructed from these materials may require an impractically long active waveguide. Therefore electro-refractive based modulation using plasma dispersion effect can provide higher efficiency in materials such as Si.

2.3.1 Plasma Dispersion Effect

Free carriers that can move freely in a conduction band absorb the energy of photons and jump from a filled region band below the Fermi level to an unoccupied region above the Fermi level (in the same band) [24]. As mentioned before, this is different from intra band absorption in semiconductors where the absorption is due to the electron that is being raised from a valence band to a conduction band. If the concentration of free carrier in a semiconductor is therefore changed the absorption characteristic consequently differs from its nominal state. The amount of variation is however very much dependent on material properties and may be considerably different from one to another.

Carrier concentrations in a semiconductor or dielectric waveguides can be controlled by injection, depletion or accumulation of free electrons/holes due to an external applied voltage and its resulting electric field. For silicon waveguides, relatively high amount of free carrier density variations are required to cause an effective change in absorption coefficients. However this can be optimized against the interaction length of the modulator which is the length required for the propagating wave to interact with the newly distributed carriers. Hence normally the carrier density change has a parallel effect in refractive indices of different materials as well as their absorptive behaviour. This is described by the Drude-Lorentz equations that relate the concentration of electrons and holes to the absorption coefficient ($\Delta\alpha$) and refractive index (Δn) [7, 24]:

$$\Delta n = \frac{e^3\lambda_0^2}{4\pi^2 c 3\epsilon_0 n}\left[\frac{\Delta N_e}{\mu_e (m_{ce}^*)^2} + \frac{\Delta N_h}{\mu_h (m_{ch}^*)^2}\right] \tag{2.1}$$

$$\Delta\alpha = \frac{-e^2\lambda_0^2}{8\pi^2 c 3\epsilon_0 n}\left[\frac{\Delta N_e}{(m_{ce}^*)} + \frac{\Delta N_h}{(m_{ch}^*)}\right] \tag{2.2}$$

where c is the velocity of light in vacuum, λ_0 is free space wavelengthϵ_0 is permittivity of free space,μ_e is the electron mobility,μ_h is the hole mobility, m_{ch} is the effective mass of holes , m_{ce} is the effective mass of electrons, ΔN_e and ΔN_h are the change of free electrons concentration and the change of free holes concentration respectively.

2.3.2 Dispersion Effect in Silicon Based Modulators

In this work the study of the effectiveness of Si and compatible materials is particularly desired because of the unique advantages that Si and Si technology offer. Soref and Bennett explained that the relative efficiency of the electric Kerr effect and the Franz-Keldysh effect are small in Si and the Pockels effect is almost negligible [13]. They concluded that the only remaining candidate for high speed modulation through using Plasma dispersion effect in Si. In particular, they focused on the communications wavelengths of 1.3 μm and 1.55 μm and worked out the following empirical expressions which were in a good agreement with Drude- Lorenz model [13].

For 1.55 μm:

$$\Delta n = \Delta n_e + \Delta n_h = 8.8 \times 10^{-22} \Delta N_e + 8.5 \times 10^{-18} (\Delta N_h)^{0.8} \tag{2.3}$$

$$\Delta \alpha = \Delta \alpha_e + \Delta \alpha_h = 8.5 \times 10^{-18} \Delta N_e + 6.0 \times 10^{-18} (\Delta N_h) \tag{2.4}$$

And for 1.3 μm:

$$\Delta n = \Delta n_e + \Delta n_h = 6.2 \times 10^{-22} \Delta N_e + 6.0 \times 10^{-18} (\Delta N_h)^{0.8} \tag{2.5}$$

$$\Delta \alpha = \Delta \alpha_e + \Delta \alpha_h = 6.0 \times 10^{-18} \Delta N_e + 4.0 \times 10^{-18} (\Delta N_h) \tag{2.6}$$

Once again from Equations 2.3, 2.4, 2.5 and 2.6 it can be seen that any changes in free carrier density of the waveguide can cause a change in both refractive and absorptive characteristics of the waveguide. In these expression it can be also observed that the effect of hole concentration change refractive index of Si is much higher than its effect caused by electron concentration change. The reason is because the effective mass of hole is larger than electron. Also Electron mobility is higher than hole. However electron and hole concentration change both have comparable effects on absorption behaviour of Si. It is also important to remember for a functional modulator it is desired that only one feature (either refractive or absorptive) of the waveguide effectively changes at a time. If the electro refractive modulation is used for example the carrier concentration should be chosen appropriately so that the absorptive behaviour ideally remains unaffected or is altered minimally. Otherwise extra losses are added due to carrier absorption which can significantly disorder phase modulation process. In addition, it will undesirably distract the purposed destructive combination of phase shifted wave and original wave, hence lower modulation depth is achieved. It is therefore critically important to clearly decide whether the modulator should operate as a phase modulator (electro refractive) implemented by refractive index change or as an intensity modulator (electro-absorptive) realized by the variation in absorption coefficient. Based on that, the best trade off and

optimization of carrier distribution parameters has to be made in order to reach the requirements of the quality modulation.

2.4 Plasma Dispersion effect in literature

Based on the three mechanisms, modulation in Si waveguide has been demonstrated in a large variety of structures. In this section, I will summarize the most successful works that are widely reported and referred in the literature.

2.4.1 Phase modulators based on carrier injection

Carrier injection can be realized by applying forward bias across either a p-n junction or through using highly doped p and n regions that are separated by an intrinsic or a lightly doped n region. Applying forward bias voltage across the device causes free electrons and holes to be injected into the waveguide region. The drawback in such a mechanism is the diminished switching speed due to carrier recombination in intrinsic region of the waveguide [13].

The first free carrier injection based modulator in Si was proposed by Soref and Bennett [25]. A p^+nn^+ Si rib single mode waveguide was realised which can induce a π radian phase shift is less than 1 μm length. In the first approximation it has been reported by the authors that the modulator is polarization independent [25].

The first theoretically analyzed phase modulators using dual gated metal-oxide semiconductor field-effect transistors is reported in the late 1980s [20, 26]. The injected charge is controlled by the junction field effect where the depletion width controls the effective cross-sectional area of the conducting channel.

Another Si phase modulator is realized through implying Mach-Zehnder interferometer configuration by Treyz at IBM research division as shown in Figure 2.2 [27]. The operation is multimode and characterized at the wavelength of 1.3 μm based on the carrier injection into the p type guiding area by applying a forward bias.

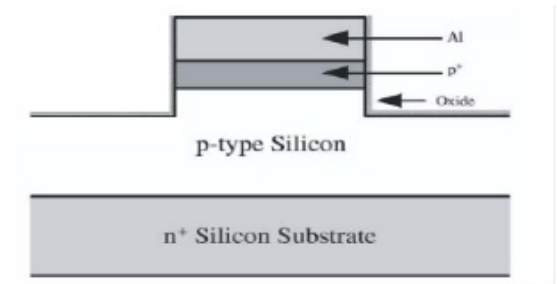

Figure 2.2: Schematic diagram of Treys model [28]

The device has relatively small response time (of about 50 ns), but experiences a high loss in comparison with Soref and Bennetts design, due to the rather weak optical confinement in the vertical direction [7]. The resulting absorption of the guided modes by n doped Si substrate is however eliminated as the fabrication of SOI based structures advanced. The vertical fabrication remained challenging, since substrate contacts could not be included [7]. Based on Treyz demonstration several more efficient vertical structures have been reported [27, 29, 30]. Works carried out by considering three terminal structures as suggested firstly by Reed at the University of Surrey [31,32]. Three terminal modulator geometry in general contains a rib waveguide which is configured in a form of a p-i-n diode as shown in Figure 2.3. This device operates in injection mode. One positive terminal is connected to the highly p doped region on top of the rib waveguide surface while the other two negative (or zero biased) terminals are configured on the horizontal surfaces of etched regions on both sides of the rib. Application a forward-bias voltage to the p-i-n structure leads to dual injection of both holes and electrons into the central region of the waveguide.

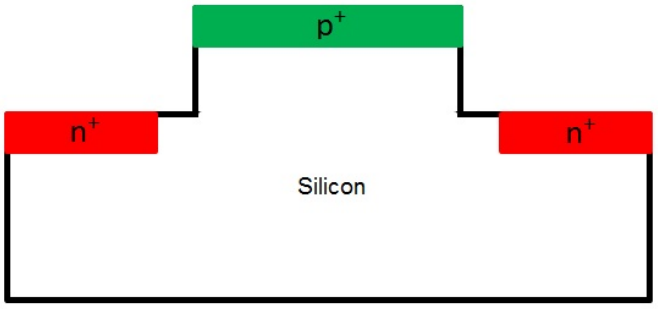

Figure 2.3: General form of a three terminal modulator structure

Several similar topologies have been theoretically and experimentally investigated in different dimensions where their design differently engineered to observe the structural effects on the performance. [33–37].

For example in 1997 a three- terminal modulator is reported by Cutolo using single mode low-loss SOI waveguide [33]. The author has however concluded that the three-terminal electronic structure is significantly more efficient than two-terminal pin diode based modulators which are also predicted previously by Reed and Tang in [31,32]. Two years later, this design has been developed by Hewitt and Reed to achieve the lowest predicted (for such a large structure) current at that time to achieve π phase shift [34,35]. They have shown that the three-terminal devices require less current (2.8 mA versus 8 mA) and are relatively faster (29 ns vs. 39 ns) than two-terminal designs for the same injection concentration [35]. This is clearly due to the more efficient carrier injection in three-terminal devices in comparison with general two terminal models. In contrast however, the three-terminal devices can suffer from extra optical attenuation caused by

the added doping contact at the top of the waveguide. This drawback could be resolved if an alternative is used, such as indium tin oxide, aluminium zinc oxide and zinc oxide that exhibit a better transparency.

Hewitt and Reed also add a new feature to their p-i-n phase modulator design [35] in order to improve the carrier confinement in the waveguide area. They have successfully applied trench isolation and report if this extends laterally on either side of a modulator it can improve both the DC characteristics by up to 74% and transient device performances by up to 18% [38]. This is shown in Figure 2.4a. An alternative method represented in Figure 2.4b was established by Day [36].

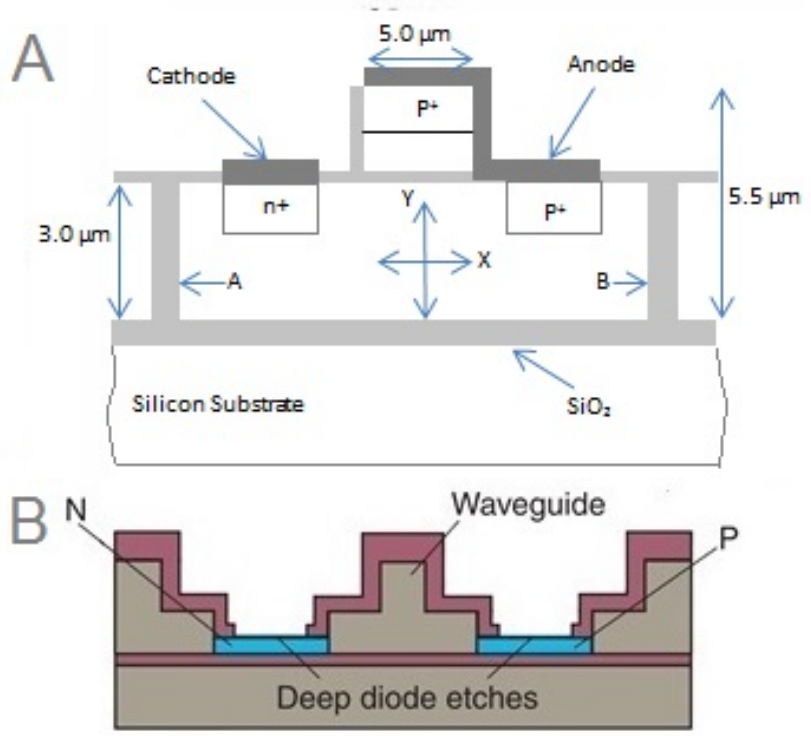

Figure 2.4: Days three terminal modulator. The regions A and B that are illustrated here are etched trenches in the intrinsic silicon [38]. b) Introduction of deep diode etch over standard doped structure by Day [36]

The argument here is to keep all injected carriers in the Si in the region between the side doped contacts for better carrier confinement. This is unlike the case where the doped regions are located in the surface of the slab region on either side of the waveguide [36]. They could improve both DC and transient performance in the order of magnitudes, although the size of waveguide was relatively large (4 μm in height) [36]. By reducing the waveguide dimensions that is shown in Figure 2.4b, Whiteman purposed to reduce the physical extent of the optical mode [37]. This allows the doping contact regions to be closer before an overlap with the optical mode occurs [37]. Hence the carriers will be injected and depleted from the optical region at higher speed. The effect has

comprehensively improved the performance from 85 mW to 50 mW power consumption and 1.7 MHz to 10 MHz bandwidth [36, 37].

In 2004, a low-loss single-mode optical phase modulator based on SOI was modelled by Png and Reed [39]. Although the reported work is purely theoretical, the bandwidth extends to the GHz region due to the size reduction of the device and optimized doping profile in the n+ regions to maximize injection efficiency (Figure 2.5). However a further slight improvement has been illustrated (3%) if the edges of the rib waveguide are rounded. This may occur within the fabrication process and helps to improve the carrier distribution profile.

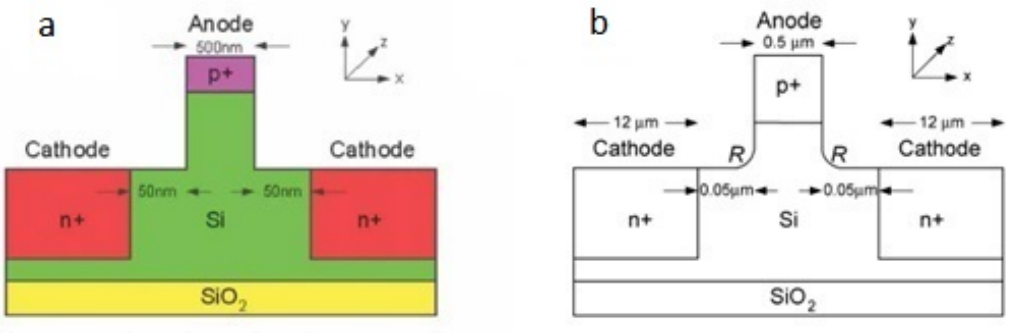

Figure 2.5: Theoretical model of a three terminal modulator design by Png and Reed [39].

Both the length and cross section of three terminal devices cannot be reduced indefinitely since this reduction posed excessive insertion loss greater cost and power consumption. Having additional metal contact in three terminal structures can induce further absorption loss or may require more complicated fabrication steps. Furthermore as discussed earlier the free carrier density and refractive index cannot be varied quite generously since it may cause the excessive absorptive characteristic to change too. With three terminal structures the effective confinement of the carrier is also challenging. Consequently the speed, length and the size of such modulator has been restricted so far [26]. However in 2005 Xu offered a new configuration to overcome this difficulty by using micro-ring resonant structures [19]. The schematic of this layout is illustrated in Figure 2.6. The authors introduced a p-i-n diode based carrier injection scheme into submicrometer waveguide based ring resonator structures, and achieved an initial data rate of 1.5 $Gbits^1$. The rings had diameters of 12 μm, providing the device size much smaller than any MZI-based modulator reported so far. The authors later reported improved results of 16 $Gbits^1$ with an extinction ratio of 8 dB using a pre-emphasis driving signal [40]. This method of modulation was also used by Green et al for a MZI based modulator, which achieved a data rate of 10 $Gbits^1$ for a device length of 200 μm [41]. The use of pre-emphasis drive signals in silicon modulator was first proposed by Ping

and Reed in 2004 [39]. This involves manipulation of the drive signal shape to reduce state transition times [39].

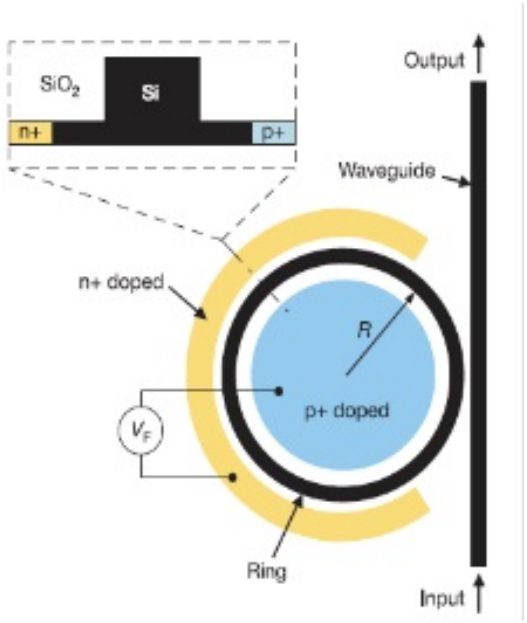

Figure 2.6: Schematic of the ring resonator-based modulator designed by Xu. The inset shows the cross-section of the ring [19].

Although significant length is saved with resonant devices, the resulting bandwidth is often reduced from more than 20 *nm* in the case of an MZI-based device to around 100 *pm* in the case of ring resonators [19]. Furthermore micro-rings are very temperature dependent and therefore temperature stabilization may be required [13].

2.4.2 Phase Modulators Based on Carrier accumulation

Carrier injection based demodulators in general, suffered from recombination of electron and holes in p-n/p-i-n junctions that inappropriately limits the switching speed narrowing the bandwidth of the device. More recently carrier accumulation has been investigated where thin insulating layer(s) of dielectric used in the waveguide structure in order to split the waveguide into capacitor(s), forms inversion layer on the sides of dielectric due to the applied external voltage.

The major milestone in Si modulators that has attracted huge media attentions was established by Liu, Intel Corp [42]. This work reports an experimental demonstration of a silicon phase modulator with a bandwidth that exceeds 1 *Ghz* for the first time. The device operates by free carrier plasma effect but based on carrier accumulation mechanism. The schematic of this device is illustrated in Figure 2.7.

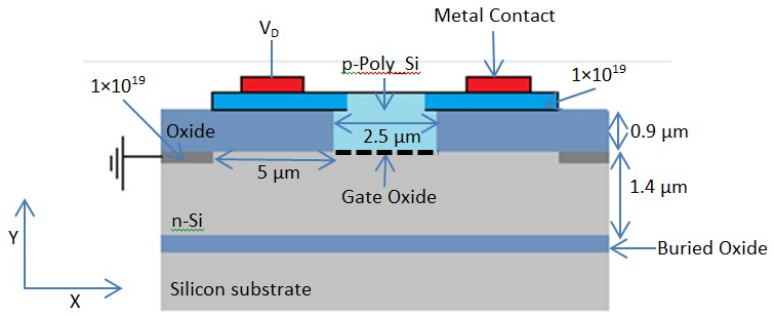

Figure 2.7: Schematic of the a MOS capacitor waveguide phase modulator using SOI technology [42]

The device structure consists a n-type crystalline at the bottom and p-type polysilicon rib layer on top that are separated by a thin insulating oxide layer. Once a positive voltage is applied to the p-type polysilicon, charge carriers accumulate at the oxide interface. This will then cause an effective change in the refractive index distribution that in turn shifts the phase of propagating wave through the device.

The bandwidth of the device faithfully exceeds 1 GHz. However, the effective length of the device is large (single 2.5 mm long phase modulator) and the reported on-chip loss is very high at 6.7 dB. Furthermore the device is highly polarization dependent because of the horizontal gate oxide. Phase modulation efficiency for TE polarization is larger than TM polarization by a factor of seven has been demonstrated [42]. Therefore all measurements reported were based on TE polarization.

Further to the Intel Corp model that predicts the device can be scaled to operate at 10 GHz, other carrier-accumulation based device was later demonstrated with such a high bandwidth. One successfully work has been demonstrated with an modulation depth of 3.8 dB [43]. Another accumulation-based modulator has been reported by Lightwire, which also achieved 10 GHz with an improved modulation depth of 9 dB [44]. The device is considerably efficient compared with Liaos model since it benefits from a significant reduction of active length from 3.45 mm for Liaos model to 480 μm [44].

Unlike carrier injection modulators, in carrier-accumulation mechanism, the modulation is not affected by the relatively long minority carrier lifetime in silicon, but is instead restricted by the device resistance and capacitance [13].

2.4.3 Phase modulators based on carrier Depletion

To avoid the speed limitation caused by the recombination of carrier, carrier depletion operation is proposed. In carrier depletion method one possibility is to form lightly doped p- and n-type regions adjoined in the waveguide to obtain a pn diode reverse bias

voltage causes the depletion area of to extend towards the edges and hence to deplete carriers from the waveguide area. Carrier depletion devices operate by allowing the propagating light to interact with the created depletion layer at a junction of reverse biased p-n diode and hence altering the initial refractive index.

A modulator with a bandwidth of 50 GHz was theoretically demonstrated by Gardes as the first carrier depletion based modulator in 2005 [45]. Intel has realized a similar three terminal device with two p-n diodes by Liu with a rate of 40 $Gbits^{-1}$ [46]. The schematic of this device shown in Figure 2.8.

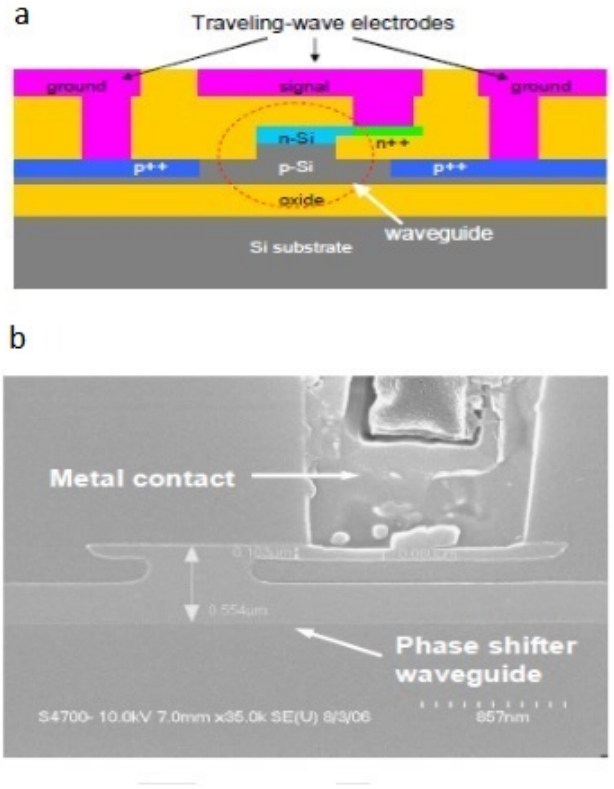

Figure 2.8: The schematic diagram of the theoretically studied of the optical phase modulator based on carrier dispersion effect [46].

The 50 $Gbits^{-1}$ silicon based modulator has been realized by Thomson and Gardes in 2012 as the fastest reported device [47]. This is shown in Figure 2.9. The device is incorporated in a Mach-Zehnder asymmetric interferometer structure. The reported extinction ratio is relatively low due to higher propagating losses. The authors expressed that small segment of the loss is due to the two multimode interference structures used to split and recombine the light from the two MZI wings. The remaining loss which has resulted a limited extinction ratio is the due to the scattering and impurities [47]. Optimisation of the fabrication process has been offered by the authors to result a lower loss device [47].

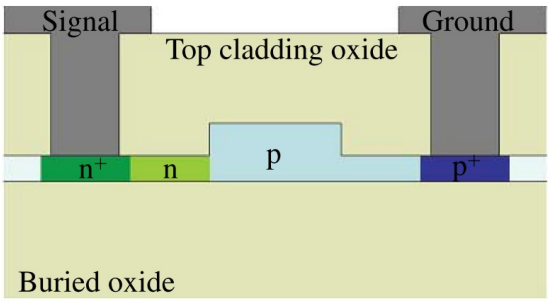

Figure 2.9: The schematic diagram of the fastest reported phase modulator implemented in a MZI structre based on free carrier depletion effect [47].

Carrier depletion approach has been studied in resonant-structure-based modulators. Successful works with the speed as high as 10 $Gbits^1$ have been demonstrated. The fastest ring resonator carrier-depletion-based structure reported so far operates at a bandwidth as high as 5 GHz [48].

It can be concluded, high modulation efficiency using an injection or an accumulation based modulator is achievable, but only at the expense of modulation speed as high as 10 GHz. However the depletion devices, in which higher modulation speeds can be reached, with longer effective length and higher power consumption. Therefore a challenge still remains to improve efficiency while maintaining other performance specifications such as propagating loss.

2.5 Loss Metrics Consideration

The overall propagating loss of a modulator is one of the fundamental metrics of the specifications mentioned in Table 2.1. In particular for nanoscale structures where the overall loss increases mainly as a result of coupling and greater modal interaction with sidewall roughness. As a result of additional loss the modulator performance is restricted due to the increased power consumption and smaller modulation depth.

Within the works reported in the last decade, there has been a promising improvement in the switching speed of Si modulators mainly as a result of considering the device operating in depletion mode. However, the loss applied to all referred works above, even the fastest reported one [47], has been considered as an issue which limits the performance of the modulator. Also the loss performance has not been very much investigated due to several reasons. The random nature of some of loss origins (such as scattering loss, which will be discussed later), the complication of the geometry and various absorptive characteristics that different materials may have are among the main reasons.

Possible solutions may be offered to boost the loss performance of the device if an accurate estimation of loss can be achieved. In some of the reported works, there have been efforts to minimize the propagating loss mainly by adjusting the parameters which directly affect the absorption behaviour of the devices. However, the influence of the other components of loss origins such as side wall roughness has yet to be individually studied and analysed. This will help to achieve an accurate comprehension of the loss effect in the modulation performance and consider the effective solutions. Such work has not been yet reported in any literature to the best of the my knowledge.

Also, it is clear that producing all building blocks of a modulator from Si as a monolithic structure can be challenging, and commercial success depends on the device performance and demands. So far, in order to improve modulation performance metrics, only few attempts have been made to investigate the effect of the presence of other material(s) that are potentially compatible with Si. Although the plasma dispersion materials still continue to be developed, introducing new materials to silicon based device and hybrid combinations can help improvement in two ways. The addition of new materials can fulfil the functions (such as electro-optic, acousto-optics effects) that silicon alone is currently unable to deliver because of its natural poor optical characteristics. At the same time, they can also help to efficiently in free carrier distribution in the waveguide. As one of the main focus of this research, this reasoning leads to investigate hybrid structures that benefit from addition of potential materials, specifically ZnO and SiO_2 together with silicon itself. It can participate in modulation based on acousto optic effect, that has been successfully demonstrated in optical fibers [18, 22]. More importantly, since ZnO is naturally n-doped material, it can contribute in free carrier transitions without adding any additional loss due to its transparency at relatively wide wavelength range (700 nm to more than 1600 nm). At the same time it can be practically useful to reduce the effect of scattering loss due to sidewall roughness if purposely applied. This effect can be investigated by introducing a novel roughness model discussed in Chapter 4 into the purposely designed hybrid modulator structure as one of the original motivations of this work. Hence reliable evolution of possible improvement that effective solutions may serve in a loss performance of the modulator can be studied. This discussion will be later followed in Chapter 5 in detail after reviewing the essential background of loss in general.

Chapter 3

Scattering Loss in Optical Waveguides

3.1 Introduction

It is mentioned in Chapter 2 that the loss performance is an important metric to consider in the design of an optical modulator operating based on plasma dispersion effect. In order to establish a well detailed insight of the role of the loss in the performance of such an optical modulator it is necessary to study the meaning of optical loss. In addition, the major sources of optical losses that contribute to the operation of the modulators have to be assigned accurately. In this Chapter a brief introduction to these potential sources will be given. The effect of sidewall surface roughness in the quality of propagation will be highlighted as one of the main sources of loss in state of the art optical modulator designs. Most successful experimental and analytical approaches to estimate scattering loss due to the waveguide roughness will also be reviewed.

3.2 Optical Loss Terminology

In different photonics devices various properties of propagating light may be altered either purposely or unintentionally. Once the light is generated and starting to propagate from a source into a waveguide it is manipulated and can be expanded, reflected, polarized, observed, refracted, scattered, condensed, filtered and so on as a result of a specified process. Sometimes these manipulations are purposely done in order to serve a particular application. In some other cases they result from the limitations imposed by physical characteristics of the light and optical properties of the interactive substances(i.e. dispersion). Some of these manipulations are appropriately manageable. In some other cases they cannot be fully brought under control and hence are considered

to be unavoidable constraints. Within last decades many efforts have been made to control the light characteristics when interacting with real medium [49]. In particular in the applications where the light intensity is weakened as it propagates in a real medium due to propagation characteristics and losses. This loss may be very small in specific dielectric materials for instance, but it cannot be zero at any time.

The interpretation of optical loss is different for each individual device, system or application. It entirely depends on the overall purpose of that device, system or application. For example in sensing application, the amount of measured optical loss along an optical fibre can be used to sense particular change in core/cladding interface properties [50]. Nevertheless in data transmission and relevant devices the optical loss in optical waveguide may not be desirable. While in some applications the scattering, absorption, refraction, and reflection are useful properties, in others are considered to be the main source of undesirable attenuations.

Optical loss in different type of optical waveguides from photonic nanowires and waveguides to optical fibres, have been always contemplated as an important parameter that can affect device/system performance. Advances in optical waveguide technology allow the integration of optical devices for high-performance operations such as communications [1]. Semiconductor and dielectric optical waveguides are dominantly considered as the basic building block of complex optical integrated circuits. In comparison with optical fibres that are used in telecommunication applications, optical waveguides for integrated optical devices usually exhibit higher losses due to their smaller sizes and errors occurred in the fabrication process. Therefore extensive studies have been in progress to explain successful techniques to quantify and consequently treat optical losses in optical waveguides [6, 51, 52]. For instance advances in material researches could dramatically reduce optical loss in waveguides [9, 10]. Vast number of researches are still going on to provide possible solutions in order to efficiently reduce optical loss as the waveguides scale down toward nano regime.

In order to serve a reliable estimation and treatment of optical loss it is necessary to have a realistic perception of the origins that causes optical loss in optical waveguides. This chapter briefly emphasise the mechanism of optical losses in optical waveguides in two categories; Extrinsic and Intrinsic losses.

3.3 Optical Loss Mechanisms in Optical Waveguides

3.3.1 Extrinsic Sources

Extrinsic losses correspond to geometrical specification of an optical waveguide rather than the material and imperfections that occur within the fabrication and processing in the construction of the waveguide.

3.3.1.1 Coupling Loss

This limitation usually arises either when the light is coupled into the waveguide from the light source or another photonic component, or associated with the coupling of the output of one waveguide to the input of another one or photonic component. There are several causes which may result coupling losses in waveguides.

1. Optical Source Radiance

   When the light is coupled from the light sources, if the launched rays are outside the right angle range, only excites dissipative radiation modes. Using objective lenses and suitable focusing arrangement can realise more efficient coupling between the light source and the waveguide [53]. The amount of power launched from a source into a waveguide depends on the optical properties of both the source and the waveguide. It depends on the radiance of the optical source which is the amount of optical power emitted in a specific direction per unit time by a unit area of emitting surface [53]. For most types of optical sources, only a fraction of the power emitted by the source is launched into the waveguide.

2. Fresnel Loss

   When the light needs to couple form either a light source or one waveguide to another, optical power may be reflected back into the source or the source waveguide. This reflected power is lost. This initial facet loss is called Fresnel losses that are due to reflections at the entrance aperture. In other word, Fresnel reflection is caused by a step change in the refractive index that occurs at the waveguide joint. This step change in refractive index is usually caused by a small air gap between the facets of waveguide. The ratio R in equation 3.1 approximates the portion of incident light (light of normal incidence)that is reflected back into the source waveguide [54].

   $$R = \left(\frac{n_1 - n_0}{n_1 + n_0}\right)^2 \tag{3.1}$$

   Here n_1 is the refractive index of the waveguide and n_0 is the refractive index of the medium between the two waveguides. Loss from Fresnel reflection may be significant. They become particularly more considerable if the waveguide and source are air coupled. A substantially better coupling can be therefore achieved by utilizing index matching materials [53]. Index matching materials help to reduce the step change in the refractive index at the waveguide joint, reducing Fresnel reflection.

3. Waveguide misalignments

   Coupling loss can be also significant in where two waveguides or waveguide based components having lateral, longitudinal or angular misaligned centres. Most of

these alignment errors are caused by mechanical imperfections introduced by waveguide or fiber jointing techniques. Longitudinal misalignments or waveguide separation occur when there is air gap between the facets of two joining waveguide. It becomes lateral misalignment if there are offsets between the two joining waveguide centres rather than a gap. Angular misalignment happens when their axis is tilted and hence the two waveguides are no longer parallel. Figure 3.1 illustrates these three typical form of misalignments.

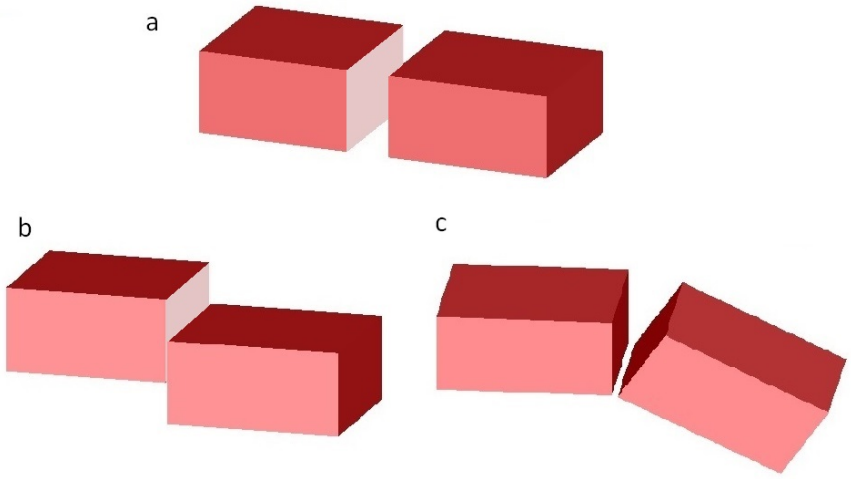

Figure 3.1: Different types of misalignment; a)Longitudinal b)lateral and c)angular

Coupling loss caused by lateral and angular misalignment typically is greater than the loss caused by longitudinal misalignments due to the relative ease in limiting the distance of waveguide separation. Coupling losses due to misalignment also depend on waveguide type and structures, the cross section size and the distribution of optical power among propagating modes. Waveguides with large Numerical Apertures (NAs reduce loss from angular misalignment and increase loss from longitudinal misalignments. Single mode waveguides are more sensitive to alignment errors than multimode waveguides because of their smaller cross sectional size. However, alignment errors in multimode waveguides joint may disturb the distribution of optical power in the propagating modes, increasing coupling loss [54].

4. Poor Facet Preparation

   The coupling loss increases as a result of poor facet finishes of the waveguides. The facets must be flat, smooth and perpendicular to the waveguide axis to ensure proper connection. If there are considerable imperfections at these facets the light will be scattered. In most cases the waveguide imperfections introduced on the facets during the sample cleaving process which will require to be polished in order to immunise the side effects. This is graphically shown in Figure 3.2.

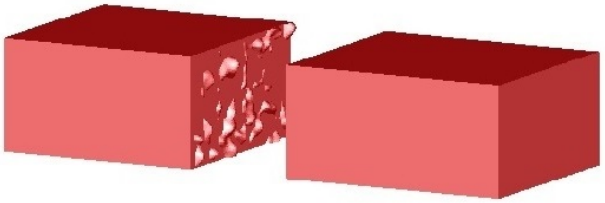

Figure 3.2: Poor facet preparation which causes that causes coupling loss

5. Waveguide Mismatches

   Coupling loss can arise independent of the waveguide jointing techniques, from the differences (or mismatch) in the inherent waveguide characteristics of the two connecting waveguides. There are three main types of mismatches in waveguide connections; waveguide cross sectional (geometry) mismatches, numerical aperture mismatch and refractive index profile difference between two waveguides. Cross sectional mismatches include the geometrical difference between the size or shape of the cross section and thickness of the core and cladding. This is graphically shown in 3.3. Numerical aperture (NA) mismatches result only if the launching waveguide has a higher NA than the receiving waveguide. Likewise coupling loss from refractive index profile difference is a case only if the launching waveguide has a larger profile parameter (real and imaginary refractive indices) compared to the receiver waveguide [54].

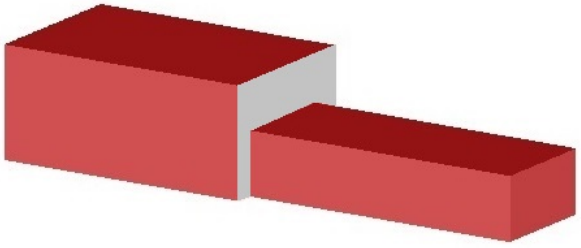

Figure 3.3: Mismatch between the cross sections of two connecting waveguides

In general, coupling loss is a parameter and can be broadly vary for different geometries and sizes. For example a reasonably designed and optimised waveguide with larger aperture may provide a more efficient coupling compared to a waveguide with tighter opening. Hence various input and output waveguide coupling mechanisms have been developed to enhance coupling efficiency [55].

3.3.1.2 Bending Loss

The light is distorted when the waveguide is deformed from the ideal straight configuration to include a bend in the propagation path. Any bend has an inner and outer radius. In order to maintain all power of the light when it propagates along a bend, the wavefront on the outer radius must travel faster than the part of the smaller inner radius. This is obviously not achievable and hence a portion of the power must be radiated away. The bending loss is different in different waveguide types such as rib, ridge or wire. It is also much greater in sharp bends in comparison with smoother bends.

3.3.2 Intrinsic Sources

Intrinsic losses are those associated with the materials that are applied in the structure of the waveguide. Unlike extrinsic losses, the total intrinsic loss is proportional to length L. Once the light is launched inside a waveguide, it is persistently attenuated as a result of the primary causes of the loss; absorption in the waveguide materials and scattering by surface roughness and material inhomogeneities.

3.3.2.1 Absorption Loss

In optical waveguides attenuation happens due to the interaction of photons with atoms, molecules and carrier charges of the medium itself or any additional impurities present in the medium. The energy that is lost in this way is usually converted to heat or acoustic waves.

Optical absorption is totally independent of the geometrical properties of the waveguide and quantified by attenuation coefficient. When the attenuation coefficient is large the light beam quickly decays as it progress through the waveguide. When it is small it means that the waveguide material is relatively transparent to propagating light. Absorption losses can widely differ from material to material and from wavelength to wavelength.

In many dielectric and semiconductor optical waveguide devices, band edge absorption and free carriers are considered as two main causes of absorption loss [56]. Absorption edge introduces an absorption discontinuity in the absorption spectrum of a substance. This discontinuity occurs at wavelengths where the energy of photon corresponds to an electric transaction between the two bands 3.4. In another world, an edge absorption happens when a photon is absorbed with energy greater than the band gap to excite electrons from the valence band. Note that when the quantum energy is smaller than the required work to eject an electron from one state the photon is not absorbed. Minimum absorption level can be obtained by selecting a wavelength that is longer than the absorption edge wavelength of the material. The optical absorption caused in this way

is an inherent loss mechanism that exists in all materials and tends to be an exponential function of length.

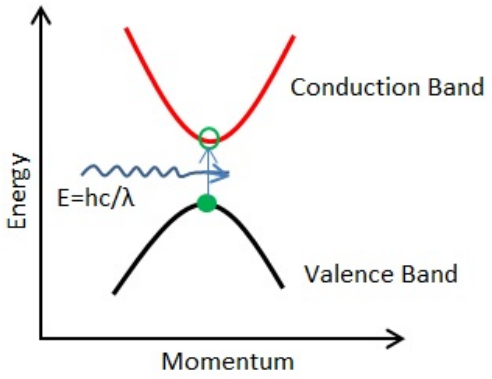

Figure 3.4: Example of interband absorption in a direct band gap semiconductor

The concentration of free carriers in semiconductor waveguides can significantly affect the absorption loss amount. Free carrier effects will be explained in more details in Chapter Four. This effect is generally explained according to Drude-Lorenz equation which describes how the variation in the concentration of electrons and hole can change absorption characteristic of a material [1].

3.3.2.2 Scattering Loss

- Scattering In General Mediums:

  Light scattering in general is a phenomenon in which the energy of light is scattered due to the interaction of light with matter. The primary sign of scattering is the deflection of the light ray from straight path. Scattering is inelastic if it involves transfer from one energy form to another. The two types inelastic scattering are Brillouin scattering and Raman scattering. The light is scattered with acoustic phonons (vibration of quanta of lattice vibrations in solids and elastic waves in liquids) in Brillouin scattering and with optical phonons (predominantly intra-molecular vibrations and rotation) in Raman scattering [57].

  On the other hand, elastic scattering occurs when there is only negligible energy transfer. Elastic scattering can be divided into three domains that are Geometric, Mie and Rayleigh scattering. This classification is based on a dimensionless size parameter $\alpha = \pi D_p/\lambda$ where πD_p is the circumference of a particle [57].

  Rayleigh scattering is a broad class of elastic scattering where the electromagnetic radiation is caused by particles much smaller than the wavelength of incident light ($\alpha << 1$). It occurs when light penetrates gaseous, liquid, or solid phases of matter. It results from electric polarizability of the particles that may be as small

as atoms or molecules. The electric field of the incident light wave acts on the charges within these particles causing them to move at the same frequency of the light. The particles therefore become small radiation dipoles whose radiations from the scattered light. The intensity of scattered light is inversely proportional to the fourth power of the wavelength of light in accordance to the following expression for single small particle [57]:

$$\alpha_s = \frac{2\pi^5}{3}\frac{d^6}{\lambda^4}(\frac{n^2-1}{n^2+2})^2 \tag{3.2}$$

where n is the substance refractive index, d is the of the particle and λ is the wavelength of the incident light. It can be seen from the expression that that the shorter wavelength in visible are scattered stronger than the longer wavelengths. Rayleigh scattering is normally considered as the main cause of signal loss in optical waveguides.

Mie scattering is a form of light scattering by mainly spherical particles of any diameter. In general, the amount of scattering is not strongly dependent on the wavelength, but is sensitive to the size of the particle. Mie scattering can correspond to Rayleigh scattering if the diameter of the particles are much smaller than the wavelength of the light. At this limit the shape of the particles no longer matters. Hence, Mie scattering are more meaningful when the particles are about the same size of the wavelength of the light ($\alpha \approx 1$). Mie scattering intensity for large particles is proportional to the square of the particle diameter [57].

Geometric scattering Restrictedly happens for particles much larger than the wavelength of light ($\alpha >> 1$).

- Scattering In Waveguides:

  When the light is scattered in a dielectric or semiconductor waveguide, the guided mode may be converted to the radiation mode, other guided modes, or the same guided mode propagating in the opposite direction. Scattering loss in the waveguides can occur in two levels:

  Subsurface scattering is due to the presence of molecular level inhomogeneities in the waveguide core medium. Inhomogeneities in waveguide material can be either structural or compositional. It is structural when the medium has disordered basic molecular structure and varying density in nature. It can be compositional when there are compositional fluctuations in the medium of the waveguide. These density fluctuations are unavoidably generated in the dielectric waveguide during fabrication and composition and by anisotropy of polarizability. In either case the effect can be observed as a variation in refractive index. As these fluctuations are generally small compared to the wavelength, the subsurface scattering is considered as a form a Rayleigh scattering. It can be a process dependent parameter and does not depend on the size and geometry of the waveguide but on the material

properties. It is inversely proportional to the fourth power of wavelength and hence short wavelength are scattered more than long wavelengths as stated before [53].

Surface scattering is due to surface roughness, sidewall roughness and waveguide geometry irregularities at the core/cladding interface of the waveguide. Light propagates inside a waveguide in accordance to total internal reflection when the wavefront interact with the core/cladding interface. The propagating light will be distorted by scattering if the interface is not ideal and has roughness or irregularities on the surface. Hence a portion of the light power is radiated away leading additional losses in the waveguide. The surface scattering loss is proportional to the length and hence as the light propagate in the waveguide it decays with higher amount. The extent of distortion depends on a number of factors including roughness properties, wavelength, mode polarization state and geometry. For example, the scattering is more considerable when the optical path is long and also for small and high index contrast waveguides [52].

3.4 Surface Roughness As a Major Source For Scattering Loss

As an intrinsic source of loss the surface scattering can be quite distinct from other loss origins in optical waveguides because of several reasons:

Unlike subsurface scattering, absorption and extrinsic losses, surface scattering originates from roughness that is a technological process dependent factor. Roughness may be reduced or minimized by improving the fabrication process and treatment, but cannot be fully eliminated. Regardless of the materials and the geometry a certain amount of roughness, even small, always exists anywhere in a waveguide and any optical waveguide-based device.

In addition, as the size shrinks down toward nanoscale, some physical waveguide parameters become significant and uncontrollable, hence limits the range of operation. In particular the effect of roughness becomes even more important in smaller waveguides where there is less capacity to control physical parameters in processing and fabrication levels.

The random nature of roughness is another feature which makes surface scattering a unique phenomenon amongst other origins of loss. The generation of roughness is very much dependent on the fabrication process. Therefore, certain type of approaches may be needed in order to precisely estimate the scattering loss. There have been major methods developed to either experimentally measure or analytically estimate scattering loss. Here I briefly review the most common techniques that are widely referred in literature.

3.5 Optical Waveguide Scattering Loss Measurement Techniques

Over the last two decades many methods have been developed for loss measurements in optical waveguide. There are also various techniques to split the total measured loss into its two main components, scattering and absorption, particularly for materials with high absorption characteristics or large number of impurities [58]. Here, among all loss measurement methods the most common techniques will be referred, having in mind that the scattering loss defines the ultimate lower limit on propagation loss in dielectric and low absorptive waveguides.

Traditionally optical propagation loss is extracted through measuring the transmitted power as a function of waveguide length [59]. One of the initial techniques using this approach was cutback method originally implemented for fiber optic measurements [53]. In this method the transmitted power through a longer fiber (P_1 of length L_1 is compared to the transmitted power through a shorted piece of the same fiber (P_2 of the length of L_2. The total attenuation per unit of length is then calculated as follows [53]:

$$Attenuation = \frac{10log\frac{P_1}{P_2}}{L_1 - L_2} \tag{3.3}$$

This method is later applied to waveguides on substrate [60, 61]. It involves comparison of power transmitted through waveguides of different lengths assuming identical sidewall and surface roughness. In addition, in order to obtain accurate results fixed coupling conditions must be applied to all waveguides over multiple measurements. Hence the method may suffer from lack of accuracy as the reproducibility of identical coupling conditions may become difficult over multiple measurements. In order to resolve this problem other methods are also developed that retain a high coupling efficiency and therefore are extensively used for propagating loss measurements. The most common are sliding prism and three prism techniques. In these methods prism couplers are used to couple substantial fraction of the light into the waveguide. Hence there would be no such a requirement of cutting the waveguide and polishing the faces.

Weber et al. used two prisms, one to couple light into the waveguide and the other one was a sliding output prism [62]. The input prism is pushed against the waveguide with a constant pressure. In practice, it is difficult to measure the input coupling efficiency. More accurate results can be obtained by coupling the light out of the waveguide into the detector at different position of the waveguide. Here is where s sliding prism can be used. The precisely adjusted output sliding prism is connected to the waveguide using an index matching liquid. Using such a liquid helps to maintain coupling conditions while there would be no need to press output prism against the waveguide. It is however important that a constant fraction of the intensity of the guided light is detected by the

detector. Weber et al. claimed by having sufficiently long contact length of the prism along the waveguide can provide a coupling efficiency of 100% [62]. The output light needs to be 100 percent collected by the detector and has to be not deflected or scattered otherwise [62].

The three prisms method is established by Y. H. Won et al. as an alternative which is independent of input and output coupling efficiencies, and also applicable to all planner waveguides [63]. Figure 3.5 presents the schematic experimental setup. Prism 1 and 3 are fixed while prism 2 can slide along the waveguide. The light is launched and coupled to the waveguide through Prism 1.

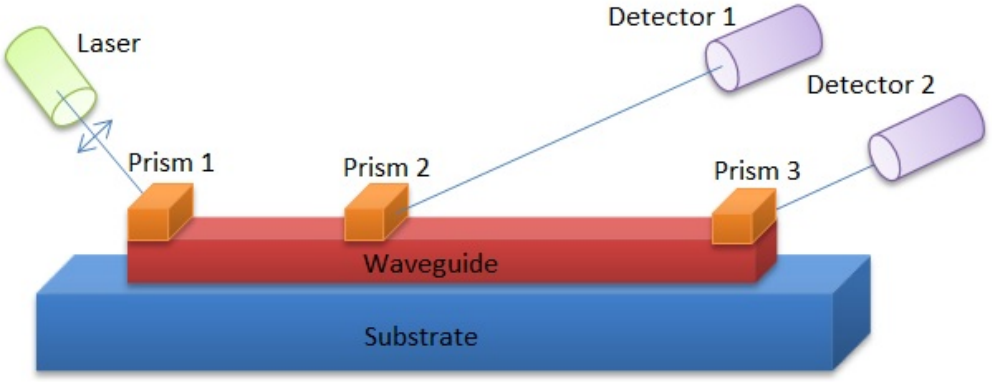

Figure 3.5: Three prisms measurement configuration.

If the coupling efficiency of prism 2 and 3 are γ_2 and γ_3 respectively, the output power of prism 2 and 3 can be calculated as:

$$P_2 = \gamma_2 I(L_2) \tag{3.4}$$

$$P_3 = \gamma_3 [I(L_2) - P_2] exp[-\alpha(L_3 - L_2)] \tag{3.5}$$

where α is the attenuation of the waveguide, $I(L_i)$ is the intensity of the light in the waveguide in the position of L_i that the place where the prism is clamped. Two distinct experiments are performed, In the first case $\gamma_2 = 0$. In the second case $\gamma_2 \neq 0$ and prism 2 is moved along the waveguide changing the collected power by ΔP_3, Using Equation 3.5 twice leads to have:

$$I_L = P_2 P_3^0 / \Delta P_3 \tag{3.6}$$

where P_3^0 is the power collected at prism 3. As can be seen this result is independent of the coupling efficiencies γ_2 and γ_3 . In both sliding prisms and three prisms methods

measurement of scattered light is performed from the top of the material along its length. The accuracy can be therefore be restricted the quality of the top surface which may not be appropriately smooth and light can be scattered at the interface consequently.

Another accurate method originally proposed by Regener and Sohler is based on Fabry-Perot(FP) resonance which does not need the information of input coupling loss [64]. This method is one of the most extensively used techniques to measure waveguide propagation loss, in particular within the last decade [61]. A cavity can be configured when the facets of the waveguides are cleaved or polished as its input and output ends. A tunable laser is coupled into the waveguide through its end facet. The power of the output light is detected by a precisely tapered single mode fibre, for a constant power of input light. The transmission has periodic resonance spectrum which is the measured transmitted power when wavelength is varying within the range of interest. The propagation loss can be then calculated using the following Equation [65]:

$$Attenuation = -\frac{1}{L}\left(\frac{1}{R}.\frac{\sqrt{\frac{I_{max}}{I_{min}}}-1}{\sqrt{\frac{I_{max}}{I_{min}}}+1}\right) \tag{3.7}$$

Where here L is the waveguide length, R is the facet reflectivity and I_{max} and I_{min} is the peak and bottom intensity of each resonance. Although the method can provide a high accuracy, it requires the estimation of the facet reflectivity. Assuming ideally polished facet one can estimate the reflectivity from effective modal simulations or index measurements using the following Equation [64]:

$$Reflectivity = \left(\frac{n_{eff}-1}{n_{eff}+1}\right)^2 \tag{3.8}$$

Another common technique which can be particularly applied to measure propagation loss in integrated optical waveguides and devices is commonly known as out of plane measurement, utilising a video camera assisted microcomputer system as shown in Figure 3.6 [66].

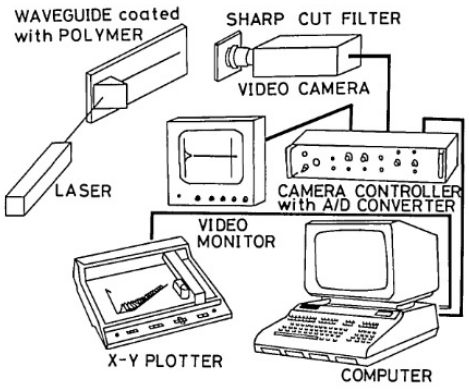

Figure 3.6: Video camera analysis system for measuring the propagation in optical waveguides [66]

Once the light is generated a prism film coupler is used to excite only one desired guided mode selectively. A video camera is set on top of the waveguide to observe the light streak scattered out of the optical waveguide. The output video signal generated by the camera is then sent to the controller and fed into both a video camera monitor and a microcomputer through a video analogue to digital converter. Using the micro-controller the steak light will be transversely scanned and hence the peak intensity variations along the streak can be obtained. Repeating the same procedure along the streak 2D information for intensity distribution can be worked out. The slope of the least mean squares fit to a decreasing exponential of the longitudinal variation of the mode power in the waveguide yields the power loss coefficient. The primary configuration of this system was insensitive to small scattering. Accordingly an improvement was achieved by employing the fluorescence of a dye to observe the optical field propagation in the waveguide. The unguided fluorescence light enables one to measure a map of the guided light intensity [66]. Also by separating the absorbed fluorescent light from the excited light using sharply cut filter , the signal to noise ration can be increased.

There are other limitedly applicable methods that have been reported in the literature where are rarely used mainly due to the instability and complexity issues. Hence, the details are beyond our interest. These are pyro-electric absorption loss measurement [67], and direct temperature measurement [68] methods.

3.6 Analytical Approaches to Calculate Scattering Loss in Optical Waveguides

As the absorption losses in waveguide for example in silicon waveguide, have been significantly reduced for 1.3 μm and 1.55 μm wavelengths, the propagation losses are mostly the result of radiation loss due to scattering by surface irregularities of the waveguide surface walls.

Various experimental techniques have been developed to measure scattering loss and optical waveguide roughness effect in light propagation. High instrumentations expenditure of measurement devices and the long time for any loss measurements, lead researchers to consider analytical approaches. Hence slightly earlier there have been attempts to theoretically evaluate scattering loss in waveguide. In this section, the most successful models used to estimate the waveguides scattering loss in the waveguide due to sidewall roughness are explained. The positive and negative features of each of these methods will be also discussed.

3.6.1 Marcuse Method

Marcuse formulated the scattering losses due to the irregularities on the waveguide walls from calculating the energy perturbation introduced by such imperfections [51]. The applicability of its model however, is limited for an infinite dielectric slab waveguide. Marcuse started by considering an ideal dielectric waveguide case where the solutions for Maxwell′s equations with the boundary condition of an ideal waveguide are a finite number of discrete guided modes. The energy carried by these modes is concentrated inside and near the waveguide. There is however additional continuum of radiation modes that extend to infinite distances from the waveguide. For an ideal waveguide, no energy conversion occurs between the propagating guided modes to other guided modes or to radiation modes. However, any imperfection in the waveguide boundaries provides a path for coupling energy between different modes, resulting in an energy loss. In this way the energy is shifted from the guided modes to radiation modes [51].

Such imperfections can be described as a variation in the index of refraction and a deviation from the straight waveguide core/cladding interface as follows. The waveguide wall irregularities are expressed by refractive index variation δn that has a spatial dependence in the propagation direction and in the direction perpendicular to the propagation direction. This is shown in Figure 3.7 where waveguide wall has δn that has x and z dependence.

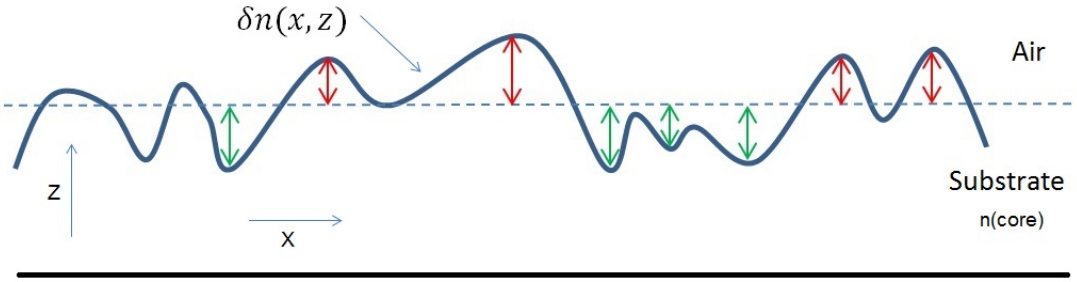

Figure 3.7: A two dimensional random profile of the sample surface roughness. The red and green arrows represents positive and negative distances of different profile points from the mean level (dash line) respectively.

In an ideal waveguide case, the propagation of light is described by Helmholtz equation. Since there is sidewall imperfection on imperfect waveguide walls, the homogeneous Helmholtz equation is modified by replacing n^2 with $n^2 + \delta n(x,z)^2$ that results in a modified homogeneous Helmholtz equation.

$$\left[\nabla^2 + k_0^2(n^2 + \delta n(x,z)^2)\right] E = 0 \tag{3.9}$$

$\delta n(x,z)^2$ is either 0 or $-n^2$, depending on the spatial position. $n^2 + \delta n(x,z)^2$ is n^2 inside the waveguide wall and 0 outside the waveguide wall in Figure 3.7. The field distribution of TE modes that satisfy the modified Homogeneous Helmholtz equation can be expressed as a summation of all guided and radiation modes as follows:

$$E_y = \sum_m C_m(z)\epsilon_m + \int_0^\infty g(\rho,z)\epsilon(\rho)d\rho \tag{3.10}$$

The expression consists of two parts. The summation part which extends over all guided modes of the discrete spectrum and the integration part that extends to infinity to cover the radiation modes of the continuum spectrum. ϵ_m and $\epsilon(\rho)$ describe the guided discrete and radiation continuum modes respectively. $C_m(z)$ and $g(\rho,z)$ are the normalised coefficients for different modes. Substituting Equation 3.10 into Equation 3.9 and rearranging to a set of two inhomogeneous differential equations for $C_m(z)$ and $g(\rho,z)$, the solutions for these differential equations can be derived as [51]:

$$C_m = C_m^+ + C_m^- \tag{3.11}$$

and

$$g(\rho,z) = g(\rho,z)^+ + g(\rho,z)^- \tag{3.12}$$

The positive sign indicates the coefficient, which waves travelling in positive z direction, while negative indicates the waves travelling in negative z direction.

Considering that the imperfect waveguide is connected to perfect waveguides at 0 and at z=L, the power loss ratio for the fundamental MODE (M=0), $\Delta P/P$, can be given as follows:

$$\frac{\Delta P}{P} = \sum_{m=1}^{\infty}\left[\left|C_m^+(L)\right|^2 + \left|C_m^-(0)\right|^2\right] + \int_0^\infty \left[\left|g^+(\rho,L)\right|^2 + \left|g^-(\rho,0)\right|^2\right] d\rho \tag{3.13}$$

Equation 3.13 calculates the total power lost by mode conversion from the fundamental mode to all modes travelling in the positive z direction at z=L and to all modes travelling in negative z direction at Z=0. However, C_m^+, C_m^-,g^+ and g^- can be only solved approximately due the complicated algebra involved. First order perturbation theory is employed by Marcuse to obtain approximate solutions in this approach. The attenuation coefficient, α, can be easily obtained from $\Delta P/P$ [51].

Due to complexity of this theory, only numerical solutions are feasible to calculate the scattering power loss. In addition, this theory assumes that the cladding is air. A more simplified model is desired to work out scattering in dielectric waveguide with non-air cladding. Furthermore, this treatment is only applicable to single mode waveguide. The work was then extended to cylindrical waveguides and experimentally evaluated by the same author [69]. The more accurate results for cylindrical waveguide have been later reported by Snyder and Love where the corrected version of Marcuse expression is published [70].

The developed expression for radiation loss by Marcuse has a complete description of both near and far fields, while the discussion of power loss requires an expression for only the far field [6]. Hence the expression is quite complex, not easily physically understood and includes many irrelevant parameters. This complexity makes it difficult to evaluate the possible effects of the variation of more general waveguide parameters such as refractive indices, optical wavelength and waveguide thickness.

Payne and Lacey developed a simplified expression by excluding the unnecessary near field term [6]. This expression is hence less complex and includes waveguide physical parameters. It is also commonly accepted as an analytical description for 2D planar waveguides and widely referred in the literature for comparison with experimental results [52,71]. There were also attempts to extend Payne and Lacey's model into 3D structures [52, 72, 73].

However, for all analytical approaches, it is necessary to access the roughness statistical data of the waveguides in spite of the faster and cost effective features of these methods. Furthermore, most of these methods were originally developed for 2D/3D optical rectangular waveguides only neglecting the possible surface roughness on top of the waveguide. In fact they only consider sidewall roughness of different type of rectangular waveguides. This may be referred as an accurate approximation considering current advanced fabrication processes, but still may cause certain limitations in specific applications where the top surface is not perfectly smooth.

Here below Payne and Lacey method is briefly reviewed as one of the most common and successful analytical methods which are widely applied to estimate scattering loss due to the roughness.

3.6.2 Payne and Lacey Analytical Method for 2D

3.6.2.1 General Overview of The Method

In order to overcome the complexity of the Marcuse′s model, Payne and Lacey formulated simpler expression from the far field pattern only. The waveguide sidewall roughness in their model is considered to create an equivalent current that acts as source of radiation from the waveguide. They modified the Helmholtz equation differently from Marcuse′s approach. By using an approximation, the resulting electric field from the current generated by surface roughness in the imperfect waveguide can be expressed as a function of the electric field for an ideally smooth waveguide. Solving the Helmholtz equations in these conditions leads to the scattering loss coefficient expression.

3.6.2.2 Mathematical Description of The Method

The dielectric waveguide in Payne and Lacey analysis can be either air or non-air cladding waveguide. Hence the modified homogeneous Helmholtz expression that Marcuse derived in equation 3.9, $n^2 + \delta n(x,z)^2$ are replaced by

$$n^2 = n_1^2 + (n_0^2 + n_1^2)U[d + f(z) - |x|] \tag{3.14}$$

Where n_1 and n_0 indicates cladding and core respectively, l is the half width of the waveguide, and $f(Z)$ is the roughness function as shown in Figure 3.8. Here $U[d + f(z) - |x|]$ is unit step function where

$$U[d + f(z) - |x|] = 0, \qquad for \quad d + f(z) < |x| \tag{3.15}$$

and

$$U[d + f(z) - |x|] = 1, \qquad for \quad d + f(z) > |x| \tag{3.16}$$

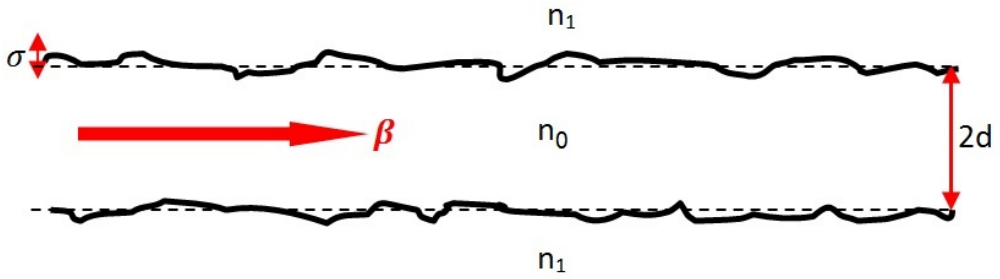

Figure 3.8: Top view of a 2D waveguide with random roughness on sidewalls. β is the light propagation constant, σ is RMS roughness, d is half of the waveguide width n_1 is refractive index of the waveguide and n_2 is the cladding refractive index.

The equation 3.14 describes the waveguide sidewall roughness where there is an index step from n_0 to n_1 along the rough wall at $x = f(z)$,

$$n^2 = n_1^2, \qquad for \quad d + f(z) < |x| \tag{3.17}$$

and

$$n^2 = n_0^2, \qquad for \quad d + f(z) > |x| \tag{3.18}$$

The modified Helmholtz equation can be rewritten as:

$$(\nabla^2 + k_0^2 n_1^2)E = k_0^2(n_1^2 - n_0^2)U[d + f(z) - |x|]E \tag{3.19}$$

Since there is no analytical solution for equation 3.16 further simplification is required by approximation. In order to achieve this approximation the electric field of an ideal waveguide, E_{ideal}, with no roughness can be obtained from solving the ideal homogeneous Helmholtz equation 3.17.

$$(\nabla^2 + \omega^2 \mu\epsilon)\overline{E} = 0 \tag{3.20}$$

Since $U[d + f(z) - |x|]$ is 0 outside the waveguide core, the right hand side of equation 3.19 only depends on the field inside the waveguide. The field inside the waveguide can be approximated to be similar to that of an ideal waveguide, E_ideal. This approximation method can be found in [74]. With this approximation Equation 3.19 can be rewritten as:

$$(\nabla^2 + k_0^2 n_1^2)E = k_0^2(n_1^2 - n_0^2)U[d + f(z) - |x|]E_{ideal} \tag{3.21}$$

Payne and Lacey explained that Equation 3.21, can be better viewed by considering the antenna theory [6]. The terms on the right hand side, which is the product of a random function and the known electric field, describes the current created by the random roughness on the waveguide sidewall. This current acts a source of radiation from the waveguide. E on the left hand side represents the resulting electric field in space due this current. Since only radiation field in space is the only concern, the far field solution is considered in solving equation 3.21. The far field solution is superposition of monochromatic plane waves. Assuming that the roughness is much smaller that the waveguide dimension ($f(z) < d$) the radiation field can be obtained following lengthy algebra. It should be noted that E is the sum of radiated plane waves and hence the

radiated power, P_{rad}, is related to the ensemble average of the radiated field intensity, $< |E|^2 >$. The resulting attenuation coefficient will be then obtained as:

$$\alpha = \frac{P_{rad}/(2L)}{P_g} \tag{3.22}$$

where $2L$ is the waveguide total length, $P_{rad}/(2L)$ is the radiated power per unit length of the waveguide, and P_g is the total guided power. Finally, the scattering loss coefficient due to the roughness effect, $\alpha_{radiation}$, can be derived as [6]:

$$\alpha_{radiation} = \phi^2(d)(n_0^2 - n_1^2)^2 \frac{k_0^3}{4\pi n_1} \times \int_0^{\pi} R(\beta - n_1 \tilde{k}_0 cos\theta) d\theta \tag{3.23}$$

Where n_0 and n_1 are refractive indices of core and cladding respectively, β is propagation constant, k_0 is the wavenumber in free space, σ is root mean square deviation from the flat surface, and $\phi(d)$ is the modal field evaluated at the waveguide surface. This can be calculated either using simulation tools or theoretically as will be shortly noted. θ is the incident angle which relates to the propagation mode number.

$R(\tilde{\Omega})$ is the spectral density function of the sidewall roughness which is used to describe the random nature of the roughness function $f(z)$. The spectral density function is related to the autocorrelation function, $R(z)$, of the sidewall roughness through the Fourier transform as will explained in the next section.

3.6.2.3 Spectral Density Function and Autocorrelation Function

Sidewall roughness has a random nature whose details are very much dependent on the fabrication process. In order for the roughness to be analytically modelled it should be considered as random process. As a random process it can be characterised by the spectral density function $R(\tilde{\Omega})$. The spectral density is a function of frequency which described how a random value is related to the frequency. The spectral density function is related to the autocorrelation function(ACF). The correlation function explains how two random variables are related and ACF describes how a random variable is related to itself. The following example will reveals the relation between a random function (sidewall roughness for an instance) to the autocorrelation function.

Consider that the following 2D profile shown in Figure 3.9 A is extracted from the edge the waveguide ($f(z)$).

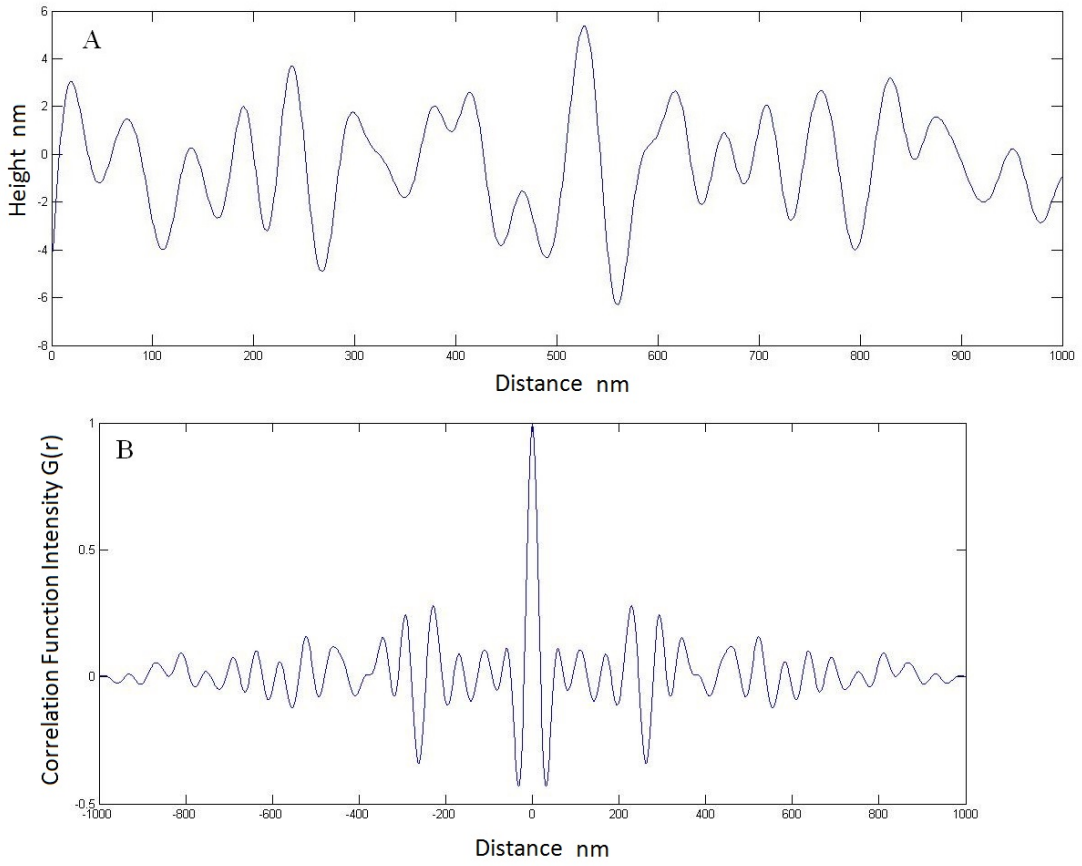

Figure 3.9: *A*) Random 2D profile generated using Matlab *B*) The corresponding ACF of the 2D random profile in A

The corresponding normalised autocorrelation function is calculated using Matlab as shown in Figure 3.9B. The ACF here is a measure of similarity of a random process such as the above roughness profile to itself when a space-lag applied to it. There the ACF is defined as a function of space or displacement. When the profile has zero space lag with respect to itself the ACF is maximum and equals to 1. As the lag starts to increase a larger mismatch occurs and hence the correlation function begins to fall from its peek. The largest mismatch occurs when the correlation function approaches zero level. From that point the ACF will be constantly oscillating in small range around the zero level, no matter how big the lag is. This point defines the correlation length. This explanation is however true for non periodic signals and random profiles. Hence the correlation length is in fact a measure of the minimum length required to achieve the highest mismatch between the profile and itself when a certain lag is applied to it. Figure 3.10 shows the ACF and the filtered curve. The correlation length is approximately 60 nm.

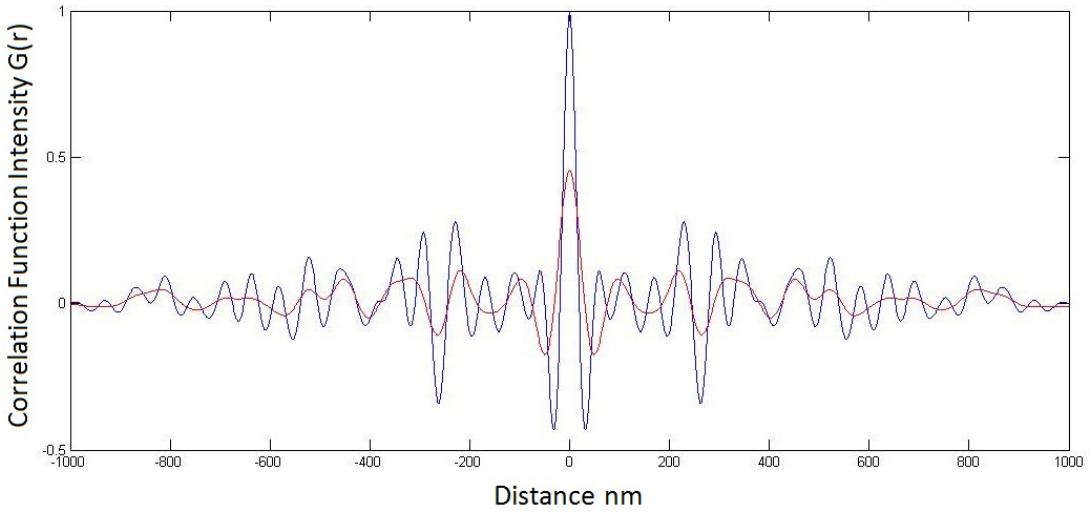

Figure 3.10: The ACF of the 2D profile and its filtered curve (red)

The correlation length is necessary to explain the variation in the vertical as a function of horizontal position. It illustrates the horizontal length which is needed so it fully covers one 'average' bump. The correlation length value however cannot fully explain the roughness statistics on its own. In fact, the ACF of random roughness is characterised by correlation length L_c, and mean square deviation $\sigma^2 = R(0)$ from a flat surface without irregularity.

RMS stands for the Root Mean Square deviation of a surface. Considering the 2D profile of a surface called S in Figure 3.9B (also in Figure 3.7), for any point on the profile, there is a distance (deviation) from the mean level. All distances above the mean level are positive and distances below the mean are negative. The mean is the value where superposition of all distances equals zero. The RMS is in fact the Root Mean Squared distances (positive values) and can be found for a 3D surface from the following expression;

$$\sigma = \sqrt{\frac{1}{MN}\sum_{j=1}^{M}\sum_{i=1}^{N}\eta^2(x_i, y_j)} \tag{3.24}$$

Where M is a number of points per profile (scan line), N indicate indicates the number of profiles and η is the amplitude at (x_i, y_j) point.

The spectral density function $R(\Omega)$ of the surface roughness is related to ACF through the Fourier transform by Wiener's hypothesis [6].

$$\tilde{R(\Omega)} = \int_{\infty}^{\infty} R(z)exp(i\Omega z)dz \tag{3.25}$$

Hence the autocorrelation function of such a roughness profile is an important identity that can be used to interpret the roughness behaviour. As the ACF can be obtained by multiplying the signal by itself when it lags by a variable amount, it can reveals how the points on this profile are related together.

3.6.2.4 Exponential and Gaussian Autocorrelation Functions

The most commonly used ACF for roughness are exponential and Gaussian ACF which are respectively represented as follows [6].

$$R_{Exponential}(u) = \sigma^2 exp(-\frac{|u|}{L_c}) \tag{3.26}$$

$$R_{Gaussian}(u) = \sigma^2 exp(-\frac{u^2}{{L_c}^2}) \tag{3.27}$$

Here, the lateral surface correlation length of interest which is defined as the range where two points are correlated together. Hence any point beyond this length is assumed to be uncorrelated and acts independent to the original point.

Atomic Force Microscopy (AFM) measurements of silicon waveguide surface roughness have demonstrated that an Gaussian ACF is a realistic model for the statistics of surface roughness of a variety of waveguides [75]. Hence in this work I only used Gaussian ACF.

Substituting the exponential and Gaussian ACFs in Equation 3.26 and 3.27 into Equation 3.23 result the relevant radiation loss coefficients to be respectively worked out as:

$$\alpha_{radiation,exp} = \phi^2(d)(n_2^2 - n_1^2)^2 \frac{k_0^3}{4\pi n_1} \times \int_0^{\pi} \frac{2\sigma^2 L_c^2 \theta}{1 + L_c^2(\beta - n_2 k_0 cos\theta)^2} d\theta \tag{3.28}$$

$$\alpha_{radiation,Gaus} = \phi^2(d)(n_2^2 - n_1^2)^2 \frac{k_0^3}{4\pi n_1} \times \int_0^{\pi} \sigma^2 L_c \sqrt{\pi} \times exp\left[\frac{-L_c^2(\beta - n_2 k_0 cos\theta)^2}{4}\right] d\theta \tag{3.29}$$

These expressions are then simplified and rearranged to include normalized ν the frequency, U and W, the normalized wavenumber in core and cladding of a two dimensional planer waveguide [6].

$$\alpha_{radiation} = \frac{\sigma^2}{\sqrt{2} k_0 d^4 n_1} g.f \tag{3.30}$$

The function g is purely determined by the waveguide geometry (u and w) that is:

$$g(V) = - = \frac{U^2V^2}{1+W} \tag{3.31}$$

Where U, V and W, the normalized waveguide parameters are:

$$U = d\sqrt{n_1^2k_0^2 - \beta^2} \tag{3.32}$$

$$W = d\sqrt{\beta^2 - n_1^2k_0^2} \tag{3.33}$$

$$V = k_0 d\sqrt{n_1^2 - n_2^2} \tag{3.34}$$

The function f is obtained according to the ACF (exponential or Gaussian) where the relevant L_c also contributes in addition to U and W.

In addition, the modal field in equation 3.23 can be theoretically calculated from the following equation.

$$\phi^2(d) = \frac{2k_0 d\sqrt{n_0^2 - n_1^2}(1-b)}{d\left[k_0 d\sqrt{n_0^2 - n_1^2} + 1/\sqrt{b}\right]} \tag{3.35}$$

3.6.3 Effective Index Method For 3D Structure

The above analytical approach is derived for a 2D dielectric slab waveguide where the light is only confined horizontally. In order to consider 3D structures the expression in Equation 3.30 is modified using effective index method. The detail description of this method can be found in the book by Gupta et al [76].

The propagating mode travels in a waveguide with propagating constant β, which can be related to effective index n_{eff} by:

$$\beta = \frac{2\pi}{\lambda_0} n_{eff} \tag{3.36}$$

Since $\frac{2\pi}{\lambda_0} = K_0$, the above Equation can be written as

$$n_{eff} = \frac{\beta}{K_0} \tag{3.37}$$

From equation 3.37 one can obtain n_{eff} of any waveguide structure wen β can be calculated.

In Effective Index Method the following argument is used; When n_{eff} can be calculated the number of dimension of the a particular waveguide can be reduced. For example consider a slab waveguide which the optical mode is confined in two dimensions restricted in heights and width. The confinement in either height or width can be eliminated by constructing a slab waveguide in the perpendicular dimension. the core index has to be however replaced by n_{eff}, that is the confinement in the removed dimension [76].

The method is graphically shows in Figure 3.11. The value of β and n_{eff} can be accurately calculated using numerical methods such as Finite Element or Finite different Time Domain Methods.

Besides the Effective index methods other analytical methods have been reported in literature, such as Volume Control and Poulton semi-analytical method [52, 77].

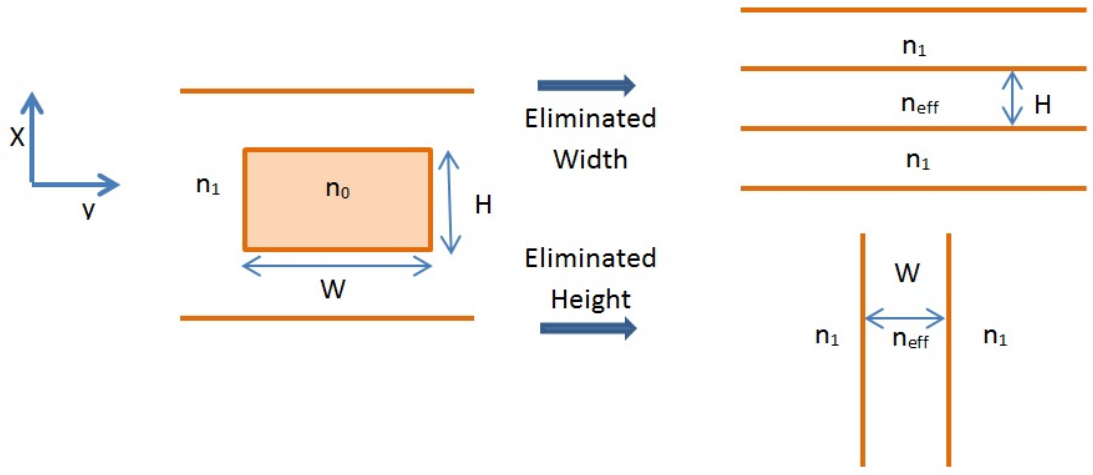

Figure 3.11: Graphical representation of Effective Index Method

3.7 Roughness Parameter Extraction

As mentioned earlier analytical estimation of roughness requires roughness statistical data of the top and sidewall of the waveguide including correlation length and RMS roughness. These data can be quantitatively extracted from given samples using two fundamental approaches. Regardless of the applied method the main remaining challenge is to access the sidewall surfaces, as the top surface data can be easily scanned. The most common instruments to characterise roughness are as follows.

3.7.1 Scanning Electron Microscopy (SEM)

This method was originally applied to etched sidewall profile. The instrument has four electron detectors which are positioned facing each other. The surface roughness at electron beam incident point is obtained by measuring the four output signals of the four detectors. The complete roughness profile can be found by scanning the sample in xy plane. Since SEM measurement is performed in a noncontact manner the sample will retained undamaged in this measurement process. The scanning resolution is however limited to 1 nm [78].

3.7.2 Atomic Force Microscopy (AFM)

Atomic force microscopy or AFM utilizes a cantilever probe tip to detect weak forces on a sample. The instrument is designed to scan the sample horizontally in xy direction and detect extremely small repulsion forces detected by the probe. These forces are due to the vertical displacement of the pointed end of the cantilevered probe following the shape of the scanned area surface. The pointed end of the cantilevered either make contact or function in noncontact mode with the scanned area. The surface data is collected by lasers, piezo electric sensors or photoelectric sensors. The piezo electric sensor realise the surface image by sending the generated voltage due to movement of the cantilever to a transducer. The photoelectric sensor however measures movement based on the variation in the incident angel caused by movement of the cantilever. Laser also works based on the same principle that photoelectric sensors operate. This configuration is shown in Figure 3.12.

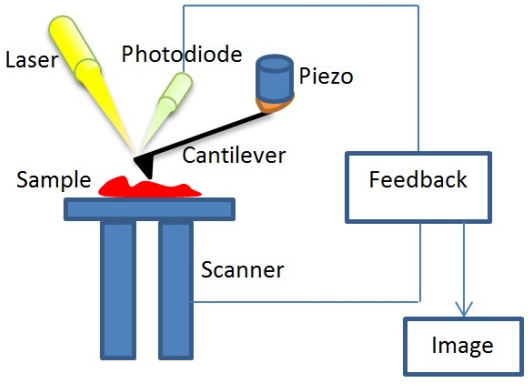

Figure 3.12: Sidewall roughness scanning configuration using AFM

The AFM has better sensitivity in characterising planer surface morphology than SEM and TEM, but is difficult to use for measuring the vertical surfaces such as sidewalls in 3D structure. This limitation has been recognised over the past 25 years, in particular for

submicron geometries. There have been three main solutions developed for full imaging sidewall and submicron measurements.

1. Use special high aspect ratio boot shaped AFM tip which is mounted in a AFM system with 2D servo control system which the details can be found in [79]. Precise results have been obtained using this approach, but the scanning area was limited by the length of the special AFM tip as well as the resolution of the scanned image.

2. Mount a sample at a title angle. The details are reported [80]. This method needs a complex calibration in order to obtain correct results, especially for deep etched optical waveguides.

3. Cleaving a sample in an appropriate way where a good segment of a waveguide is placed very close to the cleaved wafer edge [81]. This method may need a specially designed line pattern of waveguide in a sample. In addition it requires applying serious damage into the sample that certainly restricts its use in many applications.

A complete realisation of sidewall roughness is an advantage and can provide various data of the property of sidewall irregularities. In many case however all of these data are not required. For this work it may not be required to fully recover the complete image of a sidewall roughness. As I will discuss later, I only need certain statistics of roughness, since the roughness extraction process does not fully map the image of the sidewall. These are RMS roughness and correlation length which can fulfil all of our requirements in order to model sidewall roughness having identical roughness statistics and data. In this case a large enough scanning area in horizontal direction at the optimum edge of a waveguide can provide enough information to characterise the roughness in all over of the sidewalls. This approximation is valid as long as we have identical correlation length and RMS roughness in vertical direction of sidewall.

The theoretical models discussed here are either complex and/or very limitedly applicable. The experimental techniques are weak to differentiate between scattering losses and other source of loss such as absorption. In addition, they cannot directly explain the relation between roughness statistics and the scattering loss. In the next chapter a novel approach will be introduced and discussed which is applicable to a wider range of geometries and can provide a well presented illustration of the interdependence between scattering loss and statistical roughness data.

Chapter 4

Scattering Loss Estimation Using a New Approach

4.1 Introduction

In the previous chapter, a number of the most successful analytical and accurate experimental techniques to estimate scattering loss were briefly reviewed. Despite of the advantageous that each of these methods offers, they have various limitations. For example, in order to use loss experimental measurement techniques real waveguide samples are needed. The fabrication process of waveguide samples may require expensive clean room instrumentations and time consuming steps. On the other hand analytical methods can be faster, but their applicability is limited because of the approximations used in these methods. For an instance Payne and Lacey′s method was originally developed for 2D waveguides, and the effective index method can only exhibit accurate results in waveguides with high aspect ratio cross sections [52]. There are also other methods reported in the literature, such as Volume control and Poulton semi-analytical method [52, 77]. The early representation of volume current method is only accurate for low index contrast waveguides with small refractive index difference and may be not applicable in strongly guiding structures [82]. The method has been revised and corrected by Johnson et al [83] for high index contrast waveguides. Nevertheless these methods have been only developed for certain waveguide geometry mainly rectangular or circular. In addition, the corrected version of volume current method has been originally developed for periodic structures such as the photonic crystal where the waveguide scattering loss performance is not characterised by pure random roughness. Hence it does not provide a clear explanation of the relation between the roughness statistical data and the resulting scattering loss [83].

In this section the development of a simulation based method is explained which estimates scattering loss due to numerically and randomly generated 3D waveguide roughness. The method to deduce roughness is developed based on an inspiration from the fundamentals of Magnetic Resonance Imaging (MRI) technique [84]. The later result shows that the model has a comparable accuracy to the commonly used scattering loss estimation technique developed by others, in particular Payne and Lacey′s [6]. Depending on the fabrication process optical waveguides that can have either isotropic or anisotropic sidewall roughness [61]. Using this approach one can study 3D structures with roughness that are anisotropic, isotropic and mixture of isotropic and anisotropic roughness. It can also be implemented on arbitrary high/low index contrast waveguide shapes such as cylindrical, triangular geometries and multilayer structures. In order to extend its applicability to wider applications using this model, roughness can be induced on top surface of the waveguide if it is required. In addition the model can be widely used in any commercial software using Finite Difference Time Domain algorithm in order to estimate the effect of scattering loss for various photonic devices and applications. The method will also make one able to study the scattering effects independently from other source of losses such as absorption. The method provides a useful tool to optimise fabrication process before initiating, since it gives an accurate estimation of the scattering loss based on the roughness statistical parameters. The accuracy can be improved by using a cluster with more available memory space, where it requires, in order to provide finer mesh settings.

In summary, the model uses two main parameters, which characterize the roughness as mentioned in Chapter two, the correlation, L_c length and the root mean square (RMS) roughness σ, which is given by a standard deviation from the averaged flat surface. The image is considered as a two-dimension signal where its frequency spectrum can be represented using two-dimensional Fourier transform. The 2D Frequency spectrum contains horizontal and vertical spatial frequencies, which is a simplification model of k-space in MRI. A randomly generated spatial frequency spectrum is virtually created in commercial software, Lumerical script environment, which can also be independently executed in any scripting environment such as Matlab. It is then filtered according to the real roughness data extracted from AFM (Atomic Force Microscopy) images, and the actual surface is realized by applying the Inverse Fourier Transform (FT) on the filtered spatial frequency spectrum (k-space). The generated surface roughness is then implemented as sidewalls of the optical waveguide, although can be also applied on top surface of the waveguide if it is required. With this technique, anisotropic and isotropic sidewall roughness features can be accurately modelled based on surface data extracted from AFM. I use 3D FDTD technique to estimate the induced loss due to the sidewall roughness of the optical waveguide structure. The accuracy of the loss model is also dependent on the mesh resolution in FDTD.

4.2 Generation of random roughness on rectangular optical waveguide sidewalls

The method used to model roughness is based on Fourier analysis approach used in magnetic resonance imaging (MRI) technique [84, 85]. The roughness of a flat surface can be considered as a two dimensional image where each point (pixel) has different random correlated intensities. The correlation of these distributed random intensities in the image is defined by a suitable Auto Correlation Function (ACF). The ACF explains how random variables are related statistically. The ACF of a random roughness is characterised by correlation length, L_c, and mean square deviation, σ^2, from a flat surface without irregularities. The most commonly used ACFs for process dependant roughness are exponential and Gaussian [6]. AFM measurements on silicon optical waveguide surface roughness have demonstrated that Gaussian ACF is a realistic model for the statistics of surface roughness of a variety of waveguides [6]. The Gaussian ACF is expressed as:

$$R(u) = \sigma^2 exp(-\frac{u^2}{L_c^2}) \tag{4.1}$$

Here the image is treated as a random two-dimensional signal. The 2D Fourier spectrum of this signal is generally represented in a different coordinate so called k-space. K-space is a 2D Fourier transform of the original image in real space as illustrated in Figure 4.1. Each point in k-space is a number where its horizontal and vertical spatial frequencies are obtained in horizontal and vertical axis respectively. Hence, it contains all required spatial frequency and corresponding amplitude information of the image (all signals forming a 2D image) in real space. This is shown in Figure 4.1 for only one of the many 2D signals in real space. The created k-space typically has the same number of rows and columns as the original image. Low frequency components (long wavelengths) are stored around and at the middle and high frequencies components (short wavelengths) near the periphery region. Low spatial frequencies produce only waves with long wavelength characteristics in the original image. These waves only provide contrast information and give little data for the edges. On the other hand high spatial frequencies eliminate many contrast information in the resultant image, since contrast data is mainly formed by waves only with long wavelength characteristics. Image features that change in intensity over short image distances are high frequency component. Those image features that change in intensity over long image distance are low frequency component.

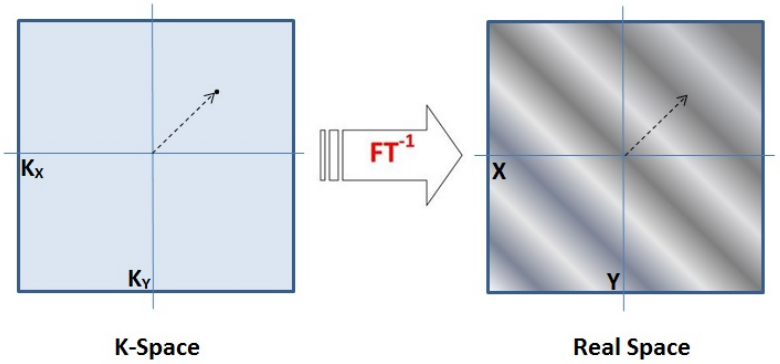

Figure 4.1: The real and K-space relation according to Fourier analysis; in the K-space we express the magnitude and the wavelength of all signals that exist in the real image. Here as an example, the Fourier spectrum of a 2D signal in the real space is expressed as a dark point in k-space. The brightness of the point expresses its amplitude and its position in K-space represents horizontal and vertical spatial frequencies. Note that the direction of a 2D signal in real space matched to the direction of position vector of its Fourier equivalent in K-space.

The general important properties of 2D spatial frequency spectrum is as follows:

1. One element in the spectrum contains information about all pixels of the real image.
2. The size of the 2D spatial frequency spectrum is the same as the size of the real image signal as defined by the 2D discrete Fourier transform 4.4;
3. In order to include all statistics of the roughness, the image and hence the frequency spectrum size has to be chosen appropriately large, so it is always grater than the selected correlation length.

The above argument for two dimensional signal is also valid for a one dimensional signal. It is therefore helpful to initially consider a simplified scenario dealing with a one dimensional signal before going through the details of how a two dimensional random signal can be generated through accessing its spatial frequency properties. This will more clarify how changing the spatial frequency properties of the signal through applying the Fourier representation (PSD) of ACF in frequency spectrum can modify the original signal properties in real space. Assume that we have been given a discrete number of one dimensional random constituent periodic waves of this signal are shown in Figure 4.2.

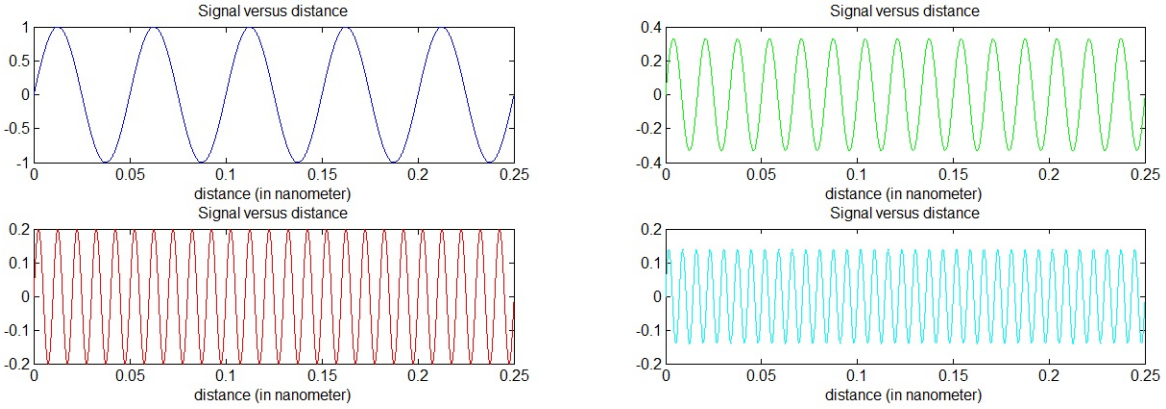

Figure 4.2: Random constituent periodic waves

The constituent periodic waves differ from each other according to their specific spatial frequencies and perhaps the relating amplitudes (or phase). Based on Fourier analysis, the superposition of these constituent waves will form a periodic one dimensional that is shown in Figure 4.3. Note that for non-periodic signals the Fourier spectrum can be continuous containing a range of frequencies.

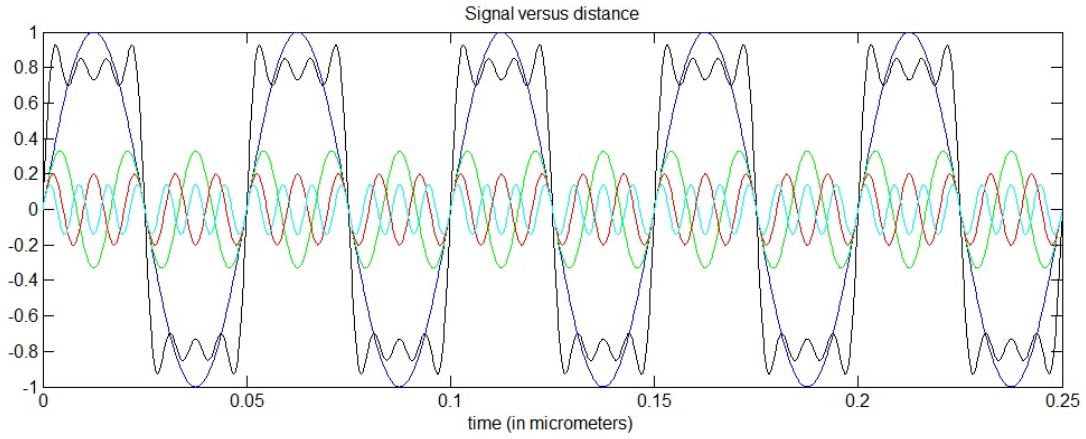

Figure 4.3: The constructed one dimensional signal created from the superposition of its constituent elements

Figure 4.4 illustrates the relating Fourier representation of this one dimensional signal. The graph includes information about the amplitude corresponds to each spatial frequency component (real values only). This Fourier representation of one dimensional signal is the equivalent for of the k-space representation of two dimensional signals.

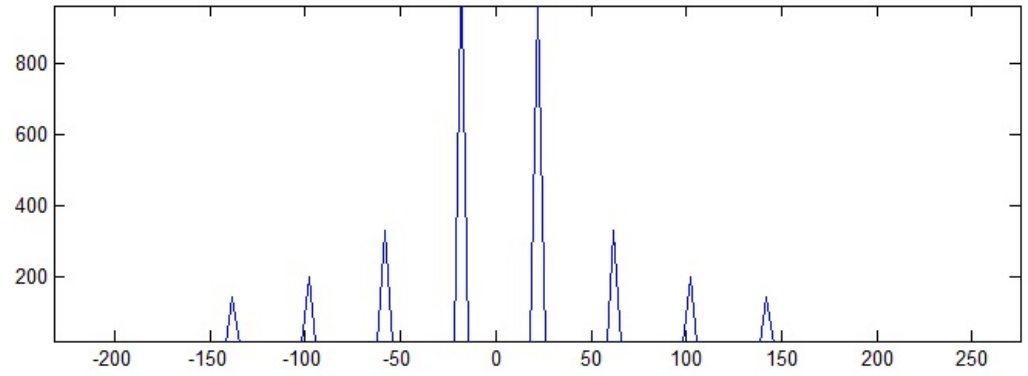

Figure 4.4: Frequency domain representation of the original signal

As can be seen in Figure 4.4 contains both high and low spatial frequency components. However, in Figure 4.5 a and b we can see how the signal in real space is modified when a few of high frequency components are removed from the Fourier spectrum. As can be clearly seen the eliminating these components removes the abrupt changes in space domain leaving the signal with more gentle and smoother osculations, pretty much like a simple sine wave. Contrariwise removing low frequency components from the Fourier spectrum results vanishing the contrast information from the original signal leaving the signal to more severely and abruptly change in the space domain.

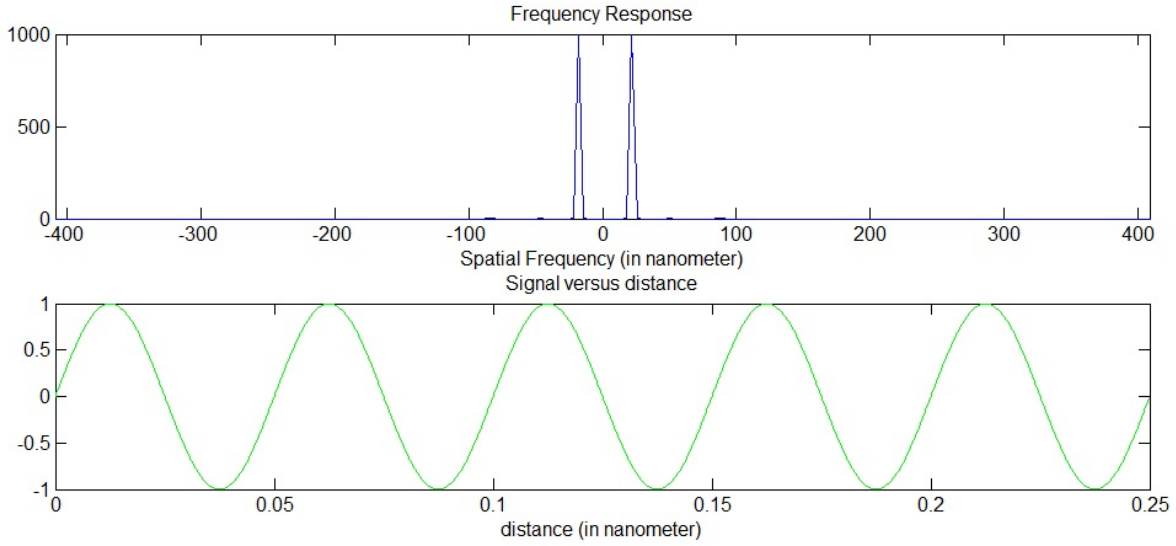

Figure 4.5: a) Frequency domain representation of the manually filtered signal
b) Manually filtered signal in real space

In Figure 4.5 the high spatial frequency component have been removed manually. Similarly frequency response of the Gaussian ACF acts as a low pass filter which removes the high spatial frequency components taking the roughness statistical data into account (the mean square root and correlation length). Figure 4.6a shows how the 1D frequency

response of the Gaussian ACF filters the high frequencies from the Fourier spectrum of the signal. The signal in space domain is also generated using Inverse Fourier Transform of the response. This is shown in Figure 4.6b.

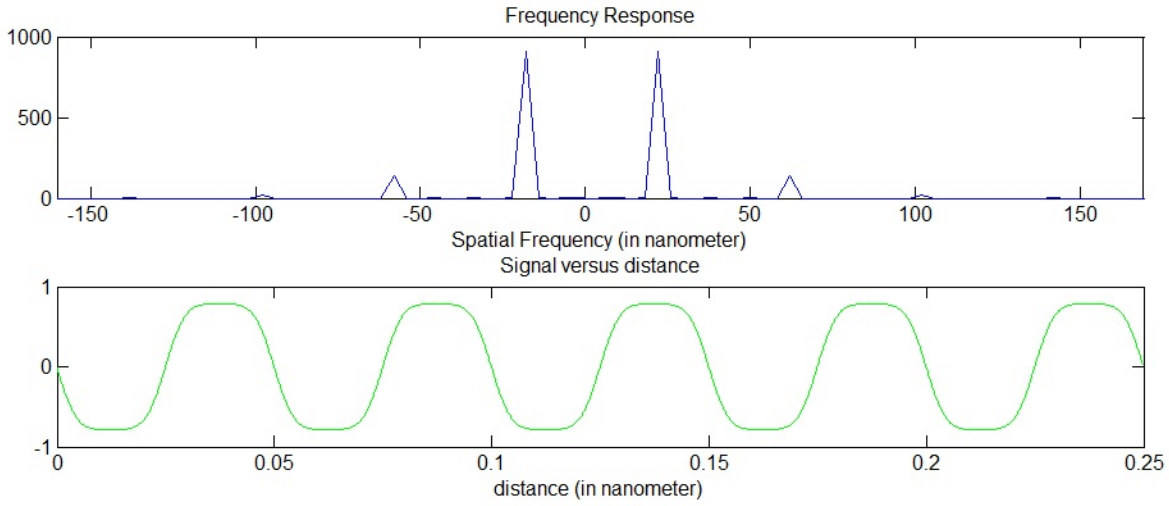

Figure 4.6: a) Frequency domain representation of the manually filtered signal using Gaussian ACF b) its equivalent in real space

As can be seen in Figure 4.6b, the edges become smoother now (high frequencies component) and the sampled points on the signal is more correlated together. As will be shortly explained the same concept is used for a two dimensional signal (an Image) to generate a realistic roughness based on the ACF defined in Equation 4.1. Figure 4.7 is the flow cart diagram explaining the steps have been taken to produce two dimensional random roughness model.

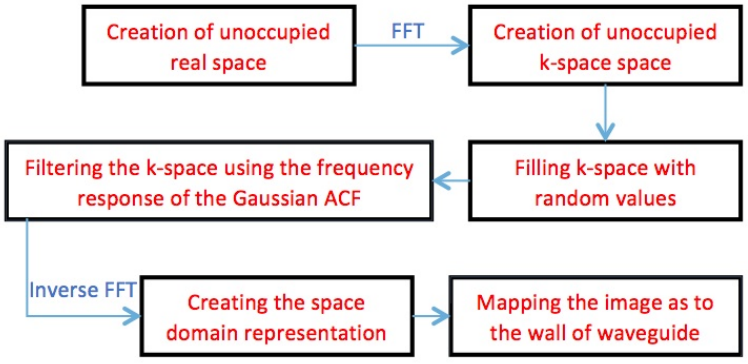

Figure 4.7: Flowchart diagram of generating roughness model

treat the a 3D roughness as a 2D image signal. Firstly, the sidewall roughness is implemented in a rectangular waveguide. Hence the roughness can be initially generated on a flat surface S that is later mapped onto the sidewalls. In order to assign position to different points of the surface S, it has to be mapped in to the real space. At this

stage the surface is considered to be completely flat without having roughness. This means the intensity of all points in the surface is initiated to be zero and hence we call it unoccupied real space. In order to expresses our proposed surface S as an unoccupied real space, the surface S is represented by an n by m matrix. The matrix represents surface S is denoted as S_c. The entries of the matrix S_c expresses the position of the associated element in the surface S while the values of the entries express the height level in the surface S. The roughness will be generated above the surface by assigning correlated intensities to the elements of the matrix S_c.

The discrete real space is virtually created by defining two position matrices X and Y that has the same dimension as matrix S_c. This is shown in Figure 4.8. They respectively represent x and y components of the position vector of each element (point) in the matrix S_c (surface). Hence any element (intensity) in matrix S_c has an equivalent element in X and Y that specify its position in x and y coordinate of the real space.

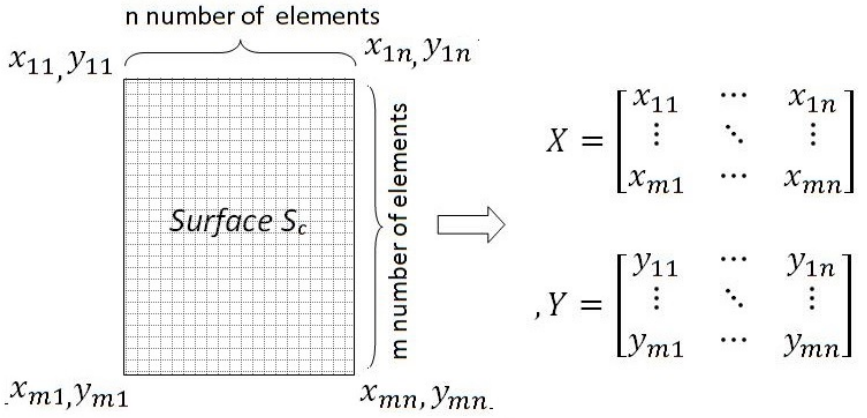

Figure 4.8: The creation of discrete real space that contains $n \times m$ elements. X and Y respectively store the x and y coordinates of all elements in the surface S_c.

In this case:

$$\begin{aligned} x_{11} = x_{21} = \cdots = x_{m1} \\ \vdots \\ x_{1n} = x_{2n} = \cdots = x_{mn} \end{aligned} \tag{4.2}$$

since X only represent the x components of the position vector and y represents the y components of the position vector;

$$\begin{aligned} y_{11} = y_{12} = \cdots = y_{1n} \\ \vdots \\ y_{m1} = y_{m2} = \cdots = y_{mn} \end{aligned} \tag{4.3}$$

Once the real space is created the relevant k-space can be also formed. Considering Equations 4.5 and 4.6 , the required spatial wavevectors $\mathbf{K}_x$ and $\mathbf{K}_y$ of an unfilled k-space associated with a Fourier transform of a function of the elements of X and Y is generated based on the 2D discrete Fourier transform relation;

$$h(K_x, K_y) = \sum_{x=0}^{n} \sum_{y=0}^{m} e^{i(K_x x + K_y y)} s_c(x, y) \tag{4.4}$$

Using equation 4.5 and 4.6. the corresponding k-space is generated which is graphically shown in Figure 4.9a.

$$K_x = \frac{2\pi}{dx.n} x_{nm} \tag{4.5}$$

$$K_y = \frac{2\pi}{dx.n} y_{nm} \tag{4.6}$$

Here dx represents the spacing between the values of the input x, n and m denote the length of x and y respectively. In addition, both x and y domain is extended from negative to positive values and therefore the generated k-space includes both negative and positive frequencies and phases. When the k-space is constructed, it is filled by uniform random numbers Figure 4.9b. Each point in k-space has now a random value, which expresses the amplitude of a trigonometric signal at specific frequency and phase. These randomly generated data are stored in a new n by m matrix that is called $\mathbf{Z}$. The matrix $\mathbf{Z}$ is used to express the k-space of the matrix S_c.It contains the intensity values correspond to $\mathbf{K}_x$ and $\mathbf{K}_y$. According to Fourier theory the superposition of all of these signals will form the final surface S_c. This is where the ACF is used. Applying ACF to the matrix $\mathbf{Z}$ provides a low pass filter that removes high frequency components in the k-space, such that the resultant image does not have much of the edge information which are defined by high spatial frequency components. This removes the edge details of previously generated uncorrelated random heights. According to Equation 4.1 the Gaussian ACF is related to the correlation length L_c, and mean square deviation σ^2. Here sigma σ is defined as:

$$\sigma = Z(n, m) - \frac{\sum_0^n \sum_0^m Z(n, m)}{n \times m} \tag{4.7}$$

However the Gaussian autocorrelation function cannot be directly applied in the 2D spatial frequency spectrum in order to filter out high frequency components. In fact the Fourier representation of Gaussian ACF, the power spectral density (PSD), has to be used. The PSD of the Gaussian ACF can be simply calculated using 3.25. The filtered data is then stored in a matrix called $\mathbf{Z}_{filtered}$.

$$Z_{filtered} = Z(n,m)^2 \times exp - (\sqrt{{K_x}^2 + {K_y}^2}/L_c)\sigma = Z(n,m) - \frac{\sum_0^n \sum_0^m Z(n,m)}{n \times m} \quad (4.8)$$

Hence a certain point at r position in k-space $\sqrt{{K_x}^2 + {K_y}^2}$can be correlated with a point at $r + \sigma$ using Equation 4.8. The resultant K-space, is shown in Figure 4.9c. The resulting values in Figure 4.93 c are then transformed back to real space, as can be seen in Figure 4.9 d;

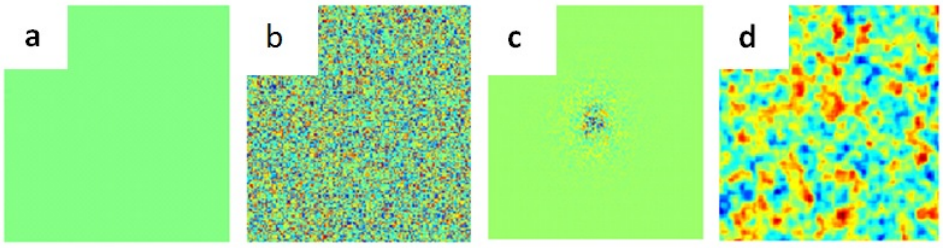

Figure 4.9: The procedure of roughness generation: a) Creation of unoccupied k-space, a coordinate showing all available wavelengths for the signals in real image b) k-space with randomly generated elements, to include magnitude of each signal c) Filtered k-space to remove the edge information d) the resulting roughness in real space.

Figure 4.9 d shows the roughness in real space is completely random and characterised to retain the isotropic behaviour. Two different randomly generated S_c surfaces can be imported on both sidewalls of 3D planar waveguide for further studies on propagation loss characteristics of optical waveguides. Isotropic sidewall roughness and loss analysis has been previously reported in [61]. Accurate AFM/SEM images show even on sidewalls with anisotropic roughness, there is still a degree of isotropic roughness features. The difference is illustrated in Figure 4.10. Such a characteristic is often ignored in scattering loss evaluations and mathematically reported roughness models.

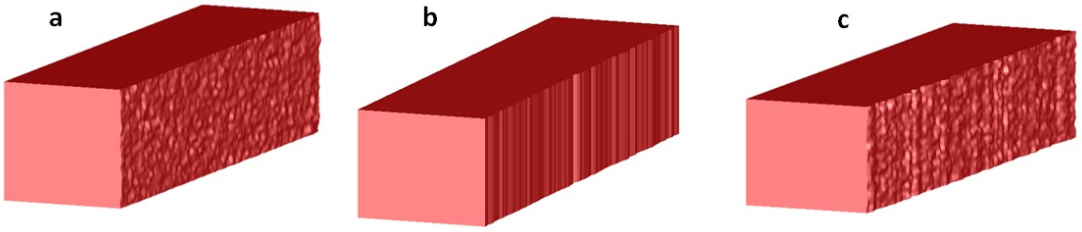

Figure 4.10: Different type of roughnesses, a) Isotropic b) anisotropic, c) mixture of isotropic and anisotropic roughness

Isotropic roughness is simply achievable, once the roughness surface in real space is created. In order to generate anisotropic roughness the following steps are performed. A one dimensional random roughness profile can be configured by accessing a random

row of the matrix that expresses surface S_c. When such a profile is repeated in y direction, a new matrix with identical row will be formed. This newly constructed matrix, $S_{anisotropic}$ is used to represent the completely anisotropic surface.

The integration of S_c and $S_{anisotropic}$ anisotropic will form the optimal sidewall roughness model that is anisotropic in y direction, but will still have a level of isotropic behaviour. This procedure is shown in Equation 4.9.

$$\begin{matrix} x_{11} \cdots x_{1n} \\ \vdots \ldots \vdots \\ x_{m1} \cdots x_{mn} \end{matrix} + \begin{matrix} x_{11} \cdots x_{1n} \\ \vdots \ldots \vdots \\ x_{11} \cdots x_{1n} \end{matrix} = \begin{matrix} x_{11} + x_{11} \cdots x_{1n} + x_{1n} \\ \vdots \ldots \vdots \\ x_{m1} + x_{11} \cdots x_{mn} + x_{1n} \end{matrix} \tag{4.9}$$

The mixture is formed by superimposing a random $S_{anisotropic}$ matrix to a random S_c. The matrices $S_{anisotropic}$ and S_c can have either different or identical roughness properties if required. All three types of roughnesses from Figure 4.10 are fully attainable using the same modelling technique. This makes the model reliable and applicable for roughness studies in arbitrary waveguide devices. In addition the isotropic roughness can be applied on top surface of the waveguide if required.

4.3 Simulation

The sidewall roughness parameters of the waveguides are extracted from its AFM measurements and imported to the roughness model in order to demonstrate the capability of this technique to estimate total scattering loss. In addition to fabricated waveguides explained above, various waveguides with obtainable roughness features can also be used for the same purpose as will be explained later.

The roughness model has been successfully configured in FDTD Lumerical software package and MATLAB environment. However, in order to run scattering loss simulations of the waveguides, it is preferred to use FDTD Lumerical package, because of its particular functionality and supportive features.

The fundamental TE mode is launched into the Si waveguides at a certain wavelength range from 1550 nm to 3800 nm. The two power monitors record the power transmission data through the waveguide. The first monitor is adjusted at the entry point of the optical waveguide to collect all transmission power information before the light is scattered with generated sidewall roughness. The second one stores these data after 10 μm length.

The calculation of loss using such a method becomes complicated since the problem is numerically very large. A very fine mesh setting is necessary in order to resolve the features on the 3D waveguide sidewalls. This substantially increases the computational

time. In addition, the time-step must be reduced in order to keep within the Courant limit which makes the calculation time even longer [88]. The Courant is a parameter that has to be satisfied in order to ensure stability of the simulation. It is given as $c\frac{\Delta t}{\Delta x}$ where c is to $1/\sqrt{\mu_0 \varepsilon_0}$, Δt is the time step and Δx is the grid size. In order to satisfy the Courant limit for three dimension structures the following condition has to be considered:

$$c\frac{\Delta t}{\Delta x} < 1/\sqrt{3} \tag{4.10}$$

Although this problem seems to be the major issue to achieve accurate results, tradeoffs in mesh settings can be applied to obtain reasonably accurate results.

4.4 Measurement of Fabricated Silicon Waveguide

Ion bombardment in the dry etch process is considered to be one of the major origins of sidewall roughness. This kind of roughness can be significantly reduced using a low gas pressure and a low radio-frequency power [86, 87]. Pattern transfer during the etch process is also considered as one of the origins of sidewall roughness [87]. Several waveguide structures from two different wafers were used to obtain experimental data. Firstly, Si strip waveguide structures were fabricated using 500 nm thick SOI and 2 μm thick buried oxide (BOX) layer. The other sample contains 220 nm width waveguides on top of the 3 μm BOX layer.

The waveguide dimensions in the first sample are 1000 nm, 1200 nm and 1400 nm in width and 500 nm in height where the waveguide can be single mode in the MIR. In the second sample waveguides have 220 nm height and 330 nm width which are even single mode at lower wavelengths around 1.3 and 1.55 μm. The SEM image of the fabricated waveguide is shown in Figure 4.11. These wavelengths have been chosen to show the applicability of this method at both NIR and MRI wavelengths.

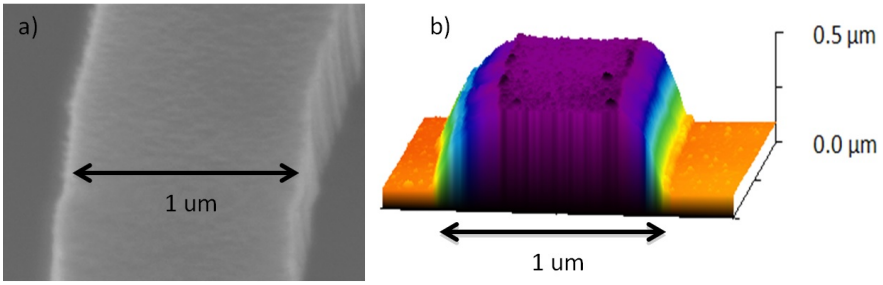

Figure 4.11: a) SEM image of the fabricated waveguide, b) 3D AFM view of the fabricated waveguide. The waveguide has 500 nm height and 1000 nm width. The roughness parameters are extracted from the sidewall of this waveguide.

Once the samples are prepared, Atomic Force Microscopic system with special high aspect ratio probe is used to generate multiple AFM images of the optical waveguide. The tip length of the probe is 2 μm and tip radius 10 nm. Figure 4.11b shows a sample 3D view of the AFM image of the silicon optical waveguides we have fabricated together with their SEM images. The two main roughness parameters, RMS and correlation length is extracted from the edge of sidewalls in AFM image.

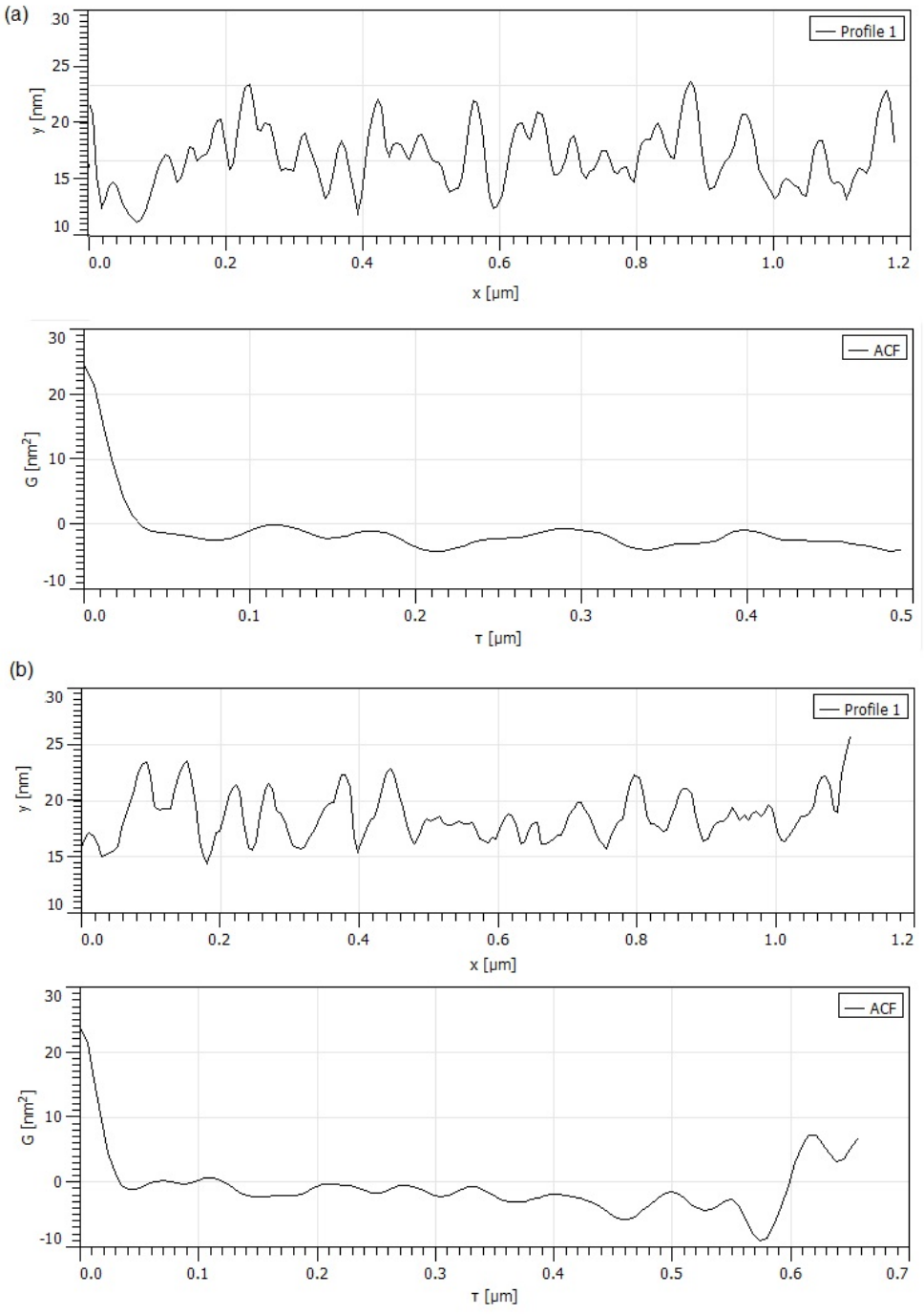

Figure 4.12: The Roughness Profile extracted form the edge of the waveguides for a) 1000 nm (width) $\times$ 500 nm (hight), b) 330 nm $\times$ 220 nm and their relevant ACF

The RMS is extracted directly from the profile and the correlation length is found from the Autocorrelation function of the profile as explained in previous chapter 3. The value of L_c equals to the minimum length that makes the normalized autocorrelation minimum ($\approx$ zero). Figure 4.12 a and b shows the extracted profiles and the corresponding ACFs of two samples of these waveguides.

Reproducible RMS of 5 nm and of L_c 45 nm were obtained from multiple samples for 1000 nm (width) $\times$ 500 nm (height) waveguides. RMS and L_c for 330 nm $\times$ 220 nm waveguides are approximately the same since the same fabrication process is used. Once these parameters are extracted from the sidewall of the waveguide they can be implemented into the simulation model to estimate the scattering loss.

The propagation losses of the waveguides are experimentally measured using Fabry-Perot techniques and cut back method in Southampton and the university of Surrey with the collaboration of Milosevic at. al. Figure 4.13 illustrates Fabry-Perot resonance response for 1000 nm $\times$ 500 nm waveguide at the 1550-1551 nm wavelength. Considering an ideally polished facet Si waveguide the reflectivity can be approximated as 0.303.

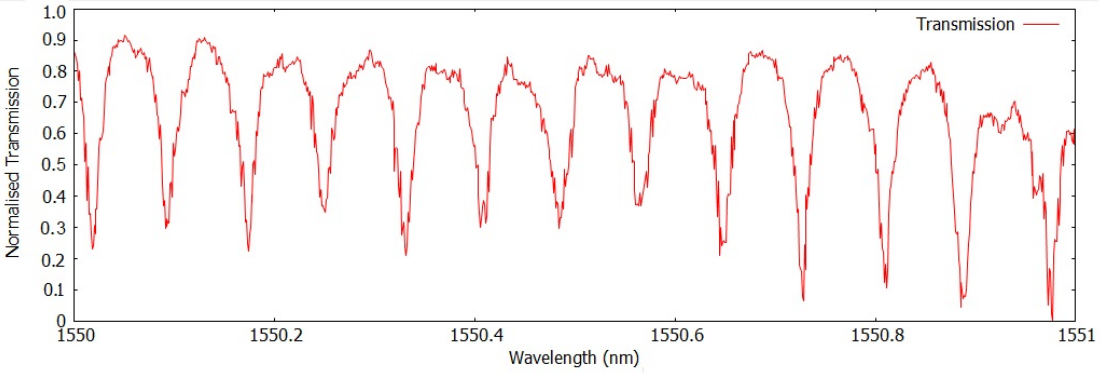

Figure 4.13: Fabry-Perot resonance in waveguides having a dimension of 1000 nm (width) $\times$ 500 nm (height)

Using the extracted values of I_{min} and I_{max} from 4.13 and equation 3.7 the propagation loss can be calculated. Having 0.25 cm long device the loss is worked as 1.2 dB/cm. The laser output power used in all measurement was 3.98 mW. A significant amount of power is lost within the coupling at the facet. The resonant spacing also confirms that the resonant peaks are generated by the facets.

The same procedure is repeated for different Si waveguides at different wavelengths. Hence one can observe the propagation loss characteristics of these waveguides as shown in Figure 4.14.

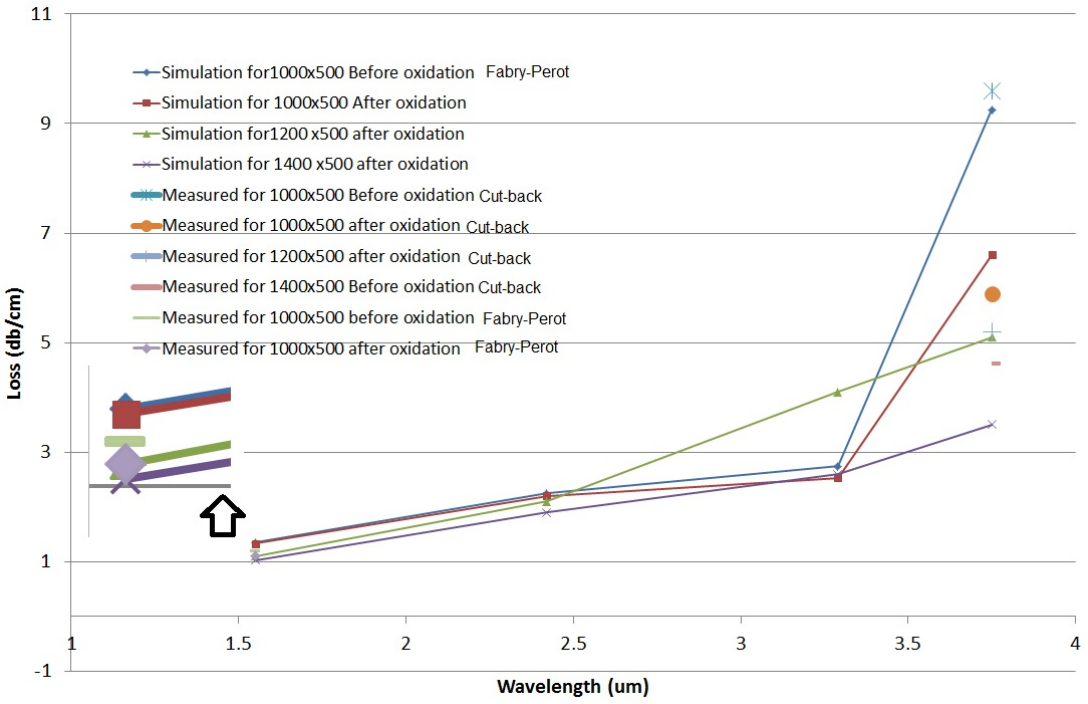

Figure 4.14: Loss measurements compared with the simulation results for waveguide ranging from 1000 nm to 1400 nm width.

4.5 Discussion

The scattering loss were obtained in dB per unit of length from simulations as shown in Figure 4.14. Figure 4.14 also shows a good agreement of simulation and measurement results that are obtained using Fabry-Perot method previously.

I demonstrate the scattering loss by simulating the propagation of light in optical waveguides with the isotropic and anisotropic roughness. In addition to fabricated waveguides explained above, various waveguides with obtainable roughness features can also be used for the same purpose. The simulation of the sidewall roughness model shows comparable results with other analytical models and a number of published results. However, the comparison with reported results in literature is difficult since there are various physical parameters that affect the losses. These parameters are mainly related to the cross section of the optical waveguide, the wavelength of light for given dimensions and also the statistical properties of the sidewall roughness.

I therefore investigate the effect of the reduction in waveguide dimensions on the transmission loss. The transmission loss is calculated in three optical waveguides with different sidewall roughness styles. The SOI waveguides have a 200 nm fixed height and width varying from 400 nm to 4.5 μm. These dimensions provides high aspect ratio guide that can help to compare results with estimations obtained by Payne and Lacey′s

2D model [6]. The roughness parameters that are used are L_c=50 nm and σ=9 nm. These are the values measured from identical waveguides in [89].

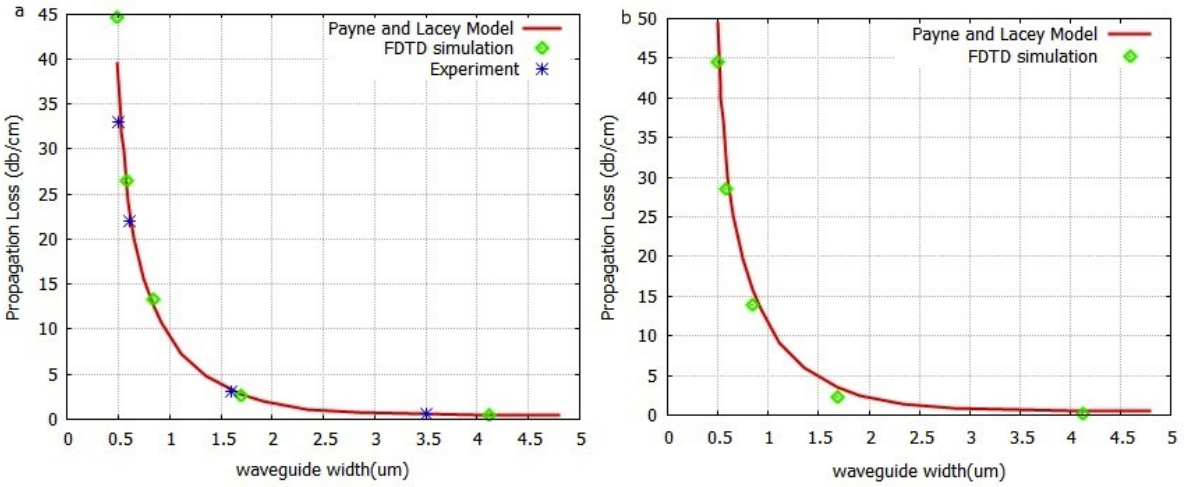

Figure 4.15: Comparison of scattering loss of Si waveguides with different widths obtained from reference [89], Payne and Lacey, and the model for a) SOI waveguide with a 1 μm deposited layer of SiO_2 added on top of the 0.2 μm. b) SOI waveguide with no additional SiO_2 layer on top. In both cases here Lc is 50 nm and σ is 9 nm.

I consider the waveguide in two different situations. Firstly a deposited 1 μm layer of SiO_2 is added on top of the 0.2 μm SOI waveguide for an accurate comparison to the experimental results obtained in [89].In the other case the additional SiO_2 layer is not applied. The effect of size reduction on propagation loss is shown for both cases in Figure 4.15 a and 4.15 b respectively. Here waveguides have anisotropic rough sidewalls with a level of isotropic behavior. The isotropic statistical components are considered to be 5 times larger than anisotropic statistical component. Having fully anisotropic sidewall roughness gives identical values to those shown in Figure 4.15 a and b. The loss is as low as 2 dB/cm for waveguide width above 2.5 μm, indicating low scattering losses at the core - cladding interface. However, losses rapidly increase for widths below 2.5 μm and are as high as 44 dB/cm for 300 nm wide waveguides due to increased interaction between the mode propagating in the waveguide and the sidewalls. Clearly from the Figure 4.15 a, this is again in excellent agreement with the results reported in [89].

In order to estimate scattering loss based on [6] approach for the above 3D structure, the effective index method has to be implied in Equation 4.11.

$$\alpha_{radiation} = \frac{\sigma^2}{\sqrt{2}k_0 d^4 n_1} g.f \tag{4.11}$$

where we therefore have:

$$\alpha = 4.34 \frac{\sigma^2}{\sqrt{2} K_0 d^4 n_1} g.f \tag{4.12}$$

The calculated scattering loss based on this approach is also shown in Figure 4.15 a and b. As can be seen the trend of the simulation results is in an excellent agreement with experimental measurements obtained by Lee et al. [89] and theoretically estimated by Payne and Lacey. The validity of this trend is completely retained for uncovered SOI waveguide in Figure 4.15 b and still stays in a good agreement with Payne and Lacey calculated results. This result shows that when the roughness data is fixed and unchanged (RMS and correlation length) and varying geometrical parameter (width or height), comparable estimations can be obtained using either Payne and lacey [6] or the approach that is used in this work.

From Figure 4.15, the interaction of the mode with the sidewalls is strongly enhanced when the waveguide cross-section is decreased. Therefore an accurate comparison can only be possible for optical waveguides with comparable cross-sections (comparable propagation constants) and geometry.

Work	Dimension nm×nm (W×H)	Correlation Length(nm)	RMS (nm)	Loss (dB/cm) measured	Loss(dB/cm)- Simulations		
					Isotropic	Anisotropic	Mixed
The First model (at 1500nm)	330×220	45±5	5±1	33	36.43	29.84	28.65
The second model (at 3800nm)	1000×500	45±5	5±1	10.2	11.06	10.04	10.53
(1550nm)	445×650	225	5	2.1±0.6	2.6	1.72	1.79
[61](1550nm)	445×220	50	2	3.6±1	5.2	4.5	4.2
[89](1550nm)	500×200	50	9	33	47.32	36.3	36.1
[71](1550nm)	800×1800	225	5	1.1±1	3.30	2.71	2.34
[71](1550nm)	1000×1800	180	9	2.25±0.1	4.2	3.65	3.32
[71](1550nm)	1000×1800	160	13.5	6.5±1.5	9.63	8.94	8.16

Table 4.1: Comparison of the results obtained from the purposed model with a number of experimental or calculated models. The first row corresponds to the waveguide that is shown in Figure 4.11 where the data is directly extracted from the AFM image and imported into the Lumerical. The second row is the result for a waveguide with arbitrary dimensions. The other results that are shown here, prove the validity of the discussed model as it matches other reported experimental work.

In table 4.1 the scattering loss simulation results, using the three roughness models, are compared to several published experimental results. I again use the two waveguide samples with two different cross section sizes (220nm×330 nm and 1000 nm×500 nm in height and width respectively). These dimensions make the cross section of the optical waveguide more identical to other reported works in the literature that I am referring to. The wavelength is again fixed at 1550 nm.

In the second work I simulate a large waveguide shown in Figure 4.16 in MIR wavelength range at 3.8 μm where the waveguides operates in single mode. All other dimensions are ranged between 300 nm to 650 nm in height and 300 nm to 1000 nm in width. The statistical parameters (RMS and L_c) are process dependent factors and therefore they vary in a wider range. From Table 4.1 it can be seen that the simulation results stay in a good agreement with experimental reported data. It is also important to remember that the loss is measured applying different measurement techniques such as Fabry-Perot or the cut-back. Our AFM measurements show that the anisotropic behaviour of the sidewall roughness is more enhanced compared to the isotropic characteristics. In all our samples the RMS we have measured for anisotropic roughness is at least 7 times larger than that for the isotropic roughness of the sidewall surface. This makes the anisotropic component dominant in the mixed case as can be seen in Table 4.1. Furthermore, from the data it seems that anisotropic roughnesses and the mixed isotropic and anisotropic roughnesses more realistically represents the form of actual sidewall roughness, as confirmed by direct measurement of the sidewall roughness.

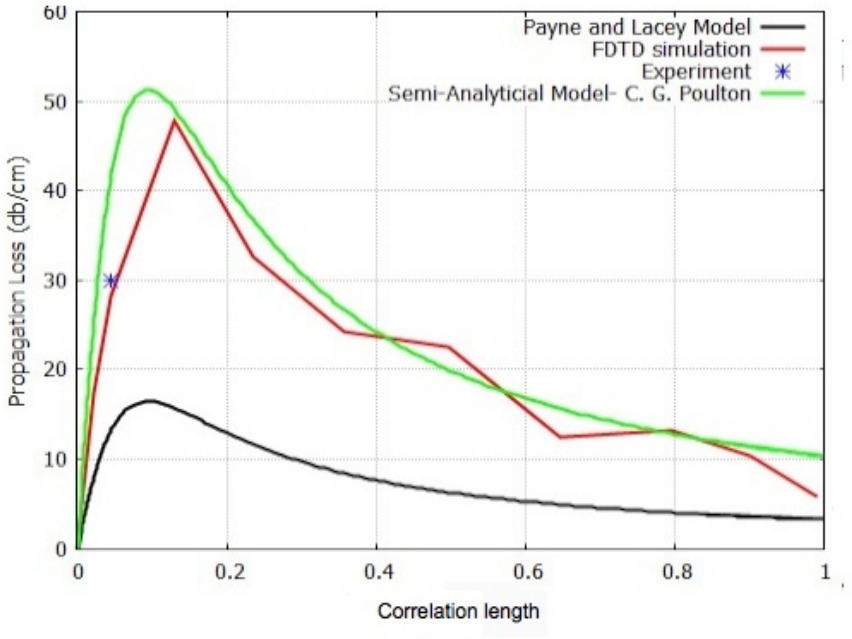

Figure 4.16: Propagation loss against the correlation length of the sidewall roughness the estimated results is compared with Payn and Lacey′s model [4] and the experimental result.

Based on the data comparison in Table 4.1 and the discussion on the relationship between the waveguide dimension and the loss, the model can be extended to show the dependence of loss on sidewall roughness parameters for given waveguides. Figure 4.16 graphically illustrates the scattering loss dependence on the variation of correlation length when σ is 5 nm. Here we used a SOI waveguide that is 220 nm by 330 nm in height and width respectively. This waveguide has relatively low aspect ratio.

Figure 4.16 shows that the FDTD simulations in comparison with the predicted results obtained by Payne and Lacey′s model [6]. The simulation results show the validity of the experimental data. However for this relatively low aspect ratio waveguide the Payne and Lacy estimations does not appropriately match to the FDTD simulations. As expected from the theory [6], when the correlation length of the sidewall roughness is longer than the effective wavelength the scattering loss is reduced. This is because the variation along the direction of propagation is much larger than the effective wavelength. However, having a very short correlation length causes the waveguide boundary to change very fast. Therefore the propagating mode cannot determine the details of the sidewall and consequently the scattering loss is reduced again. The interaction between the propagating mode and sidewall roughness is maximized when the correlation length is not very short and very long compared to the effective wavelength.

In general, all above results shows that the model has a comparable accuracy to the commonly used scattering loss estimation technique developed by Payne and Lacey [6] and others [52]. Nevertheless one important obstacle in using this model is the constraint caused by finite available memory space. The limited memory space can restrict the applicability of this approach for certain 3D structures; in particular for those structures where the dimensions are large in comparison with the extracted roughness statistical data; for example for a large dimension waveguide with very delicate sidewall surface fluctuations. The effect of roughness in such a large structure can be only obtained by having impractically large number of fine meshes and a huge indicated memory space. Such a requirement may be only realized by having a computer cluster facilities. Otherwise it is essential to precisely tune the mesh settings against the required memory space. To overcome this problem, there are mainly two possible solutions which can help to efficiently adjust the mesh settings:

1. Conformal and none uniform meshing can be applied in some cases in which the surface roughness is localised, for example on sidewalls (i.e. the top surface is considered to be ideally smooth).This can significantly reduce the required memory space. In other cased if the structure has symmetric boundaries

2. Symmetric boundary conditions are used when the user is interested in a problem that exhibits one or more planes of symmetry. Both the structure and source must be symmetric. Symmetric boundaries are mirrors for the electric field, and anti-mirrors for the magnetic field. Careful consideration must be given to whether

symmetric or asymmetric boundary conditions are required, given the vector symmetry of the desired solution. For meaningful results, the sources used must have the same symmetry as the boundary conditions.

Taking the limited available memory space in to account, the method can serve much better accuracy for smaller structures and nanoscale devices that have relatively more enhanced roughness data. The model can be however implemented in any commercial software using other FDTD and other numerical mesh-bases algorithm, in despite of the memory concern. Furthermore there are various unique advantageous correspond to this model which justify its prevailing suitability, especially in more complex structures. It has been shown using this model the effect of anisotropic, isotropic and mixture of isotropic and anisotropic sidewall roughness can be individually and independently studied. Moreover the implementation of roughness can be deduced on top surface of the waveguide if it is required. The roughness effect and scattering loss can be also studied in arbitrary waveguide shapes such as the one shown in Figure 4.17, heterostructure and multilayer configurations. This unique compatibility feature of the presented method which cannot be achieved in other analytical techniques such as Payne and Lacey and effective index approaches. In addition, using this model one can estimate the fraction of the total propagating loss corresponds to the scattering loss in in highly absorbent waveguides and devices. The works and results in this chapter have been accepted for publication in IEEE photonics Journal [90].

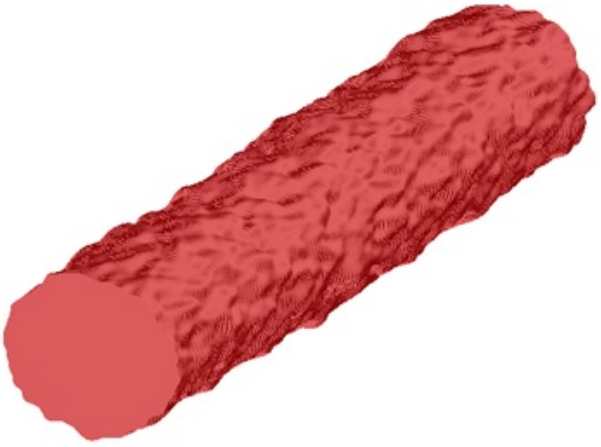

Figure 4.17: A rough circular waveguide layout.

4.6 Summary

These features together can expand the use of this model in much wider applications including more complicated waveguide based structures and photonic devices. In order to truly approve the extendable potentials of the model, it has to be therefore developed

in more advanced structures. One of the cases where the scattering loss may influence on the optimum device performance and its efficient operation is integrated carrier dispersion based optical switches. As one of the major building blocks of all optical integrated circuits , they can be realized using silicon photonic wires [42, 91, 92]. They have been extensively studied in the past few years. In order to meet the future high performance computation requirements, it is desirable to have a fast modulation and qualitative data transmission capabilities. Therefore there has been continuous trend for such an integrated modulator and other photonic components in silicon chips toward smaller scales. In contrast of this move the overall loss increases mainly as a result of coupling and greater modal interaction with sidewall fluctuations. This additional loss can undesirably reduce the final modulation depth for optical modulators for example. To have a proper analysis of performance in such devices it is extremely important to have an accurate estimation of scattering loss. Therefore in the following chapters, the performance of a novel integrated silicon modulator design operating based on carrier dispersion effect will be evaluated taking the effect of the roughness model into account.

Chapter 5

The Loss performance of a Heterojunction Optical Modulator

In this chapter the loss performance of the designed a single crystalline silicon/Amorphous zinc oxide heterojunction optical switch will be investigated. Hence, the primary aim of this design is to develop a functional model in which the effect of roughness on the device loss performance can be accurately investigated using the roughness model discussed in chapter 4. The device is purposely designed and analysed to serve low optical loss, high extinction ratio and comparable overall footprint and switching speed with recent reported works. The full design procedure is explained in details in Appendix. However more improvement in the performance may be possible by further optimization on structural aspects of the device.

5.1 General overview of the Si/ZnO pn heterojunction structure

Figure 5.1 shows the original configuration of the hybrid structure that is demonstrated.

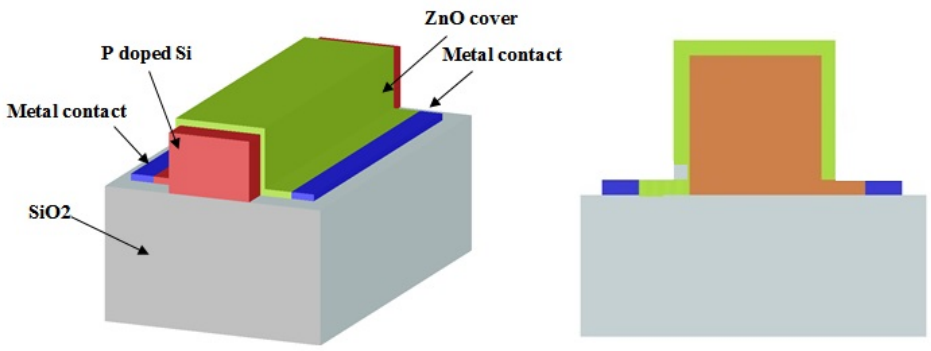

Figure 5.1: General overview of the proposed Si/ZnO rectangular heterojunction.

The two terminal configuration tends to be a part of optical integrated chip and designed as an all-wire component. It can be implemented in any specific part of the chip on extended photonic wire(s) and therefore delivers lower insertion loss. The strip like wire (waveguide) contains a p doped Si core with a naturally n-doped ZnO deposited all around. The motivation of ZnO deposition on all over the p-type Si core arises from a potential advantage; Wrapping the waveguide area with a low absorptive material may help to minimize the scattering loss caused by sidewall roughness of Si core. Using such a multilayer device simply provide a test platform to validate the new roughness model and investigate its effect on the overall loss performance of a device. It will show how the new modelling method can be successfully applied to a real and more advanced device.

The structure can be realized on top of SiO_2 wafer (SOI) as shown in Figure 5.1. The ZnO and p-doped Si pads on sides of the waveguide are used to create ohmic contacts with deposited aluminium, forming cathode and anode terminals respectively. The interface between metallic contacts and semiconductor region have to have higher doping concentration to reduce resistance between electrodes and semiconductor and optimum carrier transfers. In order to avoid additional insertion loss caused by metal contacts, the tails should be made long enough to isolate the metal contacts form the core waveguide. The heights of tails should be appropriately optimized. It should be narrow enough to retain both symmetry of the structure and wide enough for carriers to be easily transported into the anode terminal.

The interface between naturally n-doped ZnO and p-doped Si forms a heterostructure pn junction. Transparent ZnO cladding at proposed wavelength (1500 nm) where the attenuation is minimum for Si, helps to avoid any additional absorption loss. ZnO transparency feature can be maintained unchanged even at doping concentration of $4.8 \times 10^{19} cm^{-3}$ [99]. The doping concentration can even go beyond this level if AZO or Aluminium doped ZnO is used which slightly alters its transparency [100].

Since the structure tends to operate in depletion mode, only one carrier may be involved in the process of modulation and no recombination occurs offering a faster switching speed. The modulation is executed by applying zero and revered biased voltage intervals across the terminals. The carrier profile in p-doped Si can widely vary depending on the original doping concentration and reverse applied biased voltage. The wider the depletion inside the Si the higher real and imaginary refractive indices changes are induced.

When there is no bias voltage applied, the structure is assumed to be in OFF state and light interacts with holes in p-doped Si. At this condition, the company of uniformly distributed free carriers in Si and ZnO cover changes both refractive and absorptive behaviour of the wire. If the Si and ZnO are appropriately doped ($> 10^{17} cm^{-3}$), then the refractive (or absorptive) characteristic could be sufficiently affected for the propagating light to shift its original phase (or weaken its initial amplitude) when passes through the doped area.

At ON state, once the reverse biased is applied, depletion region at the interface of pn heterojunction begins to advance towards deep into the Si and ZnO. As a result the refractive and absorptive behaviour of the wire becomes more identical to undoped Si where the light can propagated with minimum interaction with free carriers. The more reverse biased voltage applied, the more free carriers deplete the wire area and hence allows the light to propagate with its intensity remains unaffected.

In order to retain the symmetry of the structure as a slab-like waveguide, the initial height of the pad connected to p-Si core was set to be as short as 50 nm. Under reverse biased conditions where the devices turns to ON state the depletion expands within the height of pad. This results a pinch to happen which obstructs the carrier free path and hence holes transportation will be affected.

The solution is to add 50 nm thin layer of SiO_2 on only top of the Si pad before the ZnO coating deposition on top of the p-doped Si core. This will isolate the ZnO and p-doped Si from each other and destroy the PN junction in that particular area and hence avoid pinch of in the region.

The physical, electrical and optical properties of the active material used in this work is summarized in Table A.1 at operating wavelength of 1550 nm [97, 98].

Operational wavelength has to be chosen so that all accompanied materials serve the lowest possible absorption. This is 1330 nm and 1500 nm for Si in which ZnO also has a transparent behaviour.

As explained in Chapter 2 the applicability of optical modulators can be limited if the the device doesn't serve a large modulation depth and high extinction ratio. One of the main factor which may limit the modulation depth is scattering loss. This is the initial motivation of choosing an optical modulator as a test platform to evaluate the

Material Properteis	Zinc Oxide	Silicon
Dielectric constant	8.5	11.7
Electron effective mass	0.23 m_e	1.18 m_e
Hole effective mass	0.54 m_e	0.8098 m_e
work function	4.5 eV	4.2 eV
Band gap	3.35 eV	1.11 eV
E lattice scattering mobility	200 $cm^2/V.S$	1471 $cm^2/V.S$
H lattice scattering mobility	180 $cm^2/V.S$	470.5 $cm^2/V.S$
E Velocity saturation	1×10^7 cm/s	1×10^7 cm/s
H Velocity saturation	1×10^7 cm/s	1×10^7 cm/s
E life time	8×10^{-9} s	3.3×10^{-6} s
H life time	8×10^{-9} s	4×10^{-6} s
refractive index	1.92	3.47
Absorption coefficient	0	0

Table 5.1: Si and ZnO electrical and optical properties

roughness model. In addition, the modulator designs that operate based on plasma dispersion effect can have a wide verity and be very complicated if the hybrid structure is considered. Due to this complication it is quite difficult or even impossible to estimate the roughness effect using experimental and theoretical approaches.

Furthermore, the surface roughness on nanoscale, leads to a higher contact resistance at the junctions, an important issue to high power and a reason to limit the high speed application of this device. Although the study of roughness effect on electrical characteristics of this device is beyond the scope of this thesis but is recommended as a potential future work.

This design is chosen to explain the several advantageous that the roughness model can offer. The model can be used to estimate the scattering loss in complicated multilayer devices where accessing the roughness data is almost impossible. The provided information on the overall loss can be also used to improve the device design further before it goes to the expensive fabrication processes and hence saves money and time. Using the information provided by applying this model can be used as a feedback data to optimise the fabrication processes regularly in order to achieve the most efficient and functional device. Furthermore, it can be used in the device characterisation process to analyse the overall performance of the device. For example it can be used to explain the amount of loss related to the roughness contribution and scattering loss and distinguish it form other sources of loss such as absorption and bending losses.

5.1.0.1 Optical response

The resulting change in refractive index due to the depletion of carrier at ON state causes almost 0.001 change in the effective index of the waveguide structure consequently. The

required length to have a π phase shift is worked out Using ($L = \lambda/2\Delta n_{eff}$), which is 775 μm.

Figure 5.2 explains the loss performance of the p-Si/n-ZnO structure, when the voltages is reversed from 0 to 07 V.

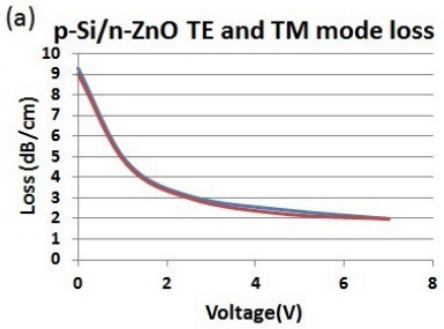

Figure 5.2: Loss performance for p-Si/n-ZnO TE mode as the reverse voltage varies from 0 to 7 V.

As TE and TM modes experience different degrees of optical confinement, the light has different powers at the waveguide core/cladding interfaces. Furthermore the propagation angles are different and therefore one mode will have more reflections at the interface than the other. Consequently the loss at the interfaces will be also different for TE and TM modes as can be seen in the Figure A.22.

Figure A.22 shows at zero biased voltage, when the device is in OFF state, the absorption loss is almost 9.3031 dB/cm. The loss is reduced by 7.33 dB/cm when the device turns to ON state by reversing the bias voltage to -7 V. Hence the change of carrier absorption loss from OFF to ON state is 7.33 dB/cm which is lower than the theoretically estimated absorption loss induced by $\approx 5 \times 10^{-17} cm^{-3}$ holes carrier change(8.86 dB/cm). Such a difference arises from the following issue: In theoretical calculations it has been assumed that the whole waveguide area is depleted from both electrons and holes. However in reality there are still free electron carriers at the interface of the pn heterojunction. This is shown in Figure A.23. Furthermore, there is an additional loss of 0.635 dB/cm related to the highly doped region under the aluminium contacts and free holes in the pad area.

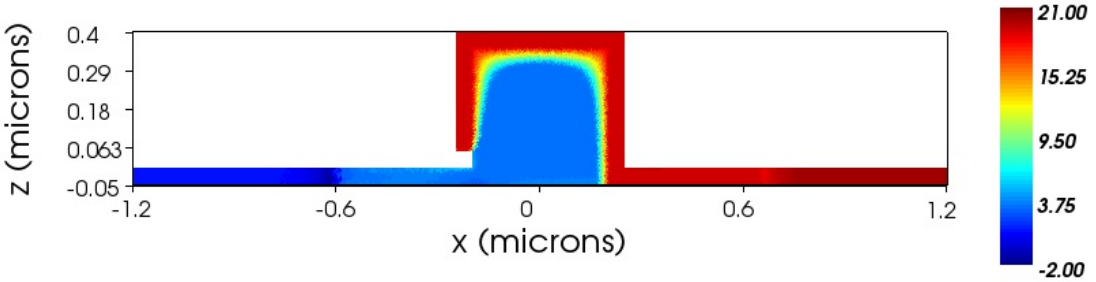

Figure 5.3: Electron carrier distribution effect at 7 V

Once the phase shifter is designed it can be implemented into the arm of MZI to convert the resulting phase change to the desired amplitude change. The MZI is designed to use MMI in order to efficiently couple the light into the arms and split it again. The detailed design of MZI and MMI coupler is explained in Appendix A. FDTD solution is applied to model the 3D MMI structure. The calculated effective index is 2.8 and having 6 μm width gives the length of 33.56 for the MMI region. The reflection into the fundamental mode of input waveguide is calculated to be 0.474 % (compared to the ideal detached waveguide). The transmission into the fundamental mode of output waveguide 1 and 2 are calculated to be 28.6245 % and 28.5923 % respectively. These two modes are then imported and propagate through the MZI arms. It can be also seen that the transmitted power into the output ports is reasonably balanced. The overall loss of the MMI region has been calculated as 0.8 dB which results in 83% of the input power transmitted to the output waveguides. Therefore the overall loss induced by coupler and splitter can be estimated as 1.6 dB.

Figure A.25 shows the scaled MZI structure under investigation. The input light is divided into two light beams propagating in the two arms of the MZI. The phase shifter is added in one arm. An identical structure can be also inserted into the other arm but must not be driven during modulation and therefore it will only balance the loss in both arms. To stay within the limit of the available memory space for simulating the structure, here the Mach Zehnder is only implemented into one arm.

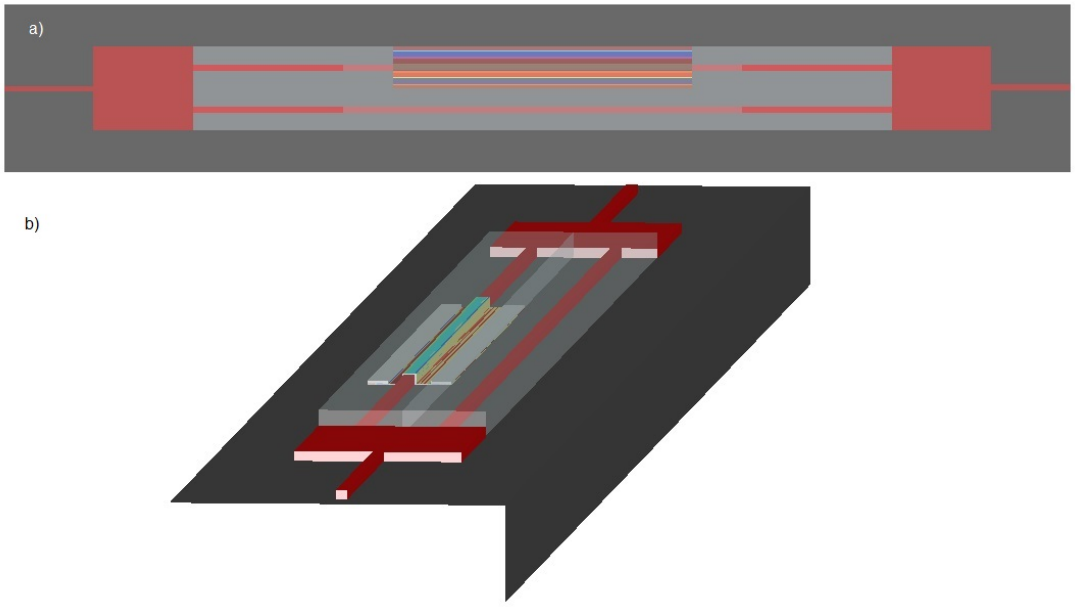

Figure 5.4: a)2D scaled MZI structure layout b) 3D scaled MZI structure. Note that the Phase shifter has been inserted into only one arm.

5.1.1 Modulation Response

Taking the calculated effective index into account the modulation response of the MZI structure can be approximated as shown in Figure A.26. Note that the transmission power ($(\frac{E}{E_0})^2$) is normalized to the range between 0 and 1.

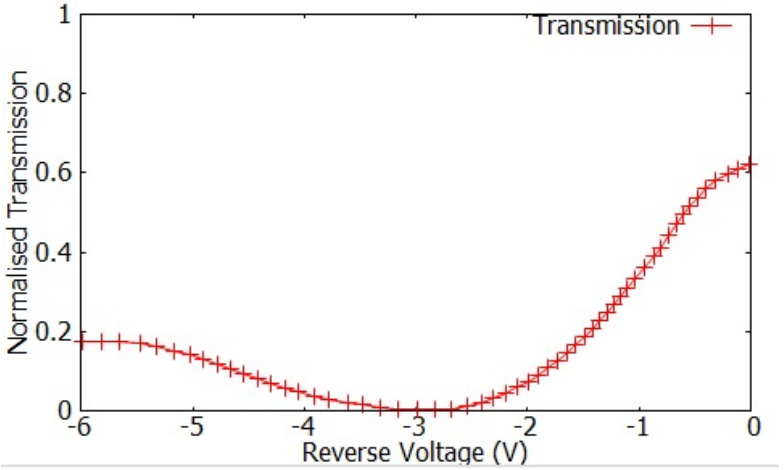

Figure 5.5: Modulation response of the Mach Zehnder.

Figure A.26 shows the MZI will have a π phase shift at 3 volts as expected when the length chosen to be 0.775 μm. The waveguide arms insertion loss and extinction ratio of the modulator is also shown in Figure A.26. This corresponds to 2.5 and 23.7 dB respectively. The insertion loss is defined as the amount of loss that the MZI have when no voltage is applied. The calculated insertion loss here include 1.6 dB of the MMI, and absorption loss caused by free carrier in off state. Furthermore the On chip insertion loss may be also taken into account which has to be added on top for practical cases. The on-chip loss associated with the entire MZI device. This on chip insertion losses in practical cases includes waveguide transmission losses caused by scattering, tapers and access waveguide losses, the losses induced by the splitter/coupler and finally the absorption losses due to free carrier plasma effect in the phase shifter. However, no tapered waveguide has been used within the MZI structure and hence tapers and access losses are not concern of this work at this stage and will be neglected. Among of other the scattering loss due to the roughness is considered for the phase shifter in the next section.

5.1.2 Scattering Loss Effects

By extracting the data from SILVACO and feeding the carrier profile changes into a mode solver in FDTD solution the transient time of the change of the effective index of optical mode can be obtained. This is shown in Figure 5.6 for TE mode where the DC simulation and characterization of optical rise time is illustrated . This is the response

to a square pulse with a voltage varying from 0 to -3 volts with a rise time of 10 ps. The voltage will stay at -3 V for 50 *ps* and decreases with 10 *ps* fall time to 0 volts.

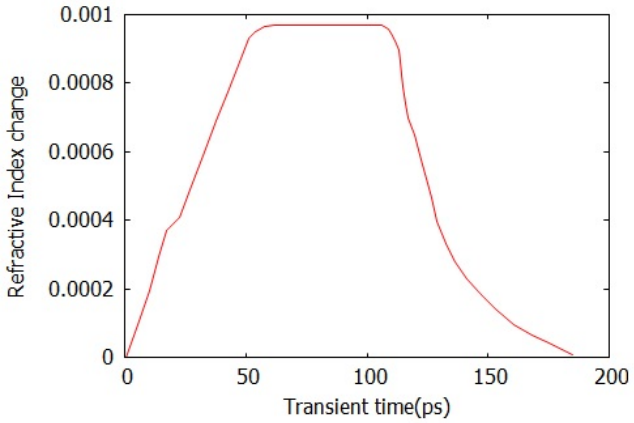

Figure 5.6: Optical effective index transient response

As can be seen the device shows a relatively slower switching speed (≈185 ps) compared to the currently reported fast switching modulators [47, 112]. The reason can be find by referring back to Figure A.9 where it can be seen the capacitance of the device is relatively high compared to these works. The additional capacitance is inserted by the top and left side p-Si/n-ZnO interface. However,the main motivation behind of this decision was to reduce the loss due the roughness which may be unavoidably generated during the etching process.

In order to demonstrate the roughness effect on the modulation performance, the generated roughness model in Chapter 4 has to be induced on the side walls of the Si waveguide at the interface with ZnO coating. Hence as well as the ideally smooth waveguide a number rough waveguides can be created with the RMS ranging from 5 to 12 nm and a fixed 100 nm correlation length that is large enough to have minimal effect on the loss behaviour.

Using the roughness model the scattering loss has been initially calculated for a passive device with different ZnO coating as expressed in Table 5.2.

-	0 nm thickness	50 nm thickness	100 nm thickness	150 nm thickness
5 nm RMS	9.4 dB/cm	5.1 dB/cm	4.8 dB/cm	4.6 dB/cm
7 nm RMS	13.1 dB/cm	7.3 dB/cm	7.1 dB/cm	7 dB/cm
12 nm RMS	24.1 dB/cm	16.7 dB/cm	16.2 dB/cm	15.9 dB/cm

Table 5.2: Scattering loss simulation for Si waveguide with different ZnO coating thickness for a passive condition

As can be seen substantially reduced the loss was introduced in to structure. Further increasing of ZnO coating will still reduce the amount of propagating loss. Even though

in this case the effect is not large enough to compensate the carrier absorption loss which is also a direct consequence of increasing the thickness of ZnO coating as discussed in previous section.

Once the refractive index data from Maltlab is imported to FDTD solutions it can be easily mapped from the 2D surface to a 3D surface that is the generated sidewalls roughness. Using the mode solver in FDTD solution shows that the waveguide absorption loss is not affected as the roughness introduced as expected. However the case is different for refractive index illustrated in Figure 5.7.

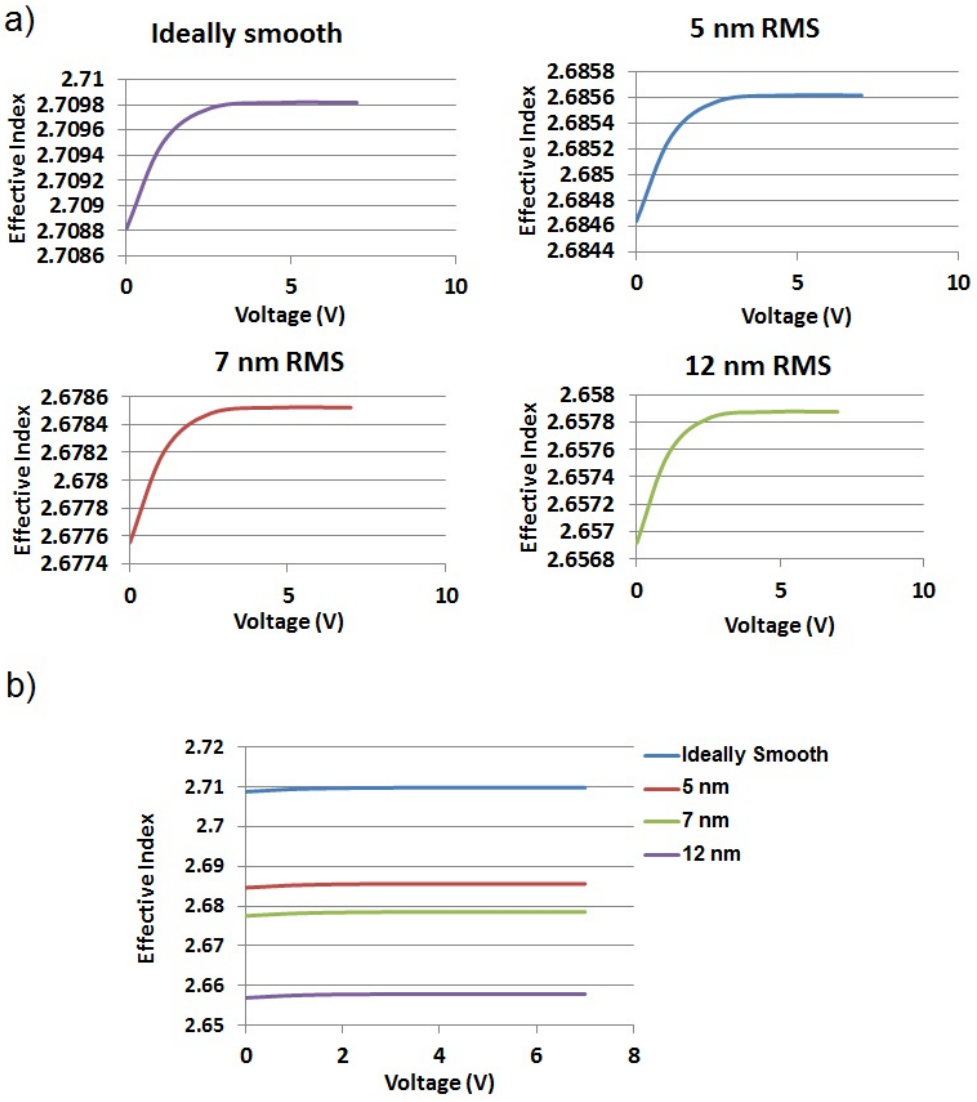

Figure 5.7: Effective index change for different sidewall roughness

The effective index is clearly varied for different roughness properties as it is shown in Figure 5.7b. However, as can be seen in Figure 5.7a, there is not a sensible change in the variation amount of effective index when the voltage is swept from 0 to 7 V. Although the effective index change in ON and OFF state is negligible, but the value of

$n_{eff1} - n_{eff2}$ differs from the case where the waveguide is ideally smooth as each arm of MZI has its own roughness characteristics. Since $n_{eff1} - n_{eff2}$ is different the transfer function also responds differently.

5.1.3 ZnO Coating Thickness Effect on The Modulation response

Figure 5.8 illustrates the condition when the transfer function is effected by sidewall roughness variations and ZnO coating thickness.

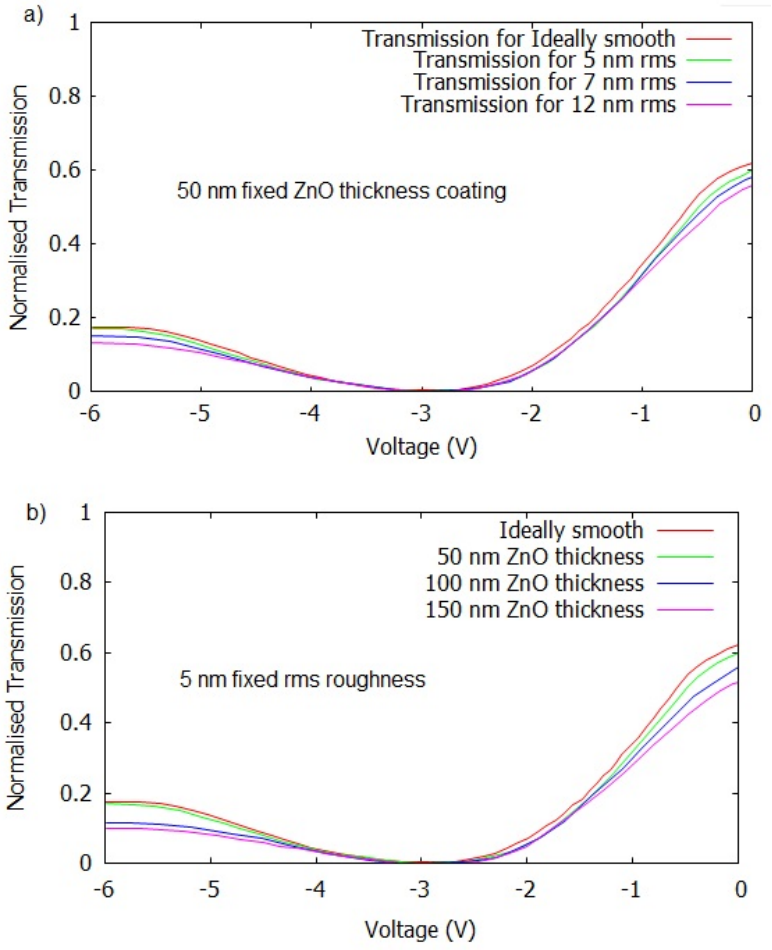

Figure 5.8: Modulation response to a) variation of rms roughness of the Si waveguide. b) variation of the coating ZnO

As shown in Figure 5.8a as the RMS of the sidewall roughness increase by 7 nm, the insertion loss also increases by 0.2 dB and extinction ratio is slightly reduced due to the additional losses. This explains how the roughness can limit the extinction ratio and hence the applicability of this device. It also agrees with theoretical expectations. The sidewall roughness in one arm of MZI destroy the balance between signals power propagating in two arms of MZI. As a result when the imbalanced and out of phase light

waves are combined a full destructive superposition will not be achieved resulting to a lower modulation depth.

However when the ZnO coating starts increasing it counterbalance the effect of roughness by almost 50 %. Hence the device experience a lower scattering loss. In contrast as explained in the last section, when the ZnO thickness increases the overall absorption loss caused by free carriers also increases. Therefore, although there is a slight improvement in the loss performance of the modulator, increasing the ZnO can not very effectively solve the scattering loss effect in the modulators.

5.2 Summary and Conclusion

In this chapter a the performance of a potential optical modulator design is investigated against the roughness effect. The roughness model has been successfully implemented to the structure to study the scattering in the modulator and its influence in the performance against changing roughness parameters and ZnO coating thickness. The results shows that the extinction ratio of the modulator can be limited by almost o.2 dB if the roughness RMS changes from 5 to 12 nm. This information can be used in order to preciosity chose appropriate processing steps to achieve a low scattering loss operation.

Furthermore, it shows that increasing the ZnO coating width can not significantly improve the extinction ratio of the design. Hence it helps to bypass the need for fabrication of several samples and test steps leading to save money and time during device design cycle.

Chapter 6

Summary, Conclusions and Future Works

The aim of this chapter is to discuss the key findings of this research in a summary; the conclusions that may be traced from them and an extension to further works.

6.1 Summary

This project has been constructively useful to the author by extending the knowledge of scattering loss effects using the novel method introduced in Chapter 3. This project has been followed in two major parts

In this part, the roughness model have been developed based on an inspiration from the technique that is excessively used in MRI field. A randomly filled K-space is generated by a general random generator function. The resulting k-space will then be filtered based on the Autocorrelation function that is mainly defined by two key components; rms and correlation length. The filtered k-space will then be transferred back to real space using inverse Fourier transform module.

The results have shown a good agreement with experimental results obtained from two different samples having a size of $1000 \times 500nm$, $1200 \times 500nm$, $1400 \times 500nm$, at 3.47 μm and $330 \times 220nm$ at 3.47 μm wavelength. In addition the results obtained using this model are compared with results by Payne and Lacey model and several other experimental published works in which all were in a good agreement.

In the second part in optical modulator is designed where later used as a case study for the roughness model. The modulator is designed to operate based on carrier plasma effect where the carrier concentration change can result a change in refractive index and absorption coefficient. It has been shown that the device can work in depletion mode

and be used a phase shifter when an appropriately large enough change in the carrier concentration causes the refractive index to vary. The device shows a refractive index change of 0.0001 in 0.775 μm length with less than 4 dB insertion loss. The transient response is yet a remaining issue in comparison with the current fast switching speed modulators that can be considered for further optimisation.

The initiative of this project leads to a parallel fabrication work of the device which is still in the progress by T. B. Masoud. The device is slightly modified as the SiO_2 could not be perfectly aligned with the waveguide sidewall. The devise has promisingly shown a switching in injection mode and forward biased conditions.

Scattering Loss Estimation Using 2-D Fourier Analysis and Modelling of Sidewall Roughness on Optical Waveguides [90].

The research in this project has led to produce 6 Publications as follows.

- Scattering loss estimation using 2D Fourier analysis and modelling of sidewall roughness on optical waveguides [90].
- Submicron silicon waveguides and optical splitters for mid-infrared applications [113].
- Silicon electro-optic switch based on n-ZnO/p-Si heterojunction structure [114].
- Compact Fabry-Perot electro-optic switch based on n-ZnO/p-Si heterojunction structure [115].
- Silicon waveguides and devices for the mid-infrared [116].
- Broadband plasmonic couplers for light trapping and waveguiding [117].

Also the following article is submitted for a publication pending for decisions in Optics Letter and IEEE photonics Journal respectively.

- Hot-Wire Polysilicon Waveguides with Low Deposition Temperature.
- Computational Studies of Si/ZnO Depletion Mode Heterojunction Optical Modulator and The Effect of sidewall Roughness.

6.2 Conclusions

The involvedness of roughness effect and its accrue estimation on the loss performance in modeling any optical device from waveguides to more complex hybrid structures can

be very adventurous in a number of ways. Among of all the very two important ones are as follows:

It can provide a more realistic characteristic of the proposed structure before the device enters to the fabrication process that may need substantial funds to spend and as well as more time.

The scattering loss in many devices can be trade off against the performance merit a particular device. Proving accurate predictions of its effect will therefore can help one to precisely adjust these merits to serve a certain applications

6.3 Future Works

The following is suggested as the highlighted plan to proceed for the future:

- The structure design of the initial switch has to be optimized to operate in single mode conditions; in addition the geometry has to be preciously modified to provide birefringence free operation. Some of these works have been partially done but not included as they were yet incomplete. Further engineering in geometrics aspects are also necessary in order to make the device feasible for fabrication.
- The effect of sidewall roughness on electrical and optical behaviour of the device, in terms of carrier profiles, transient response and dispersion characteristics are still remained as part of the this project future works.
- The fabrication and characterization of the device to compare loss measurement results with predicted results provided by the model.
- To investigate the effect of sidewall roughness on the eye diagram of the proposed optimized device that can provide the overall bit rate of the modulator and the maximum bandwidth it can offer.
- Sidewall roughness can be implemented in wider range of structures and geometries. One potential possibility is its effect in photonic crystal structures that is very little investigated

Appendices

Appendix A

The Design and Development of Hybrid Heterojunction Optical Modulator

A.1 The Design Strategy

The design and simulation strategy of this work is based on the flow chart diagram in Figure A.1.

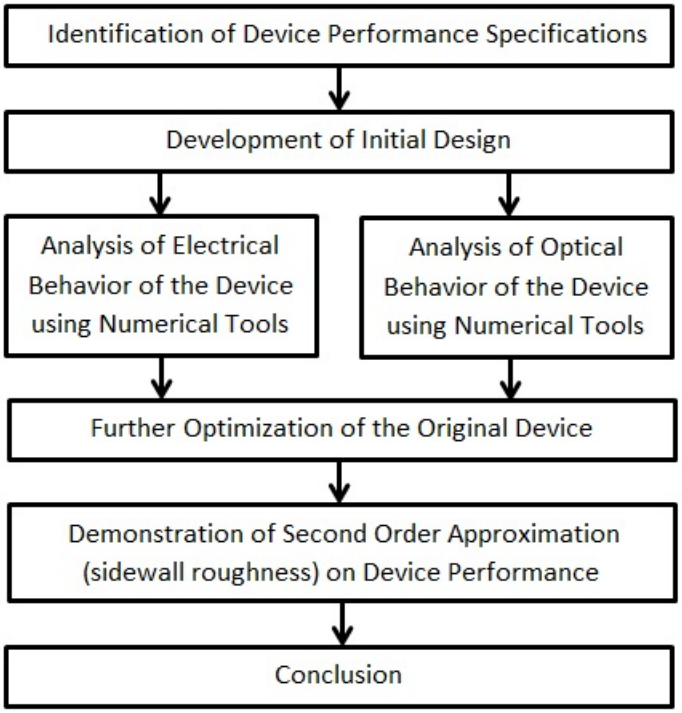

Figure A.1: Design Strategy flow chart diagram

There are several preferred priorities of the performance merits which have been considered in the preliminary design of this work. The electrical and optical behavior of the proposed device in ON and OFF state will be explored using numerical simulation tools. By analysing these primary results, further optimizations are carried out in order to boost device performance. Correspondingly, further details (in particular side wall roughness) which has been independently studied and modelled is implemented in to the structure to demonstrate its possible effects on the performance of the device. Assuming to have promising results, the device can be considered for fabrication and final characterization.

In this work ATLAS by SILVACO is used for electronic demonstration together with FDTD solutions by Lumerical for optical modelling. Matlab is also used to import the required data from ATLAS to FDTD solutions as will be explained later.

A.2 Performance Merits

In Chapter 3 the tradeoff between the performance merits in designing a structure has been explained. These performance specifications have to be prioritized based on their relevant importance in order to serve the finest possible functionality and high efficiency for a particular purpose. For example, if a low loss operation is important in a design of a plasma dispersion modulator, the parameters has to be chosen in the most appropriate way which optimally lead to a low loss performance, even though the cross section size may become larger. In this work these merits are lined up based on their priority in the following order:

1. **Loss:** A low loss performance can be only achieved by minimizing absorption and scattering loss of the device. As one the main inspirations of this work, the loss origins has to be deduced into the design for quantitative determination of their effectiveness in device performance. More importantly the structure has to be designed in such an appropriate way so that one can demonstrate the influence of possible solutions which can eliminates or at least reduce the consequence of these inserted losses.

2. **Modulation depth:** The depth of the modulation is increased as a direct result of the reduced loss in a modulator.

3. **Foot print (Effective length):** Long effective length itself can result additional loss in a modulator as the propagating light interacts longer with boundaries imperfections. It is therefore important to keep the effective length as short as possible.

4. **Modulation speed:** The modulation speed has to be feasible, practical and comparable with the many current reported modulators in the literature ($> 10GHz$),

although the very fast operation is not intended as the key purpose of this work. One clear solution that provides faster operation is to shrink the cross section size. As the cross section size is reduced more and more the loss starts to increase, since a larger portion of propagating mode interacts with the irregularities of the waveguide boundaries. This tradeoff is the major drawback in many other designs of optical devices based on wire waveguide approach.

A.3 Principle Design Consideration

In this section based on the above performance specification the principles of the purposed design is introduced and discussed:

A.3.1 Plasma Dispersion Effect Phase Modulation

It is intended to apply plasma dispersion effect as the most efficient approach for optical modulation in Si based configuration. To gain relatively high switching speed, the device is decided to operate in depletion mode. Aiming to obtain direct amplitude modulation through changing absorption coefficient in Si needs relatively large carrier distribution variations though. Because otherwise the modulation length becomes unavoidably long resulting further propagating loss. Phase modulation can be achieved within a shorter length and hence it is generally more preferred. The phase modulation can then be transformed to amplitude modulation by implementing in a MZI.

A.3.1.1 Hybrid structure

It has been explained that hybrid Si material integration can extend the capability of Si or compensate the Si optical weaknesses in general. It is therefore decided to include the advantage of such a hybrid design in the purposed structure.

A.3.1.2 Two terminal device

As mention in Chapter 2 two terminal devices may have lower absorptive behavior. Two terminal configuration is therefore chosen in the initial design.

A.4 General Waveguide Selection Criteria

In order to figure out a highly optimized optical waveguide based design, it is important to precisely select the waveguide criteria and specifications. For example, all losses in

the waveguide correspond to the material and geometry of the waveguide. For a low loss and efficient operation it is therefore essential to realize and select right waveguide materials and accurate geometry. There are three fundamental design principles which one should emphasise when selecting the material and geometry of the waveguide; Size, performance and fabrication.

A.4.1 Size Parameters

A.4.1.1 Refractive Index Contrast

The refractive index contrast or delta Δn is the difference between the core and the cladding of a dielectric waveguide. It is normally used to expresses how the quality of the waveguide boundaries can affect the confinement of the electromagnetic wave. High Δn results in very rigid boundaries and hence confining the wave strongly in the waveguide. Low Δn results in the significant part of the light to propagate outside of the waveguide boundaries. The confinement quality can be explained using confinement factor.

The confinement factor is defined as the ratio of the optical power propagating inside of the waveguide to the total propagating power [53]. The confinement factor is directly related to extinction coefficient (the imaginary part of refractive index) which describes the decay rate of the evanescent field from the core/cladding interface into the cladding [53]. Therefore confinement factors scales monotonically with Δn.

A.4.1.2 Single mode cutoff dimension

The largest waveguide dimension in which only one mode (fundamental mode) exists in the dielectric waveguide is called single mode cutoff. Single mode condition is varried for different structures and geometries. The number of supported modes in dielectric slab waveguide is worked out from [70]:

$$M = \frac{2d}{\lambda_0}\sqrt{(n_0^2 - n_1^2)} \tag{A.1}$$

Which λ is the free space wavelength, d is the waveguide width, n_0 and n_1 are the refractive index of cladding and core respectively. By setting $M = 1$ the cutoff can be calculated as :

$$d_{cutoff} = \frac{\lambda_0}{2\sqrt{(n_0^2 - n_1^2)}} \tag{A.2}$$

Since $n_0^2 - n_1^2 = (n_0 + n_1)(\Delta n)$, High Δn leads to large value of $n_0^2 - n_1^2$ and hence to a smaller single mode cutoff dimension.

For rib and strip waveguides the cross section dimensions determine the number of modes that can travel through the waveguide. This was firstly demonstrated by Soref. et. al [93] for large waveguides.

For sub-micron applications the single mode conditions needs the use of complex computational methods such as the finite difference methods, the beam propagation method and the film mode matching method [94]. Recently full vector method is used to determine the single mode condition. Applying a fitting procedure, the authors formulated an equation in which the single mode condition for both polarizations (TE and TM) at 1.55 μm operating wavelength can be worked out [94];

$$\frac{w}{h} \prec \frac{55.679h_c^2 - 45.46h_c + 10.73}{\sqrt{1 - r^2}} + 21.805h_c^2 - 21.47h_c + 5.69 \tag{A.3}$$

$$\frac{w}{h} \succ \frac{-14.9h_c^2 - 8.08h_c + 0.58}{\sqrt{1 - r^2}} + 22.28h_c^2 - 19.31h_c + 4.16 \tag{A.4}$$

where $r = h_s/h_c \prec 0.5$, w is the width of the waveguide, h_c is the height and h_s is the pads height.

A.4.1.3 Dispersion

As the loss can significantly increase in nanoscale region the importance of precise selection of cross section size is considerably enhanced. The smallest size for a waveguide in which at least one propagating mode can be supported is therefore an issue to consider when one selects a suitable waveguide dimension in nano regime. One way to obtain a minimum possible waveguide cross section size and also single cutoff dimension is through defining a parameter that is called Dispersion. The dependence of propagation constant on the frequency of propagation is called dispersion. The propagation constant is defined as:

$$\beta = n_{eff}\frac{2\pi}{\lambda_0} \tag{A.5}$$

The relationship between the normalized frequency and normalized propagation constant is referred as dispersion equation. The normalised frequency is a function of waveguide dimension, wavelength and Numerical aperture according to the following relation:

$$\nu = Kd\sqrt{(n_0^2 - n_1^2)} = \frac{2\pi d}{\lambda_0}\sqrt{(n_0^2 - n_1^2)} \tag{A.6}$$

Since there is a specific propagation constant corresponds to each propagating mode, the dispersion equation can be graphically represented as a set of universal curves. Each curve is related to a specific mode order that only exists for particular values of normalized frequency which is larger than a certain amount. This amount however varies for different modes. There is however an exception for the fundamental mode which a value of ν exists for all values of normalized propagation constant. Hence except fundamental mode, other higher order modes have cut off point at specific normalized frequency.

The normalised frequency is a function of waveguide dimension. Therefore according to dispersion relation, the variation in propagation constant can be related to the variation in dimension of the specific waveguide at specific wavelength. Based on this argument one can replace the normalised frequency by waveguide dimension and generate dispersion curves for specific waveguides.

A.4.2 Performance Parameters

A.4.2.1 Bandgap

As expressed before, the optical bandgap of the waveguide material indicates the absorption edge. If the optical energy is greater than the bandgap significant absorption occurs while optical energy lower than the bandgap shows minimum absorption. The optical bandgap determines which wavelength is the most suitable one to be used in the waveguide. Therefore for a given specific operational wavelength to be carried in the waveguide, material selection depends on optical bandgap and vice versa.

A.4.2.2 Wavelength

The right range of the wavelengths in which the waveguide shows low band edge absorption and scattering loss is another factor to consider when selecting a material. Silicon exhibit the lowest absorption behavior at around 1.3 μm and 1.55 μm wavelengths.

A.4.3 Fabrication Parameters

The materials which the waveguide is constructed from have to be compatible with current state of the art processes, CMOS. It is generally accepted that silicon on insulator (SOI) is the best platform to be used in the fabrication of silicon photonic devices.

SOI wafers can provide significant advantageous because of their multilayer semiconductor/dielectric structures. Multilayer material systems with high refractive index contrast can be realized using SOI process. Hence hybrid integration can be achieved more easily. The composition of buried oxide and air with silicon, with respectively 1.44, 1 and 3.47 refractive indices leads to highly confined layers with high aspect ratio contrast within the wavelength of interest.

Considering all above principles one can choose the right combination of geometry and materials for low loss operation for the proposed modulator design.

A.5 Waveguide Material and Geometry

A.5.1 The Urgency of Silicon in Waveguide Structures

Silicon, as an inexpensive dominant semiconductor in electronics, provides the low cost integration with electronics. Furthermore due to its unique compatibility, silicon is being the leading candidate for optical interconnect. It can provide performance enhancements resulting from electronic/photonic integration, and compatibility with CMOS as the most successful technology for producing electronics. These proven abilities together with the trend towards nanoscales integration for high speed operation and more cost effective fabrication process have opened a wide research area. The desirability is basically due to the unique combination of low fabrication costs of silicon, performance enhancements resulting from electronic/photonic integration, and compatibility with the worlds most successful electronics technology CMOS. Si in certain wavelength ranges (1350-1550 nm) also shows a low absorption characteristic.

However, Si is inherently restricted as an optical material because its indirect band gap gives inefficient light emission characteristics, the centrosymmetric crystal structure of Si and poor electro-optical properties (Kerr and Pockels effect). Hence in order to maintain the Si fundamental advantage, the in silicon must be carried out in other ways. One possible solution that is considered within this work is to investigate potential improvements of composite structures where Si and other compatible semiconductors create a hybrid Si based configuration. The benefits of such a design arise from using low cost electro-optical compound semiconductors that are appropriately compatible with silicon photonic integration and technologies.

Together with Si as a well-studied, cheap, widely used material in electrical and optical devices, the accompaniment of ZnO and SiO2 can have plenty to offer electronically and optically. The potential properties of ZnO have been also studied to investigate their use in waveguide based devices [31, 33]. Its transparency behaviour at relatively wide wavelength (700 nm to 1600 nm) leads to extend its applicability in optical and more recently in electro-optical structures. Both ZnO and SiO2 have shown that are gratefully

compatible with Si technology. The integration of both materials with Si can provide high contrast refractive index structures and strongly confined waveguides. While ZnO is naturally n-doped and hence can contribute efficiently in free carrier distribution in a heterojunction diode, it is transparent at NIR region. Therefore there would be no extra free carrier absorption added which is a very promising advantage.

The integration of ZnO in Si based devices is still very rarely referred in the literature. The promise of hybrid Si/ZnO pn heterojunction structure as UV LED and photodetectors have been demonstrated [95]. But for a similar device to operate as depletion mode optical switch is still little explored. The nearest proposed device is an electro-absorptive Si-metal structure operating based on Schottky diode configuration and shown to exhibit efficient free carrier depletion theoretically [96]. The device has short modulation length at the expense of very high insertion loss caused by the metal contact on top of it. Further studies are still required for polarization dependence, propagation loss and free carrier distribution profile in submicron scale waveguide, otherwise known as photonic wire. In spite of that it is important for the heterojunction design to be practically feasible for fabrication and characterization.

Hence Silicon CMOS process compatibility, high Δn and low absorption behaviour at 1500 nm wavelength can make the combination of SOI, SiO2 and ZnO an attractive material system for such a structure and many other low loss waveguide applications.

The physical, electrical and optical properties of the active material used in this work is summarized in Table A.1 at operating wavelength of 1550 nm [97, 98].

Material Properteis	Zinc Oxide	Silicon
Dielectric constant	8.5	11.7
Electron effective mass	0.23 m_e	1.18 m_e
Hole effective mass	0.54 m_e	0.8098 m_e
work function	4.5 eV	4.2 eV
Band gap	3.35 eV	1.11 eV
E lattice scattering mobility	200 $cm^2/V.S$	1471 $cm^2/V.S$
H lattice scattering mobility	180 $cm^2/V.S$	470.5 $cm^2/V.S$
E Velocity saturation	1×10^7 cm/s	1×10^7 cm/s
H Velocity saturation	1×10^7 cm/s	1×10^7 cm/s
E life time	8×10^{-9} s	3.3×10^{-6} s
H life time	8×10^{-9} s	4×10^{-6} s
refractive index	1.92	3.47
Absorption coefficient	0	0

Table A.1: Si and ZnO electrical and optical properties

Operational wavelength has to be chosen so that all accompanied materials serve the lowest possible absorption. This is 1330 nm and 1500 nm for Si in which ZnO also has a transparent behaviour.

A.5.2 Waveguide Geometry

All wire optical devices can provide lower insertion loss, making it suitable for high compact optical integration. SOI wafers can serve a natural one dimensional wave guiding material. However the applicability of planar waveguide is limited for particular uses and therefore two-dimensional confinements is required. Three possible structures can be considered in the design of the proposed hybrid modulator which can also serve low loss operation in a 2D confined guiding configuration:

1. Cylindrical nanowires: Cylindrical waveguides in nanoscale such as photonic nanowires exhibit a better confinement compared to the other geometries. In addition they are polarization independent and hence can support hybrid modes which is a unique advantage. Cylindrical nanowires tend to be fabricated horizontally. Although recently reported studies show that cylindrical waveguides can be realised quite smooth, the fabrication of cylindrical nanowire horizontally on silicon substrate is a very challenging process.

2. Strip Waveguide: Nanosize rectangular slab waveguides are one of the most attractive geometries that can confine the light in 2 dimensions. the third dimension of the structure provide the propagation path. They can be relatively easy fabricated as SOI structures using etching process. Either air or materials with refractive index lower than that of the Si can be used as a cladding of the etched waveguide. Nanosize rectangular slab waveguides are one of the most attractive geometries that can confine the light in 2 dimensions. They can be relatively easy fabricated as SOI structures using etching process.

3. Ridge or rib Waveguides: Likewise ridge waveguides can confine light in two dimensions, but has weaker horizontal confinement than a strip waveguide even with the same core and cladding references. In other word the effective index difference in the horizontal direction is higher for a strip waveguides. In comparison with strip waveguides, the center of ridge and rib structures can be better isolated from the metallic contacts and therefore provides a significantly reduced absorption loss.

The well developed fabrication process of strip, strip-like and ridge waveguides make them the dominant geometries for most low loss silicon waveguide based devices. It is therefore intended to use Strip-like waveguide (rib) where all wire devices with lower insertion loss can be realised, making it suitable for high compact optical integration.

The propagating mode properties are directly related to the device geometry and the materials. The propagating mode properties are discussed in section A.9.1 by investigating dispersion characteristics and polarization features of the proposed modulator design.

It is also important to note that in this work, it is aimed to initially begin with relatively nanoscale large waveguide cross section that can operate in multimode condition. Once the original structure is developed, the single mode condition can be considered by investigating the effect of size reduction against any loss increase. In the literature there has been a little interpretation of such an effect reported. The configuration can be therefore further optimized to achieve the most functional cross section size that can provide low loss as well as acceptable modulation speed.

A.6 Device Modelling Concept

Figure A.2 shows the modelling concept used to model to proposed modulator. Following this this simple results a time efficient way to model the structure.

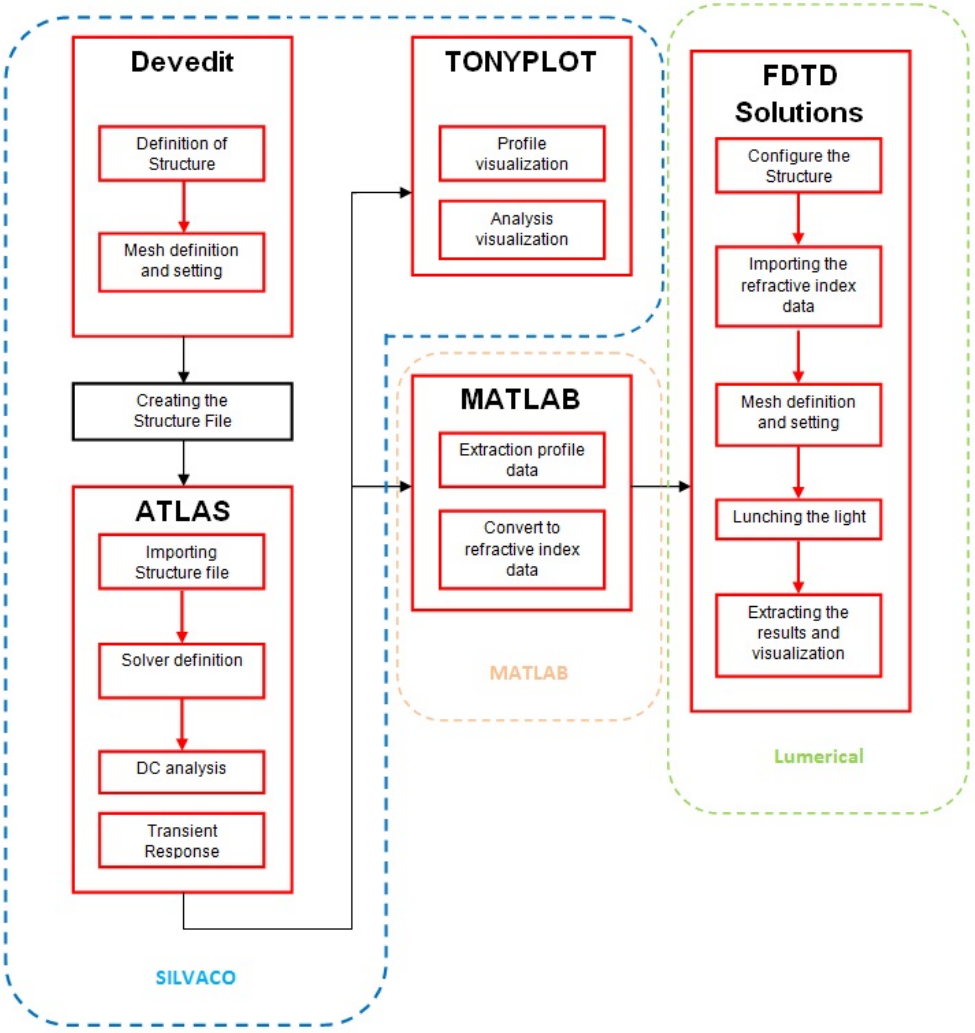

Figure A.2: Flow chart of modelling procedure of the device

SILVACO is a multipurpose semiconductor design and simulation tool. This suit consists of a number of modules that perform a certain task to design model and analyse the results.

1. Devedit is one of the three user graphic interface to define the initial structure in 2D or 3D. Materials, doping and electrodes are then specified. Once the structure is defined a mesh can be applied with mesh generation and refinement module. SILVACO uses triangular grid system and is important to define a very fine mesh in abrupt changes in which is likely to have a relatively large parameter variations; For example, at the interface of the boundaries, and regions of different doping concentration or type. Once the structure is completely created, it is saved into a structure file and can be imported to the next user graphic interface where ATLAS commands are executed.

2. Deckbuild is a runtime environment for executing ATLAS commands. ATLAS is a device simulator that can predict the device electrical behaviour under different bias conditions based on the physics of the device. ATLAS has been previously used in order to analyse electro-optic modulators in SOI waveguides in other works [101–103]. The software allows a complete statistical approach (Fermi-Dirac statistics) when heavily doped regions are considered [101]. ATLAS numerically approximates the operation of device onto a 2 or 3 dimension grid system which is previously created in Devedit. Applying a set of differential equation into this system one can simulate the transport characteristics of carriers through the entire structure. These equations are Poissons equations relating the variation in electrostatic potential to local charge densities, the charge continuity equation for hole and electrons and the transport equation which describes the way that electrons and holes densities are altered as a result of transport processes, generation and combination processes. ATLAS can be initiated in ATLAS by calling it in Deckbuilt. Additional material properties information can also be defined in ATLAS itself if required. Then depending on the imported device, ATLAS enables the user to select specific physical modes. These models are generally classified into 5 groups that are carrier statistics, mobility, recombination, impact ionization and tunnelling. The wide variety of available model is to ensure the relatively realistic results of the electrical performance of the device can be obtained. The models used in these projects are those standard SOI substrates models provided in ATLAS and are extensively used for the simulation of the pn junction relative active devices. These are carrier statistics, mobility, carrier generation and recombination and band gap narrowing which is used here. Surface recombination model is also used to take the semiconductor/insulator interface recombination into account. The numerical solving method has to be chosen after to solve these models. Here Newton and Gummal method is chosen for DC and transient simulations. Finally the user states what data need to be extracted from the simulations. DC ,

AC and transient simulation results can be obtained. Depending on the specified needed data, the extracted information will be then exported to specific files for further analysis and or displaying in TONYPLOT module.

3. TONYPLOT: TONYPLOT is a visualization module in which every solution computed by ATLAS can be analysed and displayed. This tool is very useful in particular for fast development of the initial structure. Once the device is electronically investigated the data has to be transfer to MATLAB and Lumerical FDTD solutions for optical modelling analysis.

A.7 Simple pn Heterojunction diode

The investigation was begun by considering the electrical functionality of simplified p-Si/n-ZnO heterojunction in reverse biased situation. Although the device operates in depletion mode, the carrier lifetime time has been chosen to be quite large in order to show that is not playing a dominant role in the speed of the device when it operates in depletion mode. The junction is assumed to be abrupt for simplicity [102]. Figure A.3 shows the schematics of a simplified pn heterojunction structure.

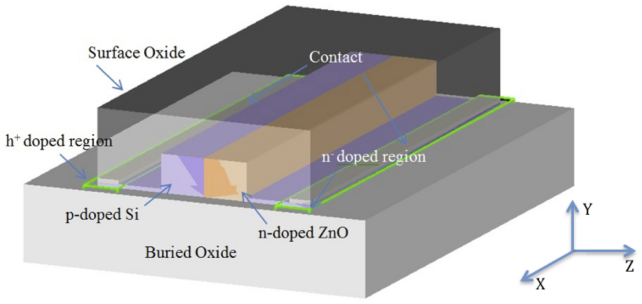

Figure A.3: Layout of a general configuration of Si/ZnO heterojunction

The waveguide structure contains a simple 0.4 μm by 0.4 μm horizontal pn heterojunction and is configured on thick buried oxide layer. Two 50 nm wide pads are used on either side of the waveguide to connect the heterojunction diode to the highly doped regions. The Aluminium contacts are deposited on top of the highly doped region. The overall structure is covered by a thick layer of surface oxide.

A.7.1 Band Diagram

Figure A.4 shows the band gap diagram of such a configuration at different bias conditions. Both Si and ZnO have $1 \times 10^{18} cm^{-3}$ primary dopant densities. The values are chosen to be reasonably practical, especially for ZnO. These values however will be

optimized in later stages to achieve the most efficient functionality. In addition, it is assumed that the temperature is fixed at 300 K.

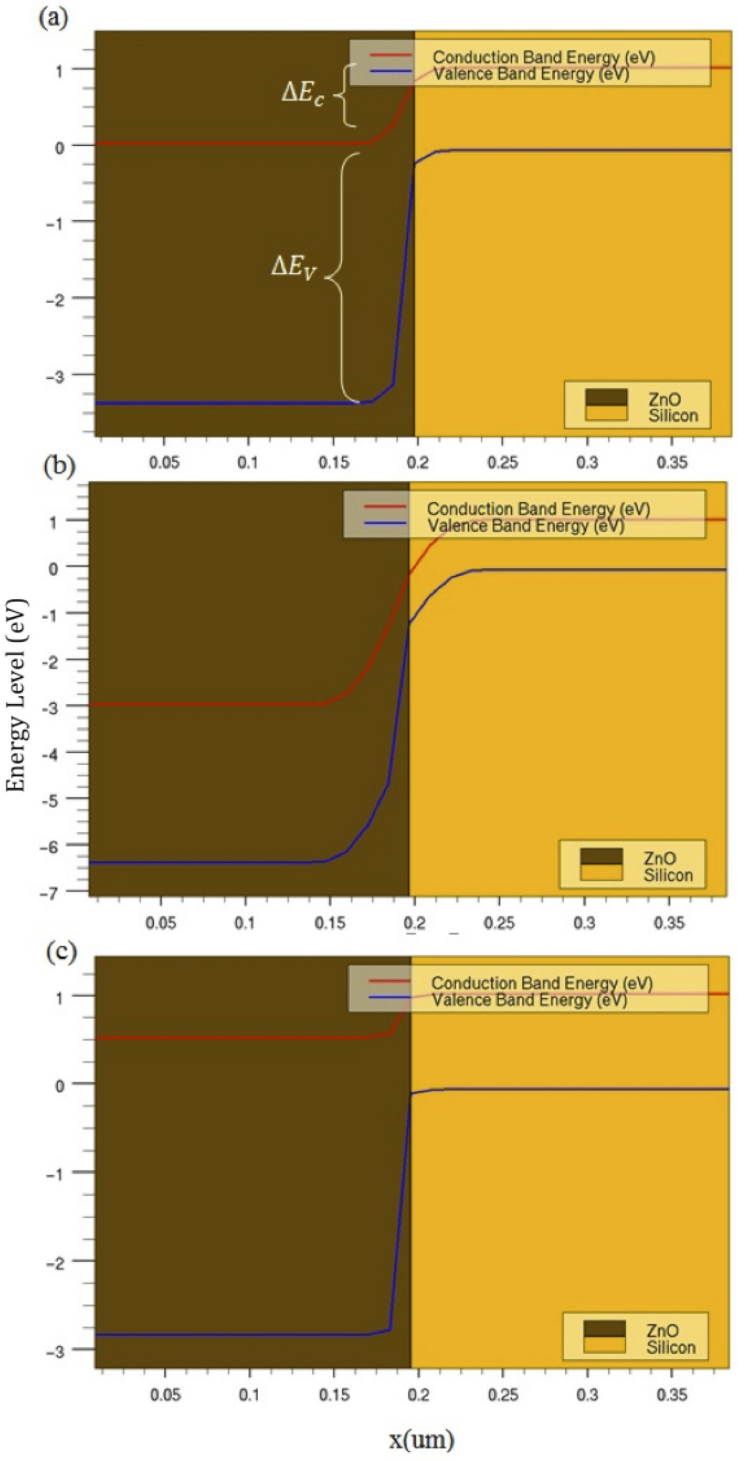

Figure A.4: Heterojunction bandgap diagram for a) zero biased b) -3 V reversed biased and c) 0.5 V forward biased conditions

Once the P-Si/n-ZnO heterojunction is configured carriers diffuse from either sides due

to a concentration gradient at the junction. When electrons diffuse from n-ZnO to p-Si, the positive space charge will remain in n-ZnO near the junction. This increases the conduction band energy and the band bends upward in n-ZnO. However, the conduction band energy in p-Si increases and thus the band bends downwards. From the electron affinity of Si and ZnO, the conduction band offset is

$$\Delta E_c = x_{ZnO} - x_{Si} = 3.7 - 4.29 = -0.59ev \tag{A.7}$$

The valence band offset is

$$\Delta E_v = Eg_{ZnO} - Eg_{Si} + \Delta E_c = 3.35 - 1.12 + 0.59 = 2.82ev \tag{A.8}$$

The large difference between ΔE_v and ΔE_c indicates that the barrier is much lower for electrons than for holes. Hence, the potential barrier is lowered and then electrons can be easily injected into the p-Si side compared to holes into the n-ZnO. This situation is likely considerable in forward biased where the heterostructure can also experience luminescence from electron-hole pair recombination [1]. This will limit the switching speed of the junction, so the forward biased operation is not the mode of this work's interest.

A.7.2 I-V characteristics

The I-V characteristic of the above p-Si/n-ZnO structure under forward biased conditions has been obtained and shown in Figure A.5. As it can be clearly observed the structure confirms a diode behaviour with the built-in potential of about 0.9 V. For the same conditions, many nearby and same values ranging from 0.8 to 1.1 V are experimentally calculated and referred widely in the literature [104]. In addition the breakdown voltage of similar structures that are indicated in the literature are relatively high mainly due to high resistivity of the p-type layer [105]. Hence there is no critical limitation on applying reverse biased voltage if the structure is reversed biased with voltages as high as 15 V. This result may change if the p-Si is heavily doped, since it makes the tunnelling current more likely to happen. Unfortunately in the situation where the structure is reversed biased the software limitations do not allow for a well interpretation of breakdown conditions. However in accordance to various experimental results reported in the literature, only small reverse leakage current is expected when the reverse biased voltage reaches 20 V in the p-Si/n-ZnO heterojunction [105].

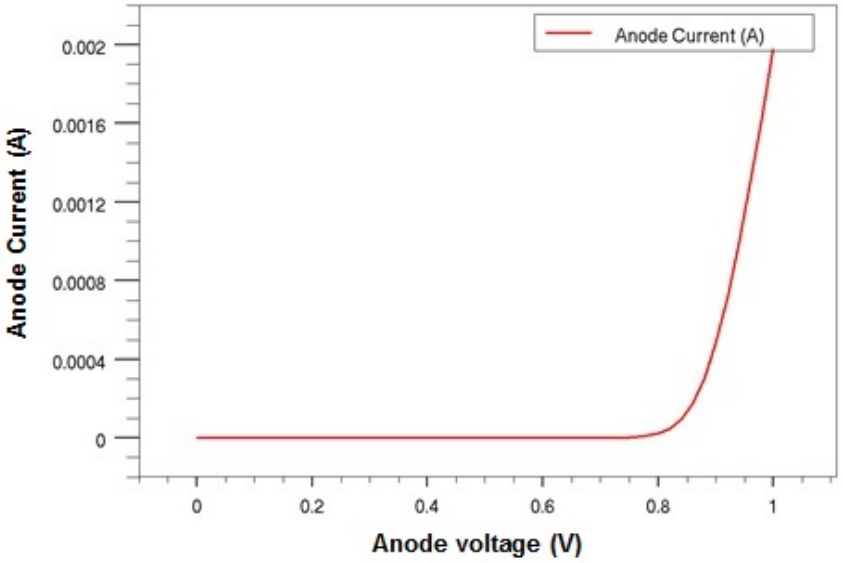

Figure A.5: I-V characteristic of the basic p-Si/n-ZnO heterojunction

A.7.3 Depletion width expansion

Depletion width describes the width of depletion region that is created by the remaining positive and negative space charges near the junction. The total depletion width is the sum of the expansion of depletion region in the n-ZnO and p-Si. For a one-dimensional geometry, taking each region separately and substituting the charge density for each region into the Poisson equation leads the depletion width can be worked out. This is predicted by Anderson model for each region and modelled as [1]:

$$W_{Si} = \left[\frac{2\epsilon_{Si}\epsilon_{ZnO}N_D}{qN_A(N_A\epsilon_{Si} + N_D\epsilon_{ZnO})}(V_{bi} - V)\right]^{0.5} \tag{A.9}$$

$$W_{ZnO} = \left[\frac{2\epsilon_{Si}\epsilon_{ZnO}N_A}{qN_D(N_A\epsilon_{Si} + N_D\epsilon_{ZnO})}(V_{bi} - V)\right]^{0.5} \tag{A.10}$$

the total expansion can be worked as as:

$$W_{total} = \left[\frac{2\epsilon_{Si}\epsilon_{ZnO}(N_A^2 + N_D^2)}{qN_DN_A(\epsilon_{Si}N_A + \epsilon_{ZnO}N_D)}(V_{bi} - V)\right]^{0.5} \tag{A.11}$$

Where ϵ_{ZnO} and ϵ_{Si} are the semiconductor dielectric constants of ZnO ($12 \times 8.85 \times 10^{-12}F/m$) and Si ($16 \times 8.85 \times 10^{-12}F/m$), N_A and N_D are the acceptor and donor concentrations in Si and ZnO respectively, V_{bi} the built-in voltage and V is the applied

bias voltage. The contour plot of depletion width growth is shown in Figure A.6 for certain ranges of N_A and N_D.

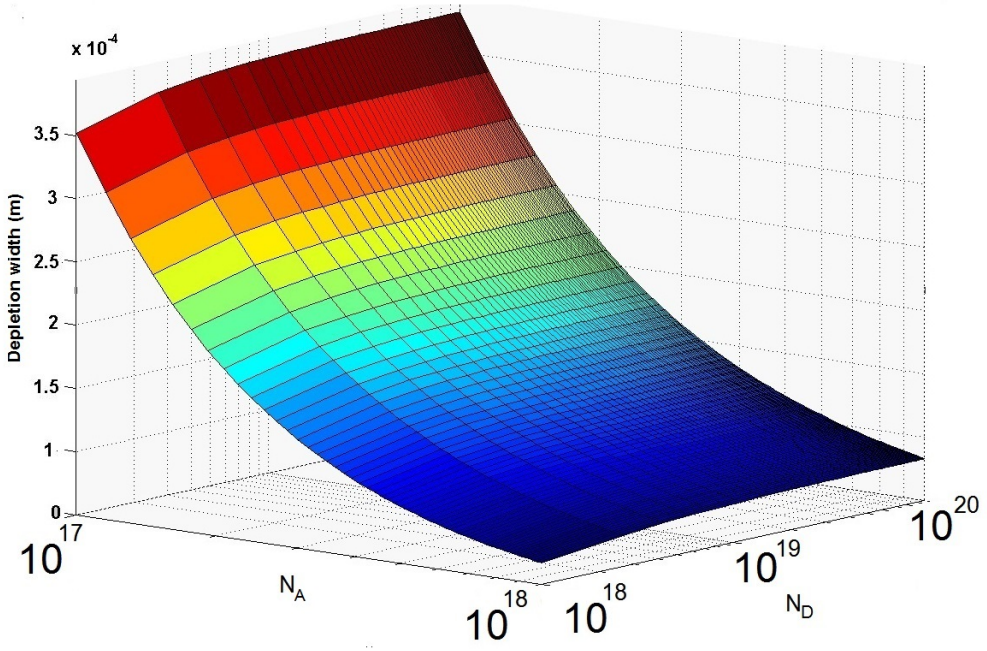

Figure A.6: Depletion width expansion contour plot

The interesting part of total depletion width is however the depletion expansion in Si, since it is the area in which the refractive index has to be altered. The parameters that may affect depletion width expansion in Si are N_A, N_D and V. From equation A.11 one can clearly understand when $N_A >> N_D$ the depletion width is minimum and only varies with N_A. In order to serve certain applications, the above parameters can be also tradeoff against each other, as will be explained shortly. For this case, the initial values of N_A and N_D have been changed to $5 \times 10^{17} cm^{-3}$and $1 \times 10^{19} cm^{-3}$. Since $N_D > N_A$ under reverse bias condition, the depletion will expand towards the p-doped Si.

A.7.4 Capacitance and transient response

Another important parameter that is mainly considered to assess the switching speed of a PN heterojunction is the capacitance. In a typical pn heterojunction, the overall capacitance consists of two components; Depletion and Diffusion capacitances.

Depletion or junction capacitance is due to the stored fixed space-charge region of a p-n junction. The ionized donor and acceptor impurities leave positive and negative fixed charges in depletion region. This can be viewed as a parallel capacitor and determines the transition junction capacitance of the pn junction. The 2D view of this electric field in the depletion region capacitance is shown in Figure A.7.

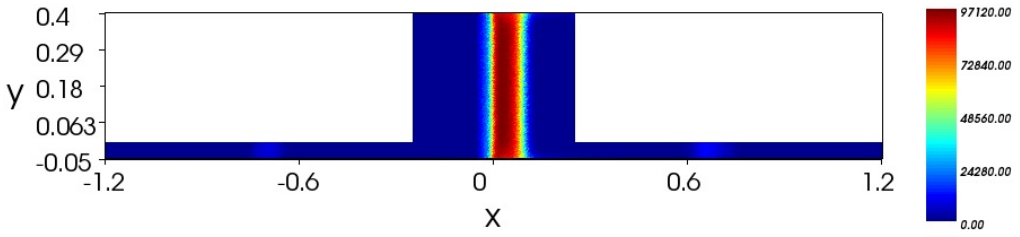

Figure A.7: Electric field in depletion region.

The transition junction capacitance per unit of area can be derived as follows:

$$C_j = \frac{dQ_s}{dV} \tag{A.12}$$

Where Q_s is the space charge on righter side of the depletion region.Considering the charge neutrality principle its value can be obtained from:

$$Q_s = qN_a x_p = qN_d x_n \tag{A.13}$$

Figure A.8 shows the corresponding space charge variation at the p-doped Si/ZnO junction relatives to the change in reverse biased voltage. According to charge neutrality condition in the depletion region, the change in positive and negative space charge in depletion region must be equal if the junction is ideal (Equation A.13). This agreement is clearly demonstrated in Figure A.8 specially at lower reverse voltages.

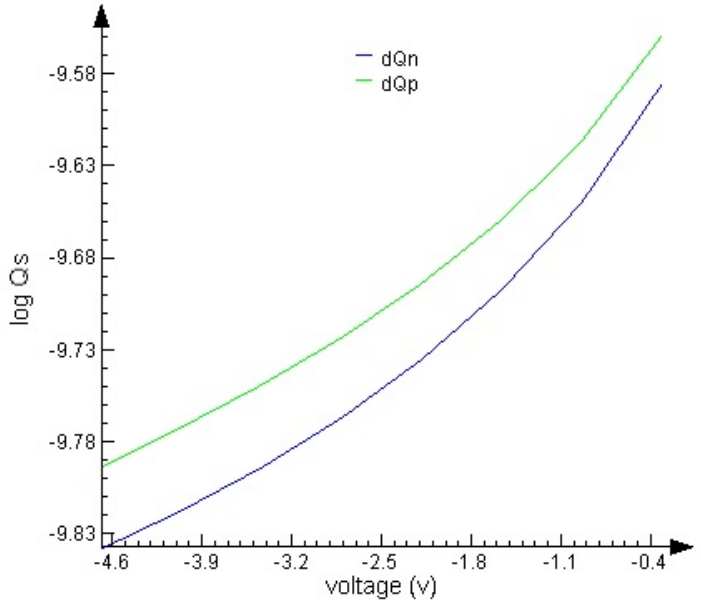

Figure A.8: The variation of positive and negative space charges in depletion region when the reverse voltage is swapped from 0 to -5 V.

However at higher voltages the doping concentrations are such that the junction width on the n side extends towards the contact. This is not the case with the p side. So the C_p is more accurately calculated correctly based on the p charge. The graphs shows that the Q_n side is smaller than Q_p. This is only happening in simulation tool where the delta function at the start of the contact is not modelled. In reality, once the depletion get closer to the n type contact, the rest of the Q_n appears in the form of a delta function, therefore, $Q_n = Q_p$.

Replacing Q_s in Equation A.12 and taking derivatives with respect to the applied voltage, the capacitance per unit of area can be therefore worked out as

$$C_J = \left[\frac{q\epsilon_{Si}\epsilon_{ZnO}N_A N_D}{2(N_A\epsilon_{Si} + N_D\epsilon_{ZnO})}(V_{bi} - V)\right]^{0.5} \tag{A.14}$$

Depletion capacitance is the dominant component when the pn junction goes into depletion mode within the ON state. Figure A.9 shows the numerically calculated capacitance based for positive and negative space charge in depletion region. This result also compared with the analytically calculated junction capacitance for ideal Si pn junction and Si/ZnO heterojunction. The capacitance trend stays in a good agreement with analytical calculations. As can be seen the numerical results shows slightly lower capacitance compared to the analytical calculations due to the presence of highly doped regions under the metallic contacts.

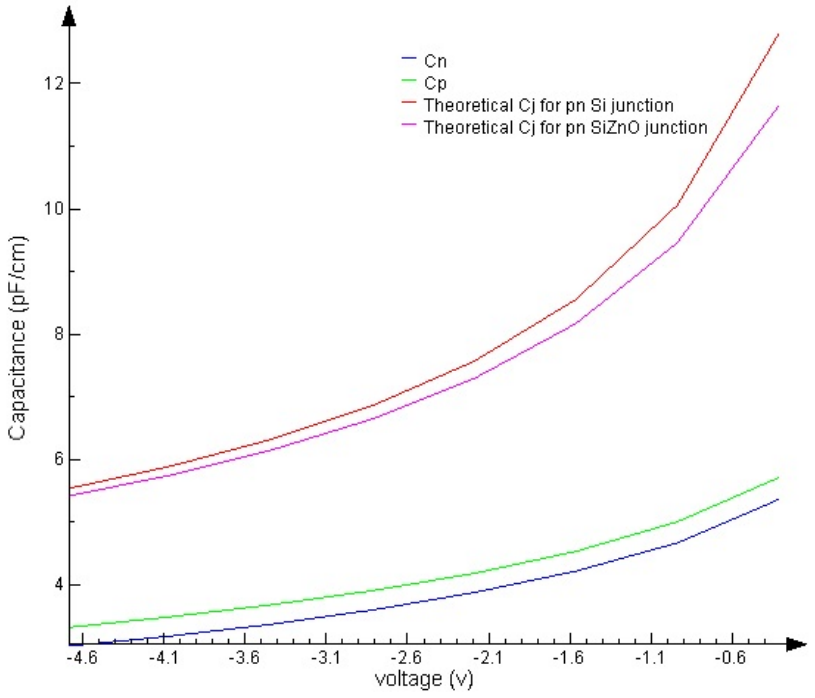

Figure A.9: CV curve of the heterojunction when the reverse voltage is swapped from 0 to -5 V.

It can be also seen that the theoretical calculation of Si/ZnO heterojunction is slightly lower than the capacitance obtained for Si pn junction, although the doping concentration has been set to be the same. This difference is due the larger built in voltage of Si/ZnO (0.9 V) heterojunction compared to Si pn junction (0.7 V). Therefore, it can be estimated that the storage charge behaviour and the transient response of the device is not substantially different in Si pn junction and Si/ZnO heterojunction.

If the proposed device is supposed to also operate in forward bias conditions a small ac signal superimposed and another component has to be taken in to account. This is known as diffusion capacitance. Diffusion capacitance is associated with the minority carriers in the natural quasi regions of the PN heterojunction.

In switching applications where the device operates in reveres biased, the small signal response will be only defined by the depletion capacitance. If the device operates in forward bias as well, the transient behaviour of the minority carrier storage in a p-n junction diode is also very important. In this case the switching time increases since the stored charge must be injected and removed from the neutral quasi regions either by net carrier flow our and/or recombination process. Hence in such an example the switching may be shortened by reducing the stored charge in the neutral quasi regions. Having either a shorter minority carrier lifetime or limited forward current flow can be considered to achieve higher switching speed. Since the proposed device only operates in reveres bias mode, the only concerning capacitance is depletion capacitance as worked out in Equation A.14.

Switching performance can be evaluated by scrutinizing the cathode current input pulse created by the reverse voltage rise time and fall time. The rise time is defined as the time required for the cathode positive current to change from 90% to 10% of the maximum value after the reverse voltage reached to its steady position. Likewise, the fall time is defined as the time required for the negative current to change from 90% to 10% of its minimum value, after the reverse voltage is removed and reached to zero level. Within the steady reverse bias mode the cathode current is negligible(maximum leakage current of 33 nA for a reverse bias of 5 volts for 1 mm device length). During switching mode, when the waveguide region is depleted of carriers, the major current is generated by the switching process. The simulated square wave voltage function in Silvaco which generates such a current is set to have a 100 % rise and fall time.

The initial 20 ps voltage rise time is chosen to demonstrate the variation of reverse voltage effect on switching speed. In Figure A.10a it has been shown that the required rise time for the junction to turn ON is 28 ps. It is also shown that having the same voltage rise time the amount of applied reverse voltage does effect on the switching time. This is expected as the transient time is purely defined by the RC constant of the device (capacitance).

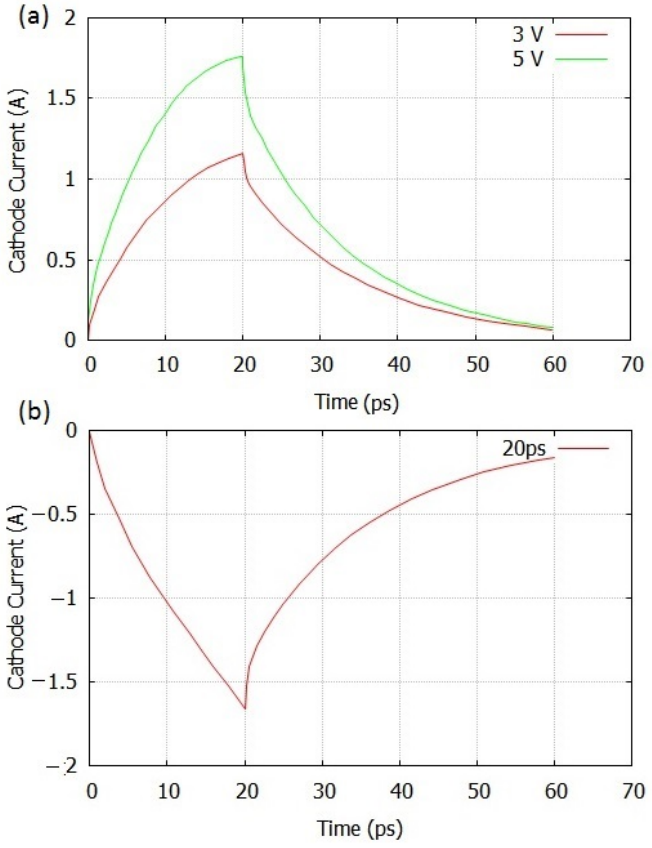

Figure A.10: a) Simulated current rise time of the pn heterojunction device for -3 and -5 V with 20 ps voltage rise time. b) Simulated fall time for -5 v reverse voltage with 20 ps fall time.

The calculated current fall time for 20 *ps* voltage fall time is also shown in Figure A.10b where has been worked out as 38 *ps*. The overall switching time for a pulse is 64 *ps* which can assure a relatively high switching speed performance ($> 15GHz$).

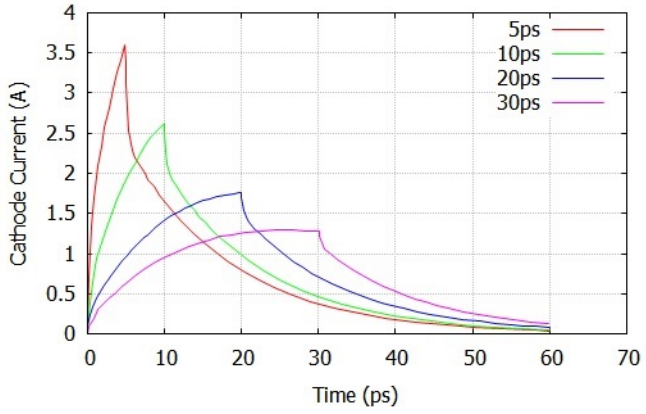

Figure A.11: Current transient response for different voltage rise times.

If the voltage rise time is varied, the switching current involved is also significantly effected as illustrated in Figure A.11.

The above results principally confirm the validity of the heterojunction structure operation and the feasibility of the physical parameters that are used to model such a switching device.

A.8 Electrical characteristic of the heterojunction modulator

A.8.1 Doping Concentration

According to the expressions 2.3- 2.6, the variation in both holes and electrons have almost comparable effects on absorption characteristic of the Si waveguide.

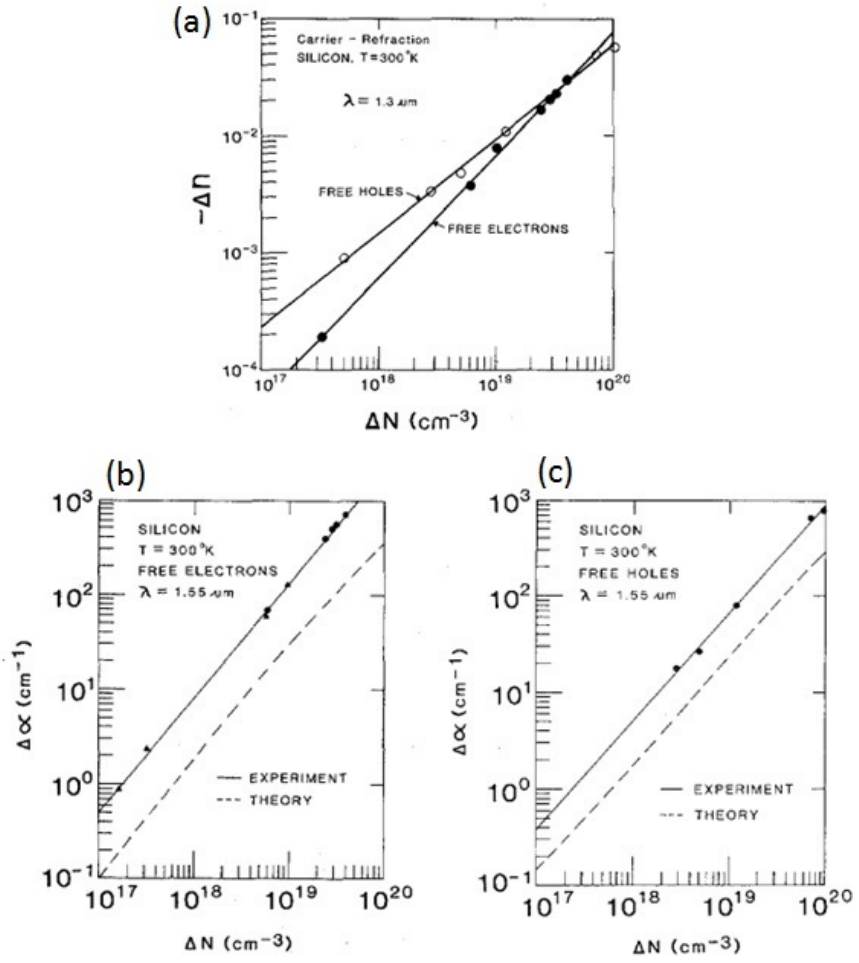

Figure A.12: Refractive index change in Si at 1550 nm wavelength as function of a) free electron and holes, Absorption change in Si at 1550 nm wavelength as function b) free electrons c) free holes [106]

.

However, free holes generally are more effective in perturbing the refractive index than free electrons in Si. This is graphically shown in Figure A.12 at 1550 nm [106]. Hence, the core of the waveguide is designed to be p doped in order to utilize the benefits of free holes effect in the real index of refraction when a phase modulator is proposed.

Maintaining a low absorption loss requires choosing accurate levels of doping concentration in both Si and ZnO. Otherwise along the desired change in refractive index the absorption coefficient may also changes effectively which increase the total amount of loss.

When a greater change occurs in doping concentration of Si, higher refractive index change is achieved; hence the phase of the propagating light wave can be varied in a desirably shorter length. Great change in doping concentration in Si requires high doping density. This higher doping density changes the absorptive characteristics of the material as well as the refractive index changes that it caused. Therefore accurate tradeoff may require to get rid of this limitation as is explained in the following major cases:

1. To keep the absorption loss induced by free carrier below an acceptable amount, the total carrier density change has to be also limited under certain level. The wave light can faithfully have pi phase shift after the wave propagates through a length of the MZI ($L = \lambda/2\Delta n_{eff}$). Here Δn_{eff} is the change in effective index induced by the change in refractive index Δn in the waveguide and is a function of applied voltage. If the doping concentration variation exceeds beyond the critical value it will also results higher absorption, hence the amplitude of the propagating wave is affected. This however reduces the required effective length for a π phase. This additional loss may be compensated in the case where the waveguide has noticeably rough boundaries, as the light is scattered by sidewall roughness in a shorter length. Thus to find the best trade off, choosing appropriate doping concentration is critically important.

2. On the other hand the interesting part of total depletion width is however the depletion expansion in Si, since it is the area in which the refractive index has to be altered. The parameters that may affect depletion width expansion in Si are N_A, N_D and V. For a better understanding of possible effect of these parameters the above expressions are simplified as follows. Consider $N_D = \alpha N_A$. then we have:

$$W_{total} \approx A\sqrt{(V_{bi} - V)\frac{1+\alpha^2}{\alpha(16+12\alpha)N_A}} \tag{A.15}$$

   Where A is constant and equals to:

$$A = \sqrt{\frac{2 \times 12 \times 15 \times 8.85 \times 10^{-12}}{1.6 \times 10^{-19}}} \tag{A.16}$$

It can be seen from these equations that a large expansion can be achieved by having smaller N_A and large value of α. As can be seen a smaller value of N_A results a larger depletion width. However for the cases where larger N_A is desired, the α should be considerably increased in order to maintain a great expansion of depletion in Si. Such an example can be observed in where a greater change of N_A is required in order to achieve a high refractive change. It should be also noted the depletion width expansion remains almost constant after certain level of incrementation in the value of α. The reason is because:

$$Lim_{\alpha \to \infty} \frac{1 + \alpha^2}{\alpha(16 + 12\alpha)} = \frac{1}{12} \tag{A.17}$$

Hence after reaching certain point any further increase in the value of N_D will not affect the depletion width. This discussion can be clearly observed in Figure A.6.

It is therefore important to consider possible tradeoff in these parameters in order to gain an efficient doping profile. As an assumption the doping concentrations in Si and ZnO profiles are now modified to be $5 \times 10^{17} cm^{-3}$ and $5 \times 10^{19} cm^{-3}$ respectively. The maximum acceptable absorption loss caused by free carriers is chosen to be 10 dB/cm. The doping concentration in Si is slightly lower than the critical doping level which causes 10 dB/mm absorption loss ($8.86 dB/cm$). Further optimization may be performed by modifying these values when the second orderre approximation and roughness data is deduced into the structure.

A.8.2 Transient response

Analytical calculation of the capacitance is very complex due to the unusual geometrical specifications. A model which has been originally developed to calculate the capacitance of rectangular transmission line may partially reflects the storage charge behaviour of this structure. Nevertheless this model it is not accurate, especially for any geometrical modifications that may differ the structure from its original form [107]. Using numerical method to obtain the capacitance characteristics of the device is also challenging. This is because at higher voltages ($> 2V$) the device is fully depleted and the carriers are accumulated at the start of the contact as a form a delta Dirac function. In FEM algorithm this condition cannot be accurately modelled and therefore the predicted result will be deviated from its real values for higher voltages. This is shown in Figure A.13 where the capacitance is calculated applying the same numerical calculation which was used to estimate the capacitance of a general Si/ZnO pn heterojunction.

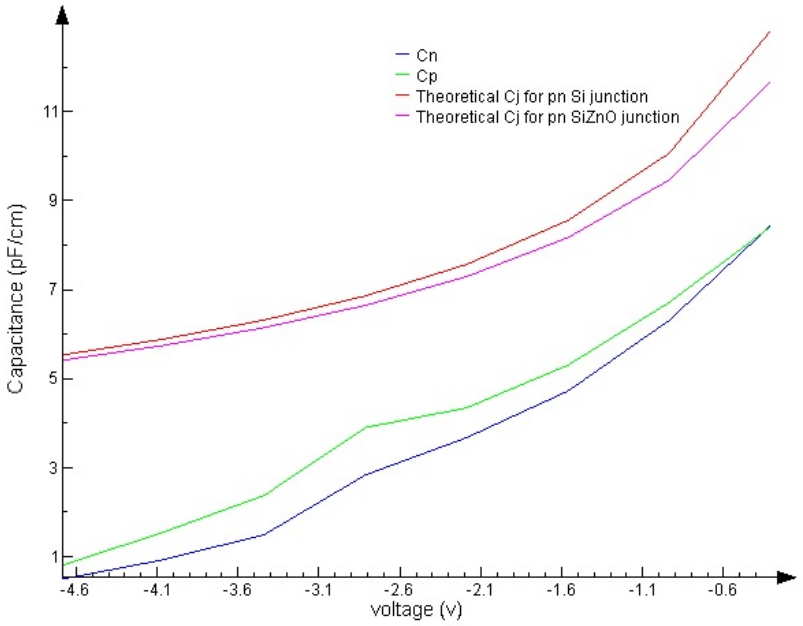

Figure A.13: CV curve of the proposed heterojunction when the reverse voltage is swapped from 0 to -5 V.

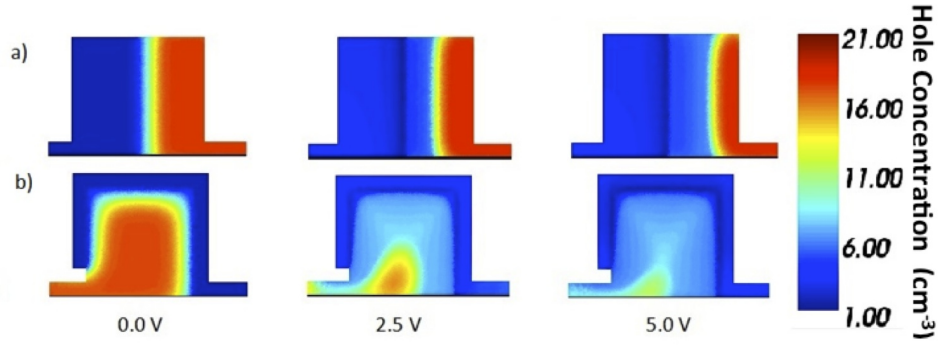

Figure A.14: Hole concentration changes at different reverse voltage conditions for a) simple heterjunction b) proposed heterojunction.

The argument above may be better understood when the carrier distribution is graphically compared in general Si/ZnO heterojunction and the proposed heterojunction under investigation. The comparison is shown in Figure Figure A.14.

For lower reverse voltages ($< 2.5V$ where the results are reliable), the estimated capacitance of the proposed structure is higher than its values for general Si/ZnO pn heterojunction. However the capacitance is still significantly smaller than that reported

in [39] which suggests an increase in performance in terms of RC cutoff frequency and confirms several GHz bandwidths.

The transient response for the device has been also calculated. The rise time and fall time results for 5 V reverse voltage are compared with the general Si/ZnO heterojunction device in figure A.15a and b.

Although a substantial reduction in the cathode current can be observed, the rise and fall time has been increased from 28 *ps* and 38 *ps* to 45 *ps* and 50 *ps* respectively. This reduction is a result of the higher junction capacitance which has fully surrounded the waveguide.

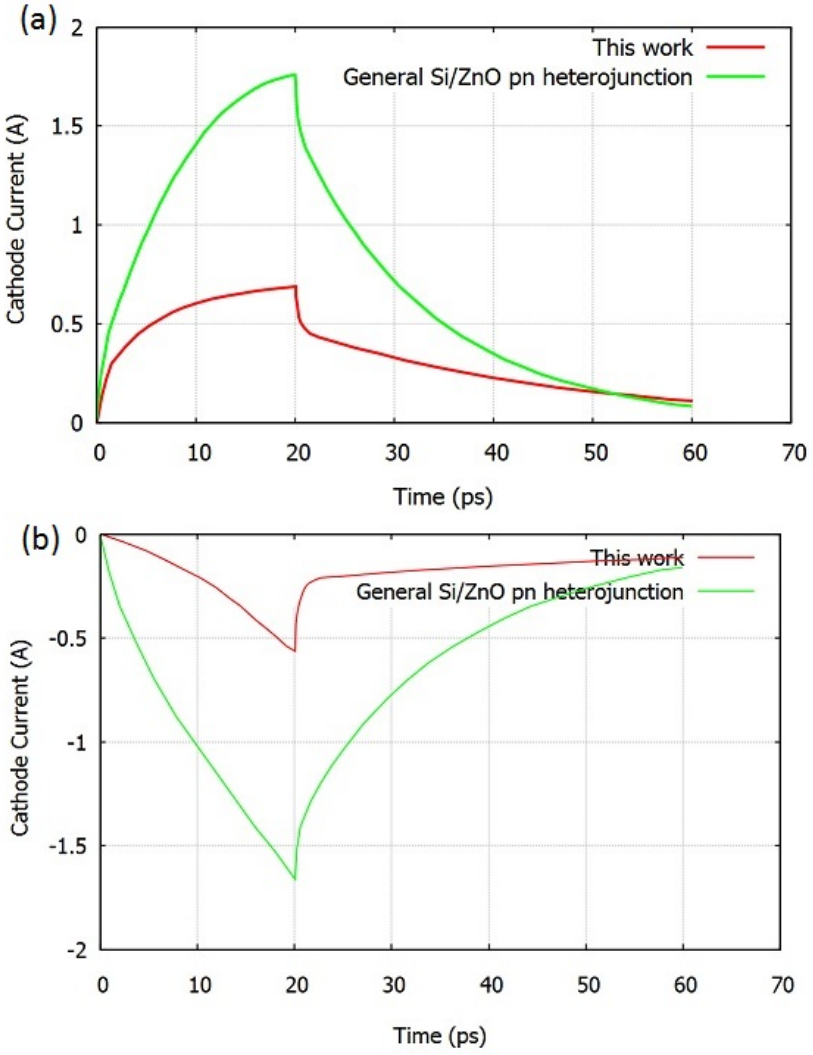

Figure A.15: Current transient response for the proposed structure in comparison with the simple pn heterojunction. a) rise time, b) fall time. The voltage rise and fall time is 20 *ps*.

Furthermore, the device experiences higher resistivity compared to a general form of heterojunction, since for the same reverse voltage (5 V) has led to a lower amount of current across the junction.

Again one can see if the voltage rise time is varied, the switching current involved can be widely effected as illustrated in Figure A.16.

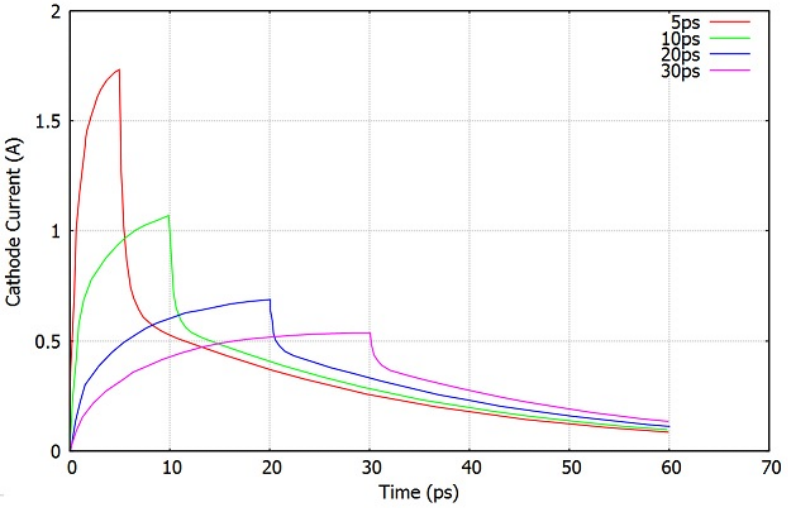

Figure A.16: Current transient response for different voltage rise times.

The above results can principally prove the potential of the proposed structure of this work. The structure is reasonably good enough to assess the roughness as a second order approximation in this work. Nevertheless further optimization in the geometry can improve its switching time performance.

In the next section the optical properties of the proposed structure is demonstrated.

A.9 Optical characteristic of the heterojunction modulator

A.9.1 Geometrical properties

Here the propagating mode properties are discussed by investigating dispersion characteristics and polarization features of the proposed device.

A.9.1.1 Dispersion

The dispersion is demonstrated in the carrier free situation where it is assumed that the free carriers have depleted Si area. Hence it can be approximated as an un-doped

material. Figure A.17 illustrated the relevant dispersion curves of the first 5 propagating mode for the structure discussed above. All results have been worked out for 1550 nm wavelength for minimum attenuation.

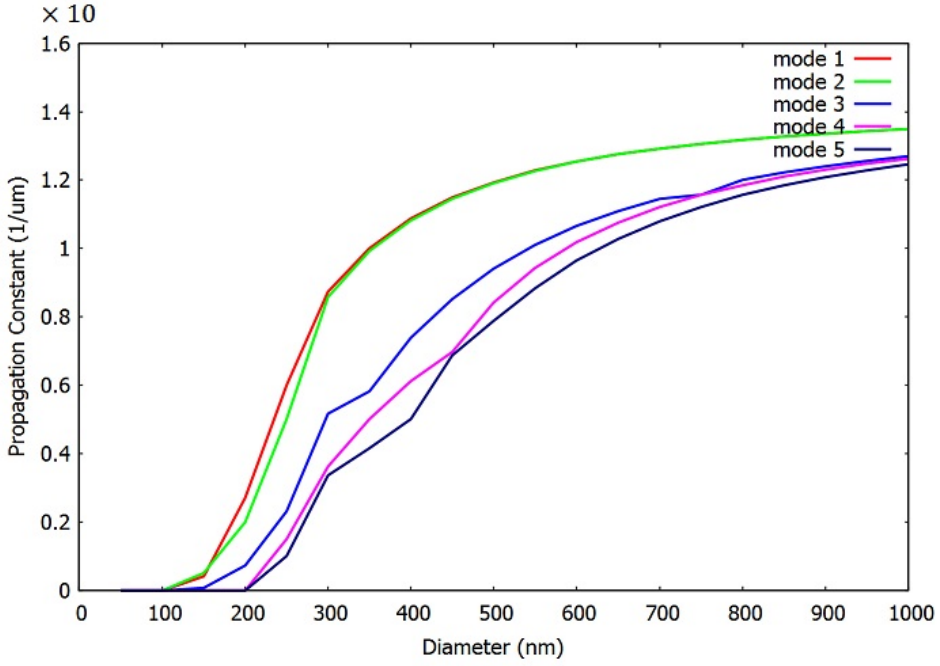

Figure A.17: Dispersion curve for the proposed structure for the first 5 propagating modes.

The waveguide is design to be square with the same width and height. For any mode to exist and be able to propagate in the structure, the effective index has to be at least more than 1. The propagation constant corresponds to the refractive index equals to 1 at 1550 nm wavelength is 4.05 μm^{-3} because [39]:

$$\beta = n_{eff}\frac{2\pi}{\lambda} \tag{A.18}$$

Hence only modes can propagate in the structure if they have a corresponding propagation constant of 4.05 μm^{-3}. This is also confirmed in Figure A.17. As it can be observed when the structure can operate at ON state in the most ideal conditions, can support 3 propagating modes at 300 nm. This is unlike the conditions with air cladding Si where the multimode operation starts after 400 nm. This is due to the added transparent 50 nm ZnO cover that increases the total width of waveguide. Hence the single mode operation of such a structure occurs at lower widths compared to air cladding Si waveguide. The single mode cut OFF of the structure is at 220 nm width and height operating at 1550 nm wavelength. Because the lower contrast of refractive index of Si core and ZnO cladding in comparison with air cladding Si, the modes are less confined in the structures and expand into the cladding area. Hence the propagation constant at wider

apertures of ZnO cladding is lower compared to air cladding Si waveguide. It is also observed that dispersion curves of the first two modes are identical. This is because the TE fundamental modes degenerate the TM modes. This degeneracy of the fundamental mode usually happens when the rectangular waveguide is symmetric and has the same width and height.

A.9.1.2 Polarization

Figure A.18 shows the supported propagating modes for the structure with the width ranging from 300 nm to 550 nm. Because the structure is rectangular, it only supports TE-like and TM-like modes. The fundamental mode is TE0 since 99% of all components of electric field are in the direction of x and perpendicular to Z, the direction of propagation. In other word, only E_x, E_y and H_z components of the total field are significant. The second propagating mode is T_M since only 2% of total electric field is in the direction of x. This confirms the prediction that the TE fundamental modes degenerates a TM mode. These two modes are shown as mode 1 and mode 2 respectively in Figure A.18. As can be seen in Figure A.17 the propagation constant are almost identical for TE and TM fundamental modes, that is the polarization dependence is minimum.

The waveguide currently operates in multimode condition at the chosen dimensions. Further optimization can be later done using relations A.3 and A.4 and the provided graphs to achieve single mode propagation.

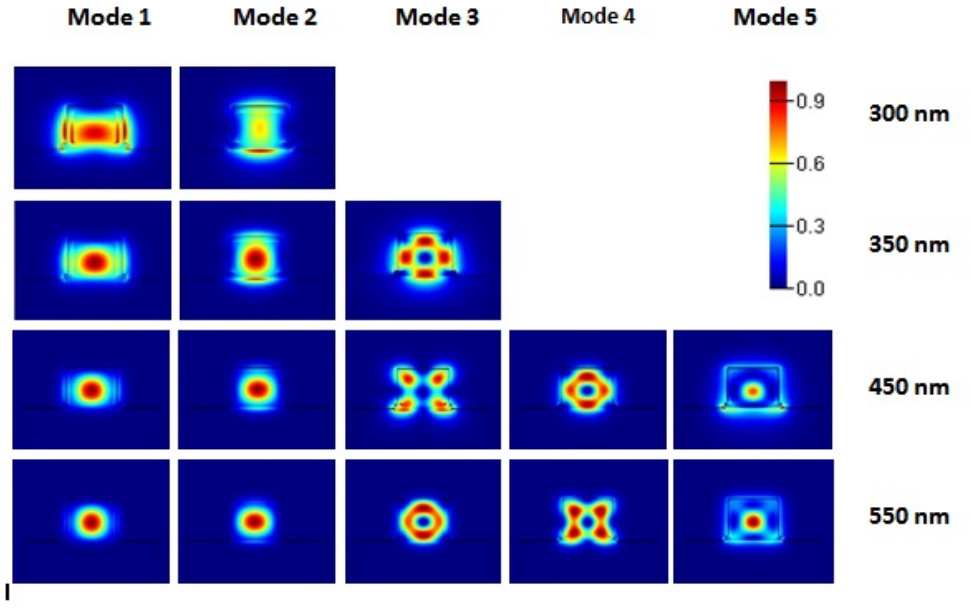

Figure A.18: Supported mode profile for different waveguide cross section width

The differences in confinement of the mode 1 (TE) and 2 (TM) is because the phase change on reflection for the TM mode is always greater than that for the TE mode.

It can be also seen, as expected, the mode profiles are better confined in wider waveguide apertures compared to narrower cross sections. The structure can then be optimized in order to achieve birefringence free behaviour where the effective index of the waveguide is equal for TE and TM polarization.

A.9.2 The effect of electrical carrier distribution in optical characteristics

To capture a clear picture of the effect of electrical carriers on optical behaviour of the device, the electrical data from SILVACO has to be extracted and imported into FDTD solution software.

A.9.2.1 Carrier distribution data extraction

In order to extract the data of the carrier distribution profile, it is required to access the information of each single mesh that is initially defined in ATLAS. Each mesh in ATLAS at specific position of the structure contains particular information of the approximated carrier density at that position. For this reason the mesh size is normally chosen to be quite small for higher accuracy (in order of 2 nm in width and height for the proposed structure).

It is therefore necessary to extract the electron and hole (n and p) profiles at each reverse voltage, interpolates the collected data from a triangular mesh grid into a rectangular mesh grid for use by MODE Solutions, subtract the results from the unperturbed carrier concentration to obtain the amount of concentration change in ON and OFF states and imports the results to a Lumerical data.

These tasks have been done using Matlab. The Soref expressions (Equations 2.3- 2.6) are applied to convert this carrier distribution profile information in to the corresponding refractive index and absorption coefficient data [106].

The absorption coefficient α has to be represented as imaginary refractive index k. If I is the measured intensity of transmitted light after the length of L through the waveguide and I_0 the incident intensity of the light, the optical absorption loss is related to absorption coefficient according to the inverse exponential power law (Beer-Lambert Law) [1];

$$OpticalLoss = \frac{I}{I_0} = e^{-\alpha.L} \tag{A.19}$$

This expression is normally preferred to be represented in dB per unit of length and hence reconfigured as;

$$OpticalLoss(dB/m) = 10log_{10}(\frac{I}{I_0}) = 10log_{10}(e^{\alpha}) = \frac{10\alpha}{Ln10} \tag{A.20}$$

Now the relationship between and k is explained based on the very important equation;

$$\alpha = \frac{n_g}{Re(n_{eff})} \times \frac{4\pi}{\lambda} Im(n_{eff}) \tag{A.21}$$

Where n_g is the group index of the waveguide, λ is the wavelength of operation, $Re(n_{eff})$ and $Im(n_{eff})$ are the real and imaginary effective index of the waveguide respectively. The first term of Equation A.19 is essentially included since the light slowly propagates due to reduced group velocity of the waveguide. This term becomes important in nanoscale waveguides because the group index is significantly larger than the effective index of the mode (in the order of 1.2 larger than effective index) [108]. However, Lumerical does not take this term into account when it calculates the optical loss and hence its effect is ignored by the software. So the relationship between absorption coefficient and imaginary part of refractive index is simply expressed by [108]:

$$\alpha = \frac{4\pi}{\lambda} Im(n) \tag{A.22}$$

This approximation by Lumerical is acceptable since the amount of change in α caused by carrier depletion in the waveguide is small enough to suppress the effect of group index.

By using equation A.22 the absorption coefficient data are converted and represented as the imaginary part of refractive index. These data together with the information of real part of refractive index are now suitable to be used in FDTD solution to optically model.

A.9.2.2 OFF state

When all contacts are grounded and no voltage is applied, the device is considered to operate at OFF state. The holes carrier distribution profile of device at this condition is shown in Figure A.19a. In this condition the maximum concentration of free carriers in p-doped Si is $5 \times 10^{17} cm^{-3}$. Figure A.19b and c illustrate the imported index profile and the corresponding propagating fundamental modes.

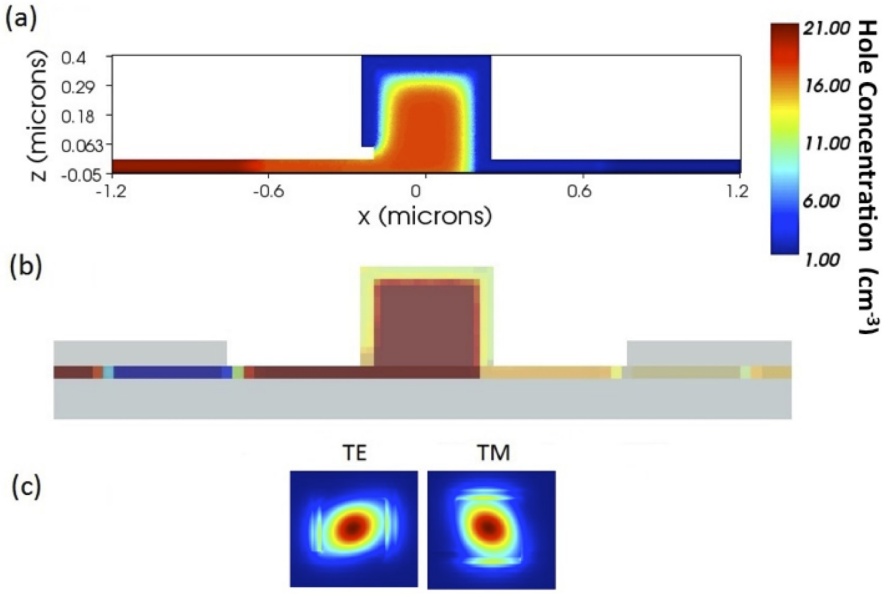

Figure A.19: a) Carrier distribution under 0 V applied voltage, b) imported index profile, c) TE and TM propagating fundamental modes.

A.9.2.3 ON state

Referring back to the original device structure shown in Figure 5.1, once the heterojunction diode turns to ON state the depletion region starts to expand as a result of electric field caused by reversed biased voltage. The bandgap structure is reconfigured as shown in Figure A.4. Consequently, the free holes are efficiently swept out from the p doped region. This provides an almost free path for the light to progress with negligible absorption, if the device is optimally revered biased.

The graphical illustration of index profile (including real refractive index and absorption coefficient) that is imported in FDTD solutions in Lumerical is shown in Figure A.20b for state. The first two supported mode profile (TE and TM) is also shown in Figure A.20c.

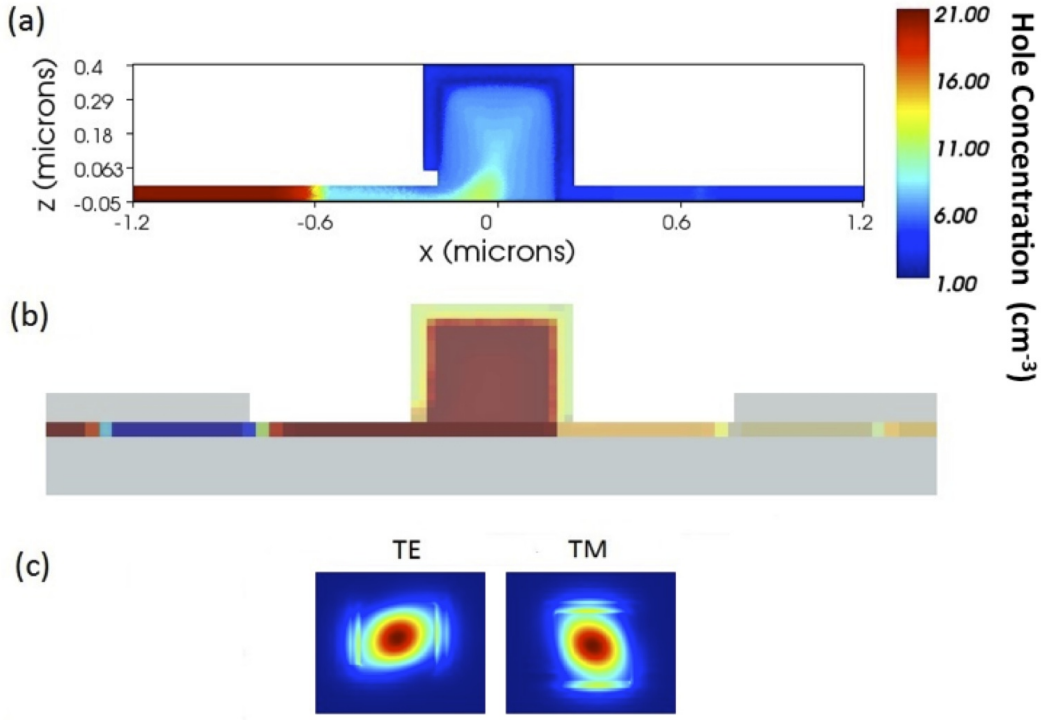

Figure A.20: a) Carrier distribution under -5 V applied voltage, b) Imported index profile, c) TE and TM propagating fundamental modes.

A.9.2.4 Optical response

The redistribution of holes density in p-Si changes the real and imaginary part or refractive indexes. This change is normally quantified by considering the variation of the effective index and the propagation loss characteristics of the structure due to the external reverse voltage. Figure A.21c and d shows the variation of effective index when the voltages is reversed from 0 to 7 V for TE and TM. In order to justify the use of ZnO in the p-Si/n-ZnO proposed structure this simulation results are also compared with similar p-Si/n-Si structure in Figure A.21a and b with identical doping concentration and dimensions.

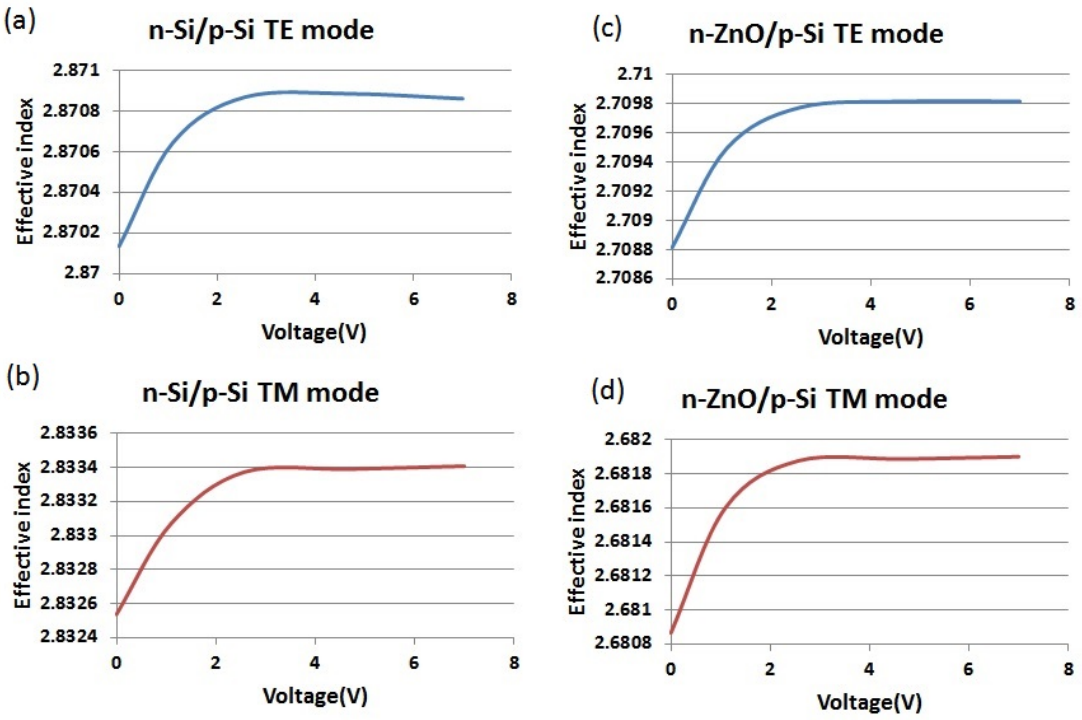

Figure A.21: Effective index change in a) p-Si/n-Si TE mode b) p-Si/n-Si TM mode c) p-Si/n-ZnO TE mode and d) p-Si/n-ZnO TM mode as the reverse voltage varies from 0 to 7 V.

Figure A.20 clearly shows that the device is fully depleted (maximum p doped concentration is ¡ $1 \times 10^{1}6cm^{-3}$) when 7 V reverse voltage is applied. From Sorefs Expression 2.3-2.6 having $5 \times 10^{-17}cm^{-3}$ change in the holes carrier concentration results a change in the real part of refractive index in the order of 0.001. In Figure A.21 such a change in refractive index causes almost 0.001 change in the effective index of the waveguide structure consequently. Although the effective index of p-Si/n-Si is higher, the effective index change within ON and OFF states shows comparable amount of change in both structures.

The required length to have a π phase shift is worked out Using ($L = \lambda/2\Delta n_{eff}$), which is 775 μm.

Figure A.22 compares the loss performance of the p-Si/n-ZnO proposed structure with the identical p-Si/n-Si device, again when the voltages is reversed from 0 to 07 V.

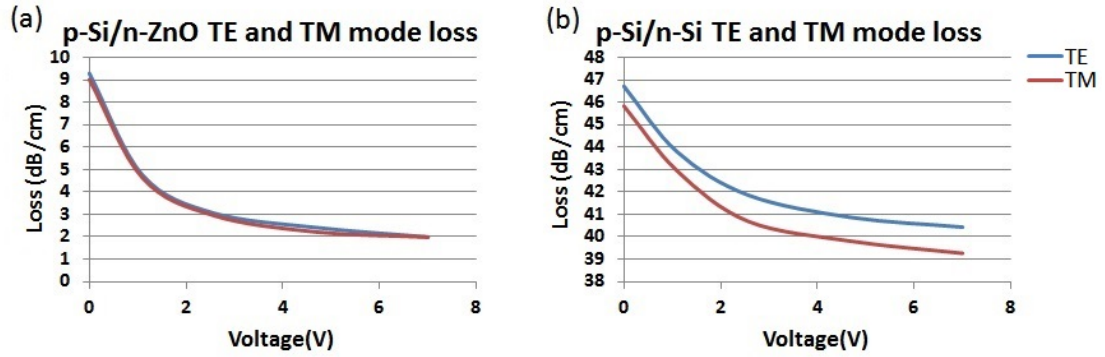

Figure A.22: Loss performance for a) p-Si/n-ZnO TE mode b) p-Si/n-Si for TE and TM mode as the reverse voltage varies from 0 to 7 V.

As TE and TM modes experience different degrees of optical confinement, the light has different powers at the waveguide core/cladding interfaces. Furthermore the propagation angles are different and therefore one mode will have more reflections at the interface than the other. Consequently the loss at the interfaces will be also different for TE and TM modes as can be seen in the Figure A.22.

Figure A.22 shows at zero biased voltage, when the device is in OFF state, the absorption loss is almost 9.3031 dB/cm. The loss is reduced by 7.33 dB/cm when the device turns to ON state by reversing the bias voltage to -7 V. Hence the change of carrier absorption loss from OFF to ON state is 7.33 dB/cm which is lower than the theoretically estimated absorption loss induced by $\approx 5 \times 10^{-17} cm^{-3}$ holes carrier change(8.86 dB/cm). Such a difference arises from the following issue: In theoretical calculations it has been assumed that the whole waveguide area is depleted from both electrons and holes. However in reality there are still free electron carriers at the interface of the pn heterojunction. This is shown in Figure A.23. Furthermore, there is an additional loss of 0.635 dB/cm related to the highly doped region under the aluminium contacts and free holes in the pad area.

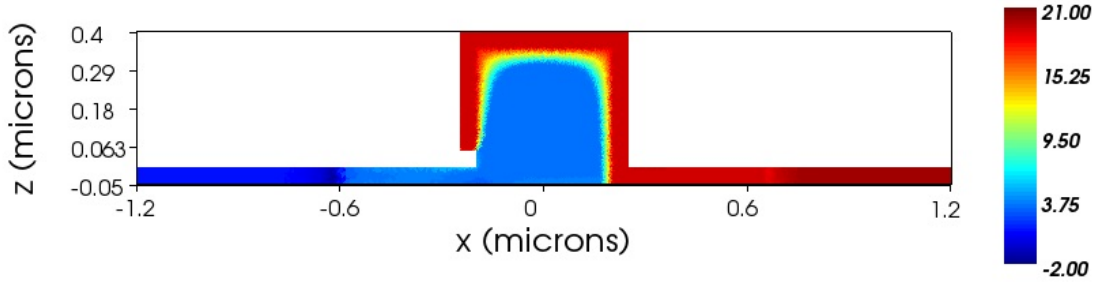

Figure A.23: Electron carrier distribution effect at 7 V

Once the phase shifter is designed it can be implemented into the arm of MZI to convert the resulting phase change to the desired amplitude change.

A.10 Mach Zehnder Interferometer

Mach Zehnder interferometer (MZI) is widely used in interference based optical modulators which split light at their input, then introduce phase delay in one arm and finally recombine them at the output. Compared to resonant based structures they offer wider bandwidths even though they scale the overall footprint. In addition, the structure is less sensitive to temperature fluctuations [1].

The quality of interference is determined by the transmitted intensity at the output. Intensity of output light (for lossless and equally divided power) is given by the well-known interferometer transfer function that is [1]:

$$I = I_0/2\,[1 + cos(\beta_2 L_2 - \beta_1 L_1)] \tag{A.23}$$

Where I is the intensity of light at the output, I_0 is the intensity of incident light at the input, , L_1 and β_1 are the length and the propagation constant in arms 1, L_2 and β_2 are the length and the propagation constant constant in arm 2. The term $\beta_2 L_2 - \beta_1 L_1$ represents the phase difference between the two arms of interferometer.

The Input light is divided into two beams propagating in the two arms of the MZI. Considering the loss factor in both arms is equivalent, having $\beta_2 L_2 - \beta_1 L_1$ equal to π when the device is at ON state, the transfer function predicts 100% extinction ratio since

$$\beta_2 L_2 - \beta_1 L_1 = \pi \tag{A.24}$$

and hence $I = 0$.

the modulation depth or modulation index is given by

$$\eta = \frac{|I_0 - I_{min}|}{I_0}(\%) \tag{A.25}$$

where I_{min} is the minimum transmitted intensity and I_0 is the value of intensity I when no electrical signal is applied.

Using an MZI structures large extinction ratio is obtainable only if the total input power is equally divided between two arms. This necessitates the use of splitters which are able to keep both arms in balance for an optimum performance. It also requires to cause low optical loss, provides relatively large optical bandwidth, small footprint, and low sensitivity to fabrication tolerances and thermal fluctuations [109].

A.10.1 Multimode Interference Splitter

The simplest way to split and couple the input light is through using Y junction. Y junction requires a small angle for minimal insertion loss. This will inherently increase the overall lengths and hence results relatively low extinction ratios [109].

Star coupler structure is another alternative [110]. The structure exhibits a low loss operation with a large optical bandwidth and small footprint. Nevertheless the device fabrication is complex and static extinction ratio was only 14 dB [110].

The other alternative is a multimode interference (MMI) structure. The optical modulator utilise MMI has been reported with an extinction ratio of approximately 25 dB [46]. MMI have typically lower lengths of about 30 - 40 μm [46]. The splitting ratio is insensitive to temperature changes.

The structure consists of one central input and two output outputs that are positioned equally separated from the centre of the MMI. In order to balance the loss in two arms, the input and two output ports are normally designed to be identical in terms of the cress sectional dimensions. The central part of MMI is designed to be large and support higher order of modes.

Here a basic form of MMI configuration that is used in [109] is designed. A more precise MMI configuration that has lower loss can be achieved by tapering the input and outputs. However, since it has been shown that the static extinction ratio of the MZI is insensitive to the MMI insertion loss there would be therefore no such an urgency in the purpose of this work.

The input and output port dimensions are the same as the actual cross section of the input waveguide (400×400 nm). Two outputs are positioned equally separated from the centre of the MMI ($3.2\mu m$). The width and length of the central part of the MMI has to be precisely designed in order to achieve minimal loss; To avoid evanescent coupling between output waveguides, the width of the MMI region has to be chosen wide enough (in the order of 6 μm) ensuring that the output waveguides are located sufficiently separated [109].

In self-imaging theory the length has a quadratic dependence to the width of the MMI region based on [111]:

$$L \approx \frac{n_{eff} W^2}{2\lambda} \tag{A.26}$$

where W and L are the required width and length of the MMI. FDTD solution is applied to model the 3D MMI structure. The calculated effective index is 2.8. Using Equation A.26 and having 6 μm width gives the length of 33.56 for the MMI region.

Figure A.24a and b respectively shows the 3D schematic of a typical MMI and the resulting electric field intensity pattern.

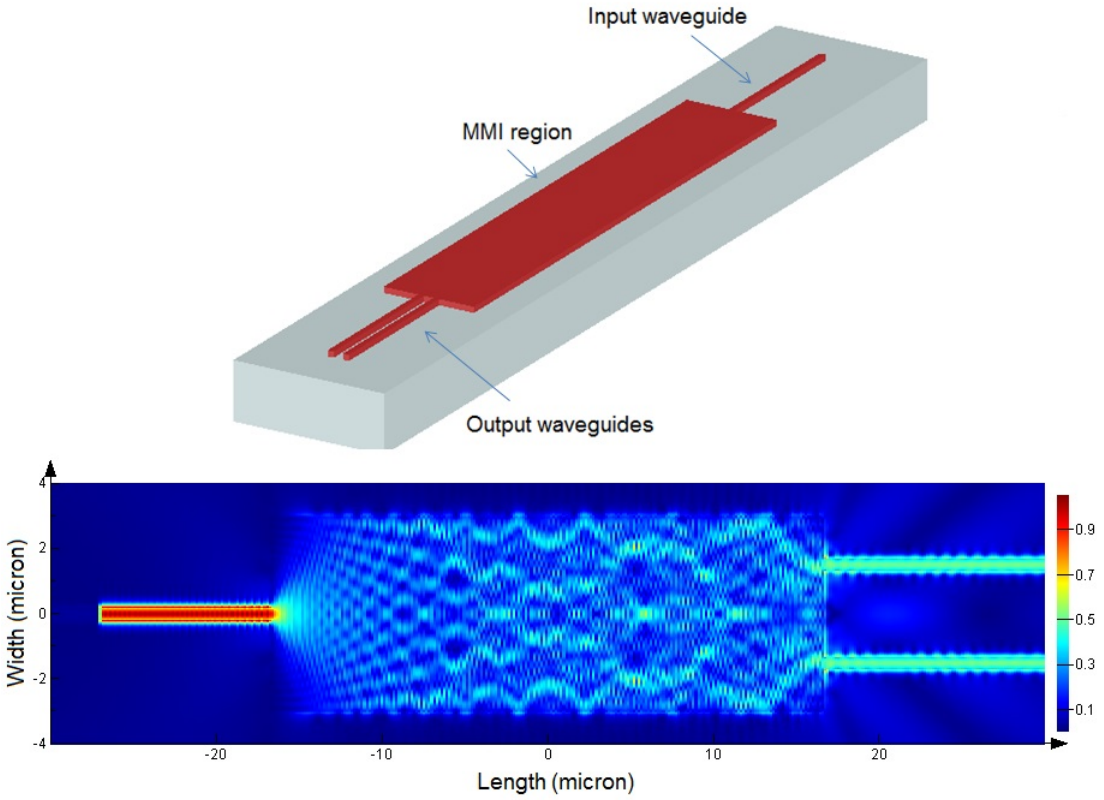

Figure A.24: a)MMI layout b) Intensity pattern inside the designed MMI. Simulations where obtained using FDTD solutions.

The reflection into the fundamental mode of input waveguide is calculated to be 0.474 % (compared to the ideal detached waveguide). The transmission into the fundamental mode of output waveguide 1 and 2 are calculated to be 28.6245 % and 28.5923 % respectively. These two modes are then imported and propagate through the MZI arms as will be explained shortly. It can be also seen that the transmitted power into the output ports is reasonably balanced. The overall loss of the MMI region has been calculated as 0.8 dB which results in 83% of the input power transmitted to the output waveguides. Therefore the overall loss induced by coupler and splitter can be estimated as 1.6 dB.

A.10.2 Modulation Response

Figure A.25 shows the scaled MZI structure under investigation. The input light is divided into two light beams propagating in the two arms of the MZI. The phase shifter is added in one arm. An identical structure can be also inserted into the other arm but must not be driven during modulation and therefore it will only balance the loss in

both arms. To stay within the limit of the available memory space for simulating the structure, here the Mach Zehnder is only implemented into one arm.

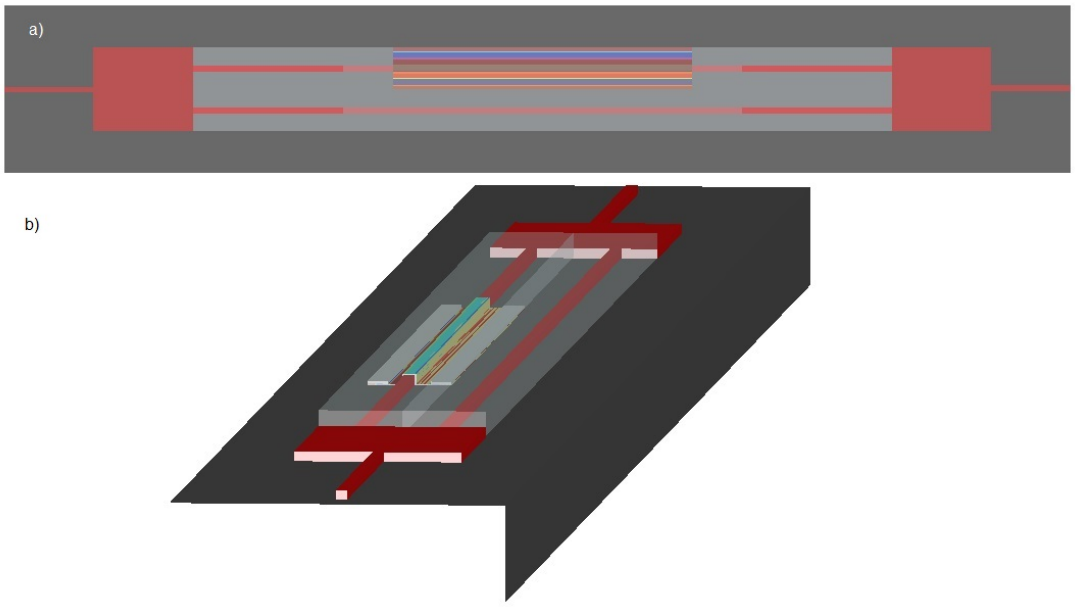

Figure A.25: a)2D scaled MZI structure layout b) 3D scaled MZI structure. Note that the Phase shifter has been inserted into only one arm.

Due to the long active length of MZI impractically large memory space would be required to simulate the entire structure in a single run. Alternatively the extinction ratio and the loss performance can be obtained using modulation transmission function that includes the loss operator as well. This contains the waveguide insertion loss and is given by:

$$E(V_1, V_2) = \frac{E_0}{1+\sigma}(\sigma e^{-\frac{2\pi}{\lambda_0} n_{eff1} L} + e^{-\frac{2\pi}{\lambda_0} n_{eff2} L}) \tag{A.27}$$

where, σ is the splitting ratio which is unity for perfect splitting, n_{eff1} and n_{eff2} are the effective index corresponds to voltage V_1, the voltage in arm 1 and V_2, the voltage applied on arm 2 respectively. L is also the length of the modulator. Taking the calculated effective index into account the modulation response of the MZI structure can be approximated as shown in Figure A.26. Note that the transmission power $((\frac{E}{E_0})^2)$ is normalized to the range between 0 and 1.

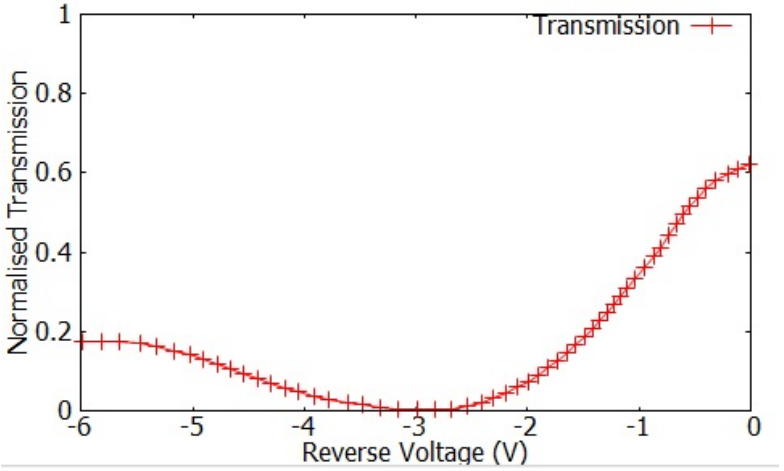

Figure A.26: Modulation response of the Mach Zehnder.

Figure A.26 shows the MZI will have a π phase shift at 3 volts as expected when the length chosen to be 0.775 μm. The waveguide arms insertion loss and extinction ratio of the modulator is also shown in Figure A.26. This corresponds to 2.5 and 23.7 dB respectively. The insertion loss is defined as the amount of loss that the MZI have when no voltage is applied. The calculated insertion loss here include 1.6 dB of the MMI, and absorption loss caused by free carrier in off state. Furthermore the On chip insertion loss may be also taken into account which has to be added on top for practical cases. The on-chip loss associated with the entire MZI device. This on chip insertion losses in practical cases includes waveguide transmission losses caused by scattering, tapers and access waveguide losses, the losses induced by the splitter/coupler and finally the absorption losses due to free carrier plasma effect in the phase shifter. However, no tapered waveguide has been used within the MZI structure and hence tapers and access losses are not concern of this work at this stage and will be neglected. Among of other the scattering loss due to the roughness is considered for the phase shifter in the next section.

A.10.3 ZnO Coating Thickness Effect

The ZnO coating thickness can alter free carrier distribution in ON state when the electric field is applied under reverse bias voltage. Thinner ZnO coating will result a more efficient transport of free electron at the contact and ZnO interface. As a result a better depletion can be observed in where the device has a thinner n-ZnO thickness as shown shown in Figure A.27.

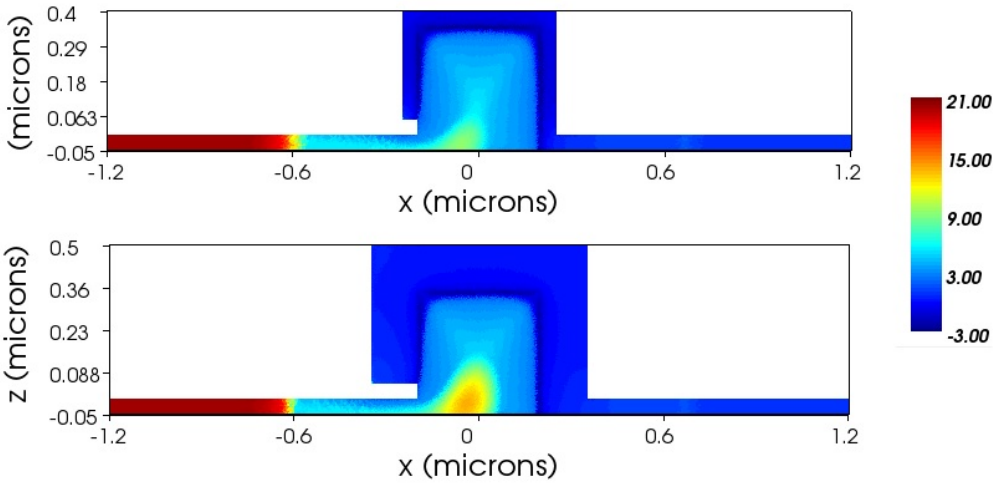

Figure A.27: ZnO thickness effect on carrier distribution under reverse voltage

The optical response in terms of loss and effective index for different ZnO thickness and applied voltage effect is shown in Figure A.28. The effective index is slightly varied when the ZnO coating thickness is changed from 50 nm to 150 nm since carrier distribution density of the p-doped Si core increases as ZnO coating becomes thicker.

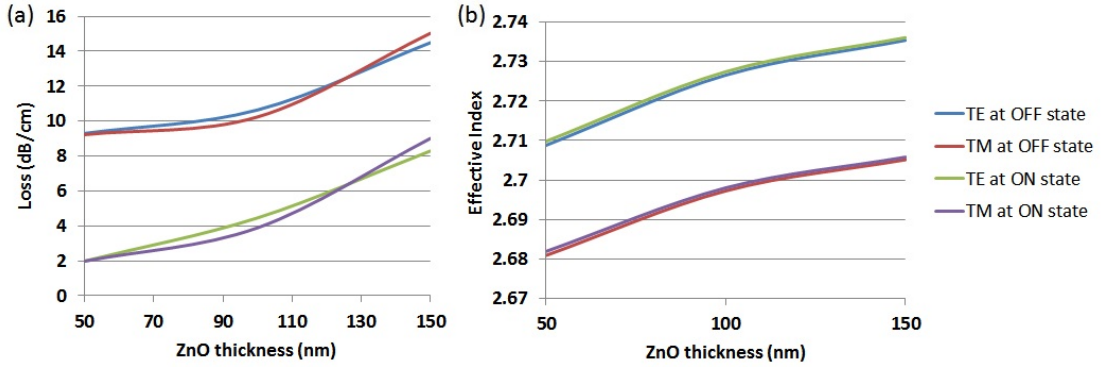

Figure A.28: ZnO thickness effect on a) loss and b) effective index

Parallel simulations reveals the same argument is valid when the Si core waveguide size becomes larger. However the rate of change of effective index due to change of ZnO coating, is higher at smaller Si core size.

The switching capability of the device is then evaluated by implementing the-Si/n-ZnO switch in the Mach Zehnder interferometer configuration.

References

[1] Ed. G. T. Reed, *Silicon photonics the state of the art*. Wiley-Interscience, 2008.

[2] C. Gunn, "CMOS photonics for high-speed interconnects," *Micro, IEEE*, vol. 26, no. 2, pp. 58–66, 2006.

[3] Y. Vlasov and S. McNab, "Losses in single-mode silicon-on-insulator strip waveguides and bends," *Opt. Express*, vol. 12, no. 8, p. 16221631, Apr. 2004.

[4] K. K. Lee, D. R. Lim, H. C. Luan, A. Agarwal, J. Foresi, and L. C. Kimerling, "Effect of size and roughness on light transmission in a Si/SiO[sub 2] waveguide: Experiments and model," *Applied Physics Letters*, vol. 77, no. 11, p. 16171619, 2000.

[5] T. Tsuchizawa, T. Watanabe, E. Tamechika, T. Shoji, K. Yamada, J. Takahashi, S. Uchiyama, S. Itabashi, and H. Morita, "Fabrication and evaluation of submicron-square si wire waveguides with spot size converters," vol. 1, 2002, pp. 287–288 vol.1.

[6] F. P. Payne and J. P. R. Lacey, "A theoretical analysis of scattering loss from planar optical waveguides," *Optical and Quantum Electronics*, vol. 26, no. 10, p. 977986, Oct. 1994.

[7] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, "Silicon optical modulators," *Nat Photon*, vol. 4, no. 8, pp. 518–526, 2010.

[8] G. Z. Mashanovich, M. M. Milo?evi?, M. Nedeljkovic, N. Owens, B. Xiong, E. J. Teo, and Y. Hu, "Low loss silicon waveguides for the mid-infrared," *Opt. Express*, vol. 19, no. 8, p. 71127119, Apr. 2011.

[9] D. K. Sparacin, S. J. Spector, and L. C. Kimerling, "Silicon waveguide sidewall smoothing by wet chemical oxidation," *Lightwave Technology, Journal of*, vol. 23, no. 8, p. 24552461, 2005.

[10] Shih-Che Hung, Eih-Zhe Liang,, and Ching-Fuh Lin, "Silicon waveguide sidewall smoothing by KrF excimer laser reformation," *Journal od lightwave technology*, vol. 27, no. 7, pp. 887–892, 2009.

[11] Azzedina Boudrioam, *Photonic Waveguide; Theory and Application.* John Wiley & Sons, 2009.

[12] J. Wilison and J. Hawkes, *Optoelectronics; An Introduction.* Prentice Hall, 1998.

[13] R. suref and B. Bennett, "Electro-optical effect in silicon," *IEEE*, vol. 23, no. 1, pp. 123–129, 1986.

[14] R. Williams, "Electric field induced light absorption in CdS," *Phys. Rev.*, vol. 117, no. 6, p. 14871490, 1960.

[15] Y. Kuo, Y. Lee, Y. Ge, S. Ren, J.E Roth, T.I Kamins, D.A.B Miller, and J.S Harris, "Quantum-confined stark effect in Ge/SiGe quantum wells on si for optical modulators," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 12, no. 6, pp. 1503–1513, 2006.

[16] L V. Keldysh, "Ionization in the field of a strong electromagnetic wave," *Lebedev Inst. of Physics, Moscow*, vol. 20, no. 5, pp. 1307–1314, 1965.

[17] S. A. Clark, B. Culshaw, E. J. Dawnay, and I. E. Day, "Thermo-optic phase modulators in SIMOX material," pp. 16–24, 2000.

[18] N. Ky, H. Limberger, R. Salathe, and G. Fox, "Optical performance of miniature all-fiber phase modulators with ZnO coating," *Lightwave Technology, Journal of*, vol. 14, no. 1, pp. 23–26, 1996.

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

[20] S. R. Giguere, L. Friedman, R. A. Soref, and J. P. Lorenzo, "Simulation studies of silicon electro-optic waveguide devices," *Journal of Applied Physics*, vol. 68, no. 10, pp. 4964–4970, 1990.

[21] G. Cocorullo and I. Rendina, "Thermo-optical modulation at 1.5 mu m in silicon etalon," *IEEE Electronics Letters*, vol. 28, no. 1, pp. 83–85, 1992.

[22] M. Imai, T. Shimizu, Y. Ohtsuka, and A. Odajima, "An electric-field sensitive fiber with coaxial electrodes for optical phase modulation," *Lightwave Technology, Journal of*, vol. 5, no. 7, pp. 926–931, 1987.

[23] A. Liu, L. Liao, D. Rubin, H. Nguyen, B. Ciftcioglu, Y. Chetrit, R. Cohen, N. Izhaky, J. Basak, and M. J. Paniccia, "Recent advances in high speed silicon optical modulator," pp. 647 710–647 710–9, 2007.

[24] H. Haug and S. W. Koch, *Quantum Theory Of The Optical And Electronic Properties Of Semiconductors*, edition 5 ed. World Scientific Publishing Company, 2009.

[25] R. A. Soref and B. R. Bennett, "Kramers-kronig analysis of electro-optical switching in silicon," *Spie Proceedings*, vol. 0704, pp. 32–37, 1987.

[26] L. Friedman, R. A. Soref, and J. P. Lorenzo, "Silicon double-injection electro-optic modulator with junction gate control," *Journal of Applied Physics*, vol. 63, no. 6, pp. 1831–1839, 1988.

[27] G. Treyz, P. May, and J.-M. Halbout, "Silicon optical modulators at 1.3 um based on free-carrier absorption," *Electron Device Letters, IEEE*, vol. 12, no. 6, pp. 276–278, Jun. 1991.

[28] 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.

[29] S. Jackson, G. Reed, C. Tang, A. Evans, J. Clark, C. Aveyard, and F. Namavar, "Optical beamsteering using integrated optical modulators," *Journal of Lightwave Technology*, vol. 15, no. 12, pp. 2259–2263, 1997.

[30] S. M. Jackson, P. D. Hewitt, G. T. Reed, C. K. Tang, A. G. R. Evans, J. Clark, C. Aveyard, and F. Namavar, "A novel optical phase modulator design suitable for phased arrays," *Journal of Lightwave Technology*, vol. 16, no. 11, p. 2016, 1998.

[31] C K Tang, G. T. Reed, A. J walton, and A G Richman, "Simulation of a low loss optical modulator for fabrication in simox material," *MRS Procedding*, vol. 298, p. 247, 1993.

[32] C. Tang, G. Reed, A. Walton, and A. Rickman, "Low-loss, single-model optical phase modulator in SIMOX material," *Journal of Lightwave Technology*, vol. 12, no. 8, pp. 1394–1400, Aug. 1994.

[33] A. Cutolo, M. Iodice, P. Spirito, and L. Zeni, "Silicon electro-optic modulator based on a three terminal device integrated in a low-loss single-mode SOI waveguide," *IEEE Journal of Lightwave Technology*, vol. 15, no. 3, pp. 5.5–518, 1997.

[34] P. D. Hewitt and G. T. Reed, "Multimicron dimension optical p-i-n modulators in silicon-on-insulator," *Proc. SPIE*, vol. 3630, 1999.

[35] P. D. Hewitt and G. T. Reed, "Improving the response of optical phase modulators in SOI by computer simulation," *Journal of Lightwave Technology*, vol. 18, no. 3, p. 443, 2000.

[36] I. Day, S. Roberts, R. O'Carroll, A. Knights, P. Sharp, G. Hopper, B. Luff, and M. Asghari, "Single-chip variable optical attenuator and multiplexer subsystem integration," 2002, pp. 72–73.

[37] R. R. Whiteman, A. P. Knights, D. George, I. E. Day, A. Vonsovici, A. A. House, G. F. Hopper, and M. Asghari, "Recent progress in the design, simulation, and fabrication of small cross-section silicon-on-insulator VOAs," *Spie Proceedings*, pp. 146–156, 2003.

[38] P. D. Hewitt and G. Reed, "Improved modulation performance of a silicon p-i-n device by trench isolation," *Lightwave Technology, Journal of*, vol. 19, no. 3, pp. 387–390, 2001.

[39] C. E. Png, S. P. Chan, S. T. Lim, and G. Reed, "Optical phase modulators for MHz and GHz modulation in silicon-on-insulator (SOI)," *Journal of Lightwave Technology*, vol. 22, no. 6, pp. 1573–1582, 2004.

[40] S. Manipatruni, Q. Xu, B. Schmidt, J. Shakya, and M. Lipson, "High speed carrier injection 18 gb/s silicon micro-ring electro-optic modulator," 2007, pp. 537–538.

[41] W. M. Green, M. J. Rooks, L. Sekaric, and Y. A. Vlasov, "Ultra-compact, low RF power, 10 gb/s siliconMach-Zehnder modulator," *Optics Express*, vol. 15, no. 25, p. 1710617113, 2007.

[42] L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, U. Keil, and T. Franck, "High speed silicon mach-zehnder modulator," *Opt. Express*, vol. 13, no. 8, p. 31293135, 2005.

[43] K. Kajikawa, T. Tabei, and H. Sunami, "An infrared silicon optical modulator of metal–oxide–semiconductor capacitor based on accumulation-carrier absorption," *Japanese Journal of Applied Physics*, vol. 48, no. 4, p. 04C107, 2009.

[44] D. Andrea, "CMOS photonics today and tomorrow enabling technology," *Proceedigns in OFC*, 2009.

[45] F. Gardes, G. Reed, N. Emerson, and C. Png, "A sub-micron depletion-type photonic modulator in silicon on insulator," *Optics Express*, vol. 13, no. 22, pp. 8845–8854, 2005.

[46] A. Liu, L. Liao, D. Rubin, H. Nguyen, B. Ciftcioglu, Y. Chetrit, N. Izhaky, and M. Paniccia, "High-speed optical modulation based on carrier depletion in a silicon waveguide," *Optics Express*, vol. 15, no. 2, p. 660668, 2007. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-15-2-660

[47] D. Thomson, F. Gardes, J.-M. Fedeli, S. Zlatanovic, Y. Hu, B. Kuo, E. Myslivets, N. Alic, S. Radic, G. Mashanovich, and G. Reed, "50-gb/s silicon optical modulator," *IEEE Photonics Technology letters*, vol. 24, no. 4, pp. 234–236, 2012.

[48] D. Gill, M. Rasras, K.-Y. Tu, Y.-K. Chen, A. White, S. Patel, D. Carothers, A. Pomerene, R. Kamocsai, C. Hill, and J. Beattie, "Internal bandwidth equalization in a CMOS-Compatible si-ring modulator," *IEEE Photonics Technology letters*, vol. 21, no. 4, pp. 200–202, 2009.

[49] Keigo Iizuka, *Elements of Photonics*. John Wiley & Sons, 2002.

[50] K. Iwamoto and I. Kamata, "Liquid-level sensor with optical fibers," *Appl. Opt.*, vol. 31, no. 1, p. 5154, Jan. 1992.

[51] M. Marcuse, "Mode conversion caused by surface imperfections of a dielectric slab waveguide," *Bell Syst. Technology Journal*, pp. 3187–3215, 1969.

[52] C. G. Poulton, C. Koos, M.f. Fujii, A. Pfrang, T. Schimmel, J. Leuthold, and W. Freude, "Radiation modes and roughness loss in high index-contrast waveguides," *IEEE Journal of slelected topics in quantum electronics*, vol. 12, no. 6, pp. 1306–1321, 2006.

[53] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. John Wiley & Sons, 2007.

[54] K. Okamoto, *Fundamentals of Optical Waveguides*. Academic Press, 2 Edition, 2005.

[55] D. V. Thourhout, G. Roelkens, R. Baets, W. Bogaerts, J. Brouckaert, P. Debackere, P. Dumon, S. Scheerlinck, J. Schrauwen, D. Taillaert, F. V. Laere, and J. V. Campenhout, "Coupling mechanisms for a heterogeneous silicon nanowire platform," *Semiconductor Science and Technology*, vol. 23, no. 6, p. 064004, 2008.

[56] A.R Hanim, H. Hazura, B. Mardiana, and Menon, "Free carrier absorption loss on p-i-n and n-p-n at 1.3 um and 1.55 um wavelegth," *IEEE International Conference on Semiconductor Electronics (ICSE)*, pp. 351 –354, 2010.

[57] E. W. Max Born, Max Born, *Principles of Optics;Electromagnetic Theory of Propagation, Interference and Diffraction of Ligh*. Cambridge University Press, 7th Edition, 1999.

[58] S. Satoh, K. Susa, and I. Matsuyama, "Simple method of measuring scattering losses in optical fibers," *Appl. Opt.*, vol. 38, no. 34, p. 70807084, Dec. 1999.

[59] P. K. Tien, "Light waves in thin films and integrated optics," *Appl. Opt.*, vol. 10, no. 11, p. 23952413, 1971.

[60] I. P. Kaminow and L. W. Stulz, "Loss in cleaved ti-diffused LiNbO3 waveguides," *Applied Physics Letters*, vol. 33, no. 1, pp. 62–64, 1978.

[61] Y. Vlasov and S. McNab, "Losses in single-mode silicon-on-insulator strip waveguides and bends," *Opt. Express*, vol. 12, no. 8, p. 16221631, Apr. 2004. [Online]. Available: http://dx.doi.org/10.1364/OPEX.12.001622

[62] H. P. Weber, F. A. Dunn, and W. N. Leibolt, "Loss measurements in thin-film optical waveguides," *Appl. Opt.*, vol. 12, no. 4, p. 755757, Apr. 1973.

[63] Y. H. Won, P. C. Jaussaud, and G. H. Chartier, "Three-prism loss measurements of optical waveguides," *Applied Physics Letters*, vol. 37, no. 3, pp. 269–271, 1980.

[64] R. Regener and W. Sohler, "Loss in low-finesse Ti:LiNbO3 optical waveguide resonators," *Applied Physics B*, vol. 36, no. 3, pp. 143–147, 1985.

[65] S. Taebi, M. Khorasaninejad, and S. S. Saini, "Modified fabry-perot interferometric method for waveguide loss measurement," *Appl. Opt.*, vol. 47, no. 35, p. 66256630, Dec. 2008.

[66] Y. Okamura, S. Sato, and S. Yamamoto, "Simple method of measuring propagation properties of integrated optical waveguides: an improvement," *Appl. Opt.*, vol. 24, no. 1, p. 5760, Jan. 1985.

[67] K. H. Haegele and R. Ulrich, "Pyroelectric loss measurement in LiNbO3:Ti guides," *Opt. Lett.*, vol. 4, no. 2, p. 6062, Feb. 1979.

[68] A. M. Glass, I. P. Kaminow, A. A. Ballman, and D. H. Olson, "Absorption loss and photorefractive-index changes in Ti:LiNbO3crystals and waveguides," *Appl. Opt.*, vol. 19, no. 2, p. 276281, Jan. 1980.

[69] M. Marcuse, "Radiation loss of the dominant mode in round dielectric slab waveguide," *Bell Syst. Technology Journal*, vol. 49, pp. 1665–1693, 1970.

[70] A. W. Snyder, W. Allan, R. A. Sammut, and A. Rowland, "Radiation modes of optical waveguides," *Electronics Letters*, vol. 15, no. 1, pp. 4–5, 1979.

[71] K. P. Yap, A. Delge, J. Lapointez, B. Lamontagne, J. H. Schmid, P. Waldron, B. A. Syrett, and S.Jan, "Correlation of scattering loss, sidewall roughness and waveguide width in silicon-on-insulator (SOI) ridge waveguides," *Journal of Lightwave Technology*, vol. 27, no. 18, pp. 3999–4008, 2009.

[72] T. Barwicz and H. A. Haus, "Three-dimensional analysis of scattering losses due to sidewall roughness in microphotonic waveguides," *J. Lightwave Technol.*, vol. 23, no. 9, p. 2719, Sep. 2005.

[73] C. Ciminelli, F. DellOlio, V. Passaro, and M. Armenise, "Fully three-dimensional accurate modeling of scattering loss in optical waveguides," *Optical and Quantum Electronics*, vol. 41, no. 4, pp. 285–298, 2009.

[74] A. Snyder and . Love, *Optical Waveguide Theory*. Springer, 1984.

[75] F. Ladouceur, J. D. Love, and T. J. Senden, "Measurement of surface roughness in buried channel waveguides," *Electronics Letters*, vol. 28, no. 14, pp. 1321–1322(1), Jul. 1992.

[76] M. Gupta, *The Handbook of Photonics*. CRC Press, 1996.

[77] M. Kuznetsov and H. Haus, "Radiation loss in dielectric waveguide structures by the volume current method," *Quantum Electronics, IEEE Journal of*, vol. 19, no. 10, pp. 1505–1514, 1983.

[78] A. Matsutani, F. Koyama, and K. Iga, "Three-dimensional electron probe roughness analysis of InP sidewalls processed by reactive ion beam etching," *Applied Physics Letters*, vol. 66, no. 1, pp. 64–66, 1995. [Online]. Available: http://link.aip.org/link/?APL/66/64/1

[79] Y. Martin and H. K. Wickramasinghe, "Method for imaging sidewalls by atomic force microscopy," *Applied Physics Letters*, vol. 64, no. 19, pp. 2498–2500, 1994.

[80] K. HosomiI, M. Shirai, K. Hiruma, J. Shigeta, and T. Katsuyama, "AFM characterization of GaAs/AlGaAs waveguides," *IEICE Transactions on Electornics*, vol. E79-C, no. 11, pp. 1579–1585, 1996.

[81] J. H. Jang, W. Zhao, J. W. Bae, D. Selvanathan, S. L. Rommel, I. Adesida, A. Lepore, M. Kwakernaak, and J. H. Abeles, "Direct measurement of nanoscale sidewall roughness of optical waveguides using an atomic force microscope," *Applied Physics Letters*, vol. 83, no. 20, pp. 4116–4118, 2003.

[82] S. Johnson, M. Povinelli, M. Soljai, A. Karalis, S. Jacobs, and J. Joannopoulos, "Roughness losses and volume-current methods in photonic-crystal waveguides," *Applied Physics B*, vol. 81, no. 2-3, pp. 283–293, 2005.

[83] M. S. A. K. S. J. J. J. S.G. Johnson, M.L. Povinelli, "Roughness losses and volume-current methods in photonic-crystal waveguides," *IEEE Photonics Technology*, vol. 81, p. 283293, 2005.

[84] C. Westbrook, C. Kaut Roth, and J. Talbot, *MRI in Practice*. John Wiley & Sons, 2011.

[85] W P Dong and K J Stout, "Two-dimensional fast fourier transform and power spectrum for surface roughness in three dimensions," *Journal of Engineering Manufacture*, vol. 209, pp. 381–391, 1995.

[86] E. Z. Liangn, C. J. Huang, and C. F. Li, "Use of SiO2 nanoparticles as etch mask to generate si nanorods by reactive ion etch," *Journal of Vacuum Science Technology*, vol. 24, pp. 599–603, 2006.

[87] P. Dumon, W. Bogaerts, V. Wiaux, J. Wouters, S. Beckx, J. Van Campenhout, D. Taillaert, B. Luyssaert, P. Bienstman, D. Van Thourhout, and R. Baets, "Low-loss SOI photonic wires and ring resonators fabricated with deep UV lithography," *Photonics Technology Letters, IEEE*, vol. 16, no. 5, p. 13281330, May 2004.

[88] Karl S. Kunz and Raymond J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*. CRC Press, 1993.

[89] K. K. Lee, D. R. Lim, L. C. Kimerling, J. Shin, and F. Cerrina, "Fabrication of ultralow-loss Si/SiO[sub 2] waveguides by roughness reduction," *Optics Letters*, vol. 26, no. 23, p. 18881890, 2001.

[90] E. Jaberansary, T. Masaud, M. Nedeljkovic, M. Milosevic, G. Mashanovich, and H. Chong, "Scattering loss estimation using 2D fourier analysis and modelling of sidewall roughness on optical waveguides," *Photonics Journal, IEEE*, vol. PP, no. 99, pp. 1–1, 2013.

[91] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia, "A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor," *Nature*, vol. 427, no. 6975, p. 615618, 2004.

[92] A. Huang and T. Pinguet, "A 10Gb/s photonic modulator and WDM MUX/DEMUX integrated with electronics in 0.13/spl mu/m SOI CMOS," *Solid-State Circuits Conference, IEEE International*, pp. 922 –929, 2006.

[93] J. S. R. A. Soref and K. Petermann, "Large single mode rib waveguides in gesi-si and si-on-sio2," *IEEE Journal of Quantum Electron*, vol. 27, p. 19711974, 1991.

[94] X. Xu, S. Chen, J. Yu, and X. Tu, "An investigation of the mode characteristics of SOI submicron rib waveguides using the film mode matching method," *Journal of Optics A: Pure and Applied Optics*, vol. 11, no. 1, p. 015508, 2009.

[95] Y.-P. Hsieh, H.-Y. Chen, M.-Z. Lin, S.-C. Shiu, M. Hofmann, M.-Y. Chern, X. Jia, Y.-J. Yang, H.-J. Chang, H.-M. Huang, S.-C. Tseng, L.-C. Chen, K.-H. Chen, C.-F. Lin, C.-T. Liang, and Y.-F. Chen, "Electroluminescence from ZnO/Si-Nanotips light-emitting diodes," *Nano Letters*, vol. 9, no. 5, pp. 1839–1843, 2009.

[96] Ali W. Elshaari and Stefan F. Prebe, "10 gb/s broadband silicon electro-optic absorption modulator," *Optics Communications*, vol. 283, p. 2829 2834, 2010.

[97] Dieter K. Schroder, *Semiconductor Material and Device Characterization*, 3rd ed. Wiley-Blackwell, 2006.

[98] Nargis Bano, "Fabrication and characterization of ZnO nanorods based intrinsic white light emitting diodes (LEDs)," 2011.

[99] D.-J. Lee, J.-Y. Kwon, S.-H. Kim, H.-M. Kim, and K.-B. Kim, "Effect of al distribution on carrier generation of atomic layer deposited al-doped ZnO films," *Journal of The Electrochemical Society*, vol. 158, no. 5, pp. D277–D281, 2011. [Online]. Available: http://jes.ecsdl.org/content/158/5/D277.abstract

[100] S. Kwon, S. Bang, S. Lee, S. Jeon, W. Jeong, H. Kim, S. C. Gong, H. J. Chang, H.-h. Park, and H. Jeon, "Characteristics of the ZnO thin film transistor by atomic layer deposition at various temperatures," *Semiconductor Science and Technology*, vol. 24, no. 3, p. 035015, 2009.

[101] C. A. Barrios, "Electrooptic modulation of multisilicon-on-insulator photonic wires," *J. Lightwave Technol.*, vol. 24, no. 5, p. 2146, 2006.

[102] C. A. Barrios, V. R. d. Almeida, and M. Lipson, "Low-power-consumption short-length and high-modulation-depth silicon electrooptic modulator," *J. Lightwave Technol.*, vol. 21, no. 4, p. 1089, 2003.

[103] S. Manipatruni, R. K. Dokania, B. Schmidt, N. Sherwood-Droz, C. B. Poitras, A. B. Apsel, and M. Lipson, "Wide temperature range operation of micrometer-scale silicon electro-optic modulators," *Opt. Lett.*, vol. 33, no. 19, p. 21852187, 2008.

[104] J. D. Lee, C. Y. Park, H. S. Kim, J. J. Lee, and Y.-G. Choo, "A study of conduction of ZnO film/p-si heterojunction fabricated by photoinduced electrodeposition under illumination," *Journal of Physics D: Applied Physics*, vol. 43, no. 36, p. 365403, 2010.

[105] X. Li, B. Zhang, X. Dong, Y. Zhang, X. Xia, W. Zhao, and G. Du, "Room temperature electroluminescence from ZnO/Si heterojunction devices grown by metalorganic chemical vapor deposition," *Journal of Luminescence*, vol. 129, no. 1, pp. 86 – 89, 2009.

[106] R. Soref, S. J. Emelett, and R. Buchwald, "Silicon photonic waveguides for mid- and long-wave infrared region," *Journal of Materials Science Materials in Electronics*, vol. 8, no. 840, pp. 1464–4258, 2007.

[107] S. Musa and M. N. O. Sadiku, "Analysis of rectangular coaxial lines," pp. 322–325.

[108] J. T. Robinson, K. Preston, O. Painter, and M. Lipson, "First-principle derivation of gain in high-index-contrast waveguides," *Opt. Express*, vol. 16, no. 21, p. 1665916669, 2008.

[109] D. Thomson, Y. Hu, G. Reed, and J.-M. Fedeli, "Low loss MMI couplers for high performance MZI modulators," *Photonics Technology Letters, IEEE*, vol. 22, no. 20, pp. 1485–1487, 2010.

[110] D. Marris-Morini, L. Vivien, J. M. Fdli, E. Cassan, P. Lyan, and S. Laval, "Low loss and high speed silicon optical modulator based on a lateral carrier depletion structure," *Opt. Express*, vol. 16, no. 1, p. 334339, 2008.

[111] L. Soldano and E. C. M. Pennings, "Optical multi-mode interference devices based on self-imaging: principles and applications," *Lightwave Technology, Journal of*, vol. 13, no. 4, pp. 615–627, 1995.

[112] F. Y. Gardes, D. J. Thomson, N. G. Emerson, and G. T. Reed, "40 gb/s silicon photonics modulator for TE and TM polarisations," *Opt. Express*, vol. 19, no. 12, p. 1180411814, 2011.

[113] T. B. M. E. J. H. C. G. R. G. M. M.M. Milosevic, M. Nedeljkovic, "Submicron silicon waveguides and optical splitters for mid-infrared applications," *Group IV Photonics (GFP), IEEE 9th International Conference on*, p. 1517, 2012.

[114] H. C. T.M. Ben Masaud, E. Jaberansary, "Silicon electro-optic switch based on n-zno/p-si heterojunction structure," *Group IV Photonics (GFP), IEEE 8th International Conference on*, p. 136138, 2011.

[115] S. S. O. C. R. G. D. B. H. C. T.M. Ben Masaud, E. Jaberansary, "Compact fabry-perot electro-optic switch based on n-zno/p-si heterojunction structure," *Nanotechnology (IEEE-NANO), 12th IEEE Conference on*, pp. 1–3, 2012.

[116] T. B. M. E. J. H. C. G. R. G. M. M.M. Milosevic, M. Nedeljkovic, "Silicon waveguides and devices for the mid-infrared," *Applied Physics Letters*, vol. 101, no. 12, p. 121105, 2012. [Online]. Available: http://dx.doi.org/10.1063/1.4753948

[117] H. M. H. C. D. M. B. F. Djidjeli, E. Jaberansary, "Broadband plasmonic couplers for light trapping and waveguiding," *Spiee International conference*, 2010.