Physical Layer Modeling and Optimization of Silicon Photonic Interconnection Networks

Meisam Bahadori

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2018

ABSTRACT

Physical Layer Modeling and Optimization of Silicon Photonic Interconnection Networks

Meisam Bahadori

The progressive blooming of silicon photonics technology (SiP) has indicated that optical interconnects may substitute the electrical wires for data movement over short distances in the future. Silicon Photonics platform has been the subject of intensive research for more than a decade now [175] and its prospects continue to emerge as it enjoys the maturity of CMOS manufacturing industry [5]. SiP foundries all over the world [102] and particularly in the US (AIM Photonics) have been developing reliable photonic design kits (PDKs) that include fundamental SiP building blocks such as wavelength selective modulators and tunable filters. Microring resonators (MRR) are hailed as the most compact devices that can perform both modulation and demodulation in a wavelength division multiplexed (WDM) transceiver design [14, 106]. Although the use of WDM can reduce the number of fibers carrying data, it also makes the design of transceivers challenging [97]. It is probably acceptable to achieve compactness at the expense of somewhat higher transceiver cost and power consumption. Nevertheless, these two metrics should remain close to their roadmap values [220] for Datacom applications. An increase of an order of magnitude is clearly not acceptable. For example costs relative to bandwidth for an optical link in a data center interconnect will have to decrease from the current $\sim$\$5/Gbps down to $<$\$1/Gbps [8, 1]. Additionally, the transceiver itself must remain compact.

The optical properties of SiP devices are subject to various design considerations, operation conditions, and optimization procedures. In this thesis, the general goal is to develop mathematical models that can accurately describe the thermo-optical and electro-optical behavior of individual SiP devices and then use these models to perform optimization on

the parameters of such devices to maximize the capabilities of photonic links or photonic switch fabrics for datacom applications.

In Chapter 1, *Introduction*, we first provide an overview of the current state of the optical transceivers for data centers and datacom applications. Four main categories for optical interfaces (Pluggable transceivers, On-board optics, Co-packaged optics, monolithic integration) are briefly discussed. The structure of a silicon photonic link is also briefly introduced. Then the direction is shifted towards optical switching technologies where various technologies such as free space MEMS, liquid crystal on silicon (LCOS), SOA-based switches, and silicon-based switches are explored.

In Chapter 2, *Silicon Photonic Waveguides*, we present an extensive study of the silicon-on-insulator (SOI) waveguides that are the basic building blocks of all of the SiP devices. The dispersion of Si and SiO2 is modeled with Sellmiere equation for the wavelength range 1500−1600 nm and then is used to calculate the TE and TM modes of a 2D slab waveguide. There are two reasons that 2D waveguides are studied: first, the modes of these waveguides have closed form solutions and the modes of 3D waveguides can be approximated from 2D waveguides based on the effective index method [202]. Second, when the coupling of waveguides is studied and the concept of curvature function of coupling is developed, the coupled modes of 2D waveguides are used to show that this approach has some inherent small error due to the discretization of the nonuniform coupling. This chapter finishes by describing the coefficients of the sensitivity of optical modes of the waveguides to the geometrical and material parameters. Perturbation theory is briefly presented as a way to analytically examine the impact of small perturbations on the effective index of the modes.

In Chapter 3, *Compact Modeling Approach*, the concept of scattering matrix of a multi-port silicon photonic device is presented. The elements of the S-matrix are complex numbers that describe the amplitude and phase relationships of the optical modes in the input and output ports. Based on the scattering matrix modeling of silicon photonics devices,

two methods of solving photonic circuits are developed: the first one is based on the itera-tion for linear circuits. The second approach is based on the construction of an equivalent signal flow graph (SFG) for the circuit. We show that the SFG approach is very efficient for circuits involving microring resonator structures. Not only SFG can provide the solution for the transmission, it also provides the signal paths and the closed-form solution based on the Mason's graph formula. We also show how the SFG method can be utilized to formulate the backscattering effects inside a ring resonator.

In Chapter 4, *Scalability of Silicon Photonic Switch Fabrics*, we develop the models for electro-optic Mach-Zehnder switch elements (2×2). For the electro-optic properties, the empirical Soref's equations [174] are used to characterize how the loss and index of silicon changes when the charge carrier density is changed. We then use our photonic circuit solver based on the iteration method to find accurate result of light propagation in large-scale switch topologies (e.g. 4×4, and 8×8). The concept of advanced path mapping based on physical layer evaluation of the switch fabric is introduced and used to develop the optimum routing tables for 4×4 and 8×8 Benes switch topologies.

In Chapter 5, *Design space of Microring Resonators*, we introduce the concept of cur-vature function of coupling to mathematically characterize the coupling coefficient of a ring resonator to a waveguide as a function of the geometrical parameters (ring radius, coupling gap, width and height of waveguides) and the wavelength. Extensive 2D and 3D FDTD simulations are carried out to validate our modeling approach. Experimental demonstration are also used to not only further validate our modeling of coupling, but also to extract an empirical power-law model for the bending loss of the ring resonators as a function the radius. By combining these models, we for the first time present a full characterization of the design space of microring resonators. Moreover, the value of this discussion will be further apparent when the scalability of a silicon photonic link is stud-ied. We will show that the FSR of the rings determines the optical bandwidth but it also impacts the properties of the ring resonators.

In Chapter 6, *Thermo-optic Efficiency of Microheaters*, we develop analytical models for the thermo-optic properties of SiP waveguides. For the thermo-optic properties, the concept of thermal impulse response is mathematically developed for integrated micro-heaters [22]. The thermal impulse response is a key function that determines the tradeoff between heating efficiency and heating speed (thermal bandwidth), as well as allows us to predict the pulse-width-modulation (PWM) optical response of the heater-waveguide system. One of the motivations behind this study was to find the highest possible efficiency for thermal tuning of microring resonators to use it in the evaluation of the energy consumption of a photonic link. The results indicate 2 nm/mW which is in agreement with the trends that we see in the literature.

In Chapter 7, *Crosstalk Penalty*, we theoretically and experimentally investigate the optical crosstalk effects in microring-based silicon photonic interconnects. Both inter-channel crosstalk and intra-channel crosstalk are investigated and approximate equations are developed for their corresponding power penalties. Inclusion of the inter-channel crosstalk is an important part of our final analysis of a silicon photonic link.

In Chapter 8, *Scalability of Silicon Photonic Links*, we present the analysis of a WDM silicon photonics point-to-point link based on microring modulators and microring wave-length filters. Our approach is based on the power penalty analysis of non-return-to-zero (NRZ) signals and Gaussian noise statistics [21, 14]. All the necessary equations for the optical power penalty calculations are presented for microring modulators and filters. The first part of the analysis is based on various ideal assumptions which lead to a maximum capacity of 2.1 Tb/s for the link. The second part of the analysis is carried out with more realistic assumptions on the photonic elements in the link, culminating in a maximum throughput of 800 Gb/s. We also provide estimations of the energy/bit metric of such links based on the optimized models of electronic circuits in 65 nm CMOS technology.

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank Professor *Keren Bergman* for providing me the opportunity to join her research group at Columbia University and constantly supporting me through four years of study and research. When I started my Ph.D. in 2014, the Lightwave Research Laboratory (LRL) was beginning to get involved with the newly established American Institute for Manufacturing Photonics Integrated Circuits (AIM Photonics). Through the years, our collaboration with AIM Photonics has definitely grown in the form of various projects that we are participating in such as designing and fabricating silicon photonic switches, developing functional testing procedures for wafer-scale testing of silicon photonic devices, and so on. Clearly, having access to a foundry is crucial in developing new ideas, devices, and systems based on silicon photonics technology. In July 2017, I had the opportunity to attend a one-week workshop at Massachusetts Institute of Technology (MIT) on the matter of design and fabrication of silicon photonic chips, and where AIM Photonics is headed in the future. I am very grateful for having such opportunity to learn and gather a working knowledge of designing silicon photonic chips.

I would like to also thank Dr. *Sébastien Rumley* for his mentorship throughout my Ph.D. work at Columbia University. Through numerous discussions and brainstormings on the topics of performance of silicon photonic interconnects and how to properly include physical layer models into the optimization of bandwidth in a large scale network, I acquired a great deal of knowledge of the current state of optical networks and the challenges and opportunities that optical communication is facing.

I would like to also extend my gratitude to Dr. *Robert Polster* and Dr. *Qixiang Cheng* for many fruitful discussions about analyzing, designing and testing silicon photonics devices and circuits. Dr. *Cheng* is truly an expert in integrated photonic switches and it was through his guidance that I looked into the impact of physical layer on the routing performance of large scale silicon photonic switches. What we realized in this study was that it is possible to take advantage of the element of repetition to optimize the path mapping strategy of a silicon photonic switch fabric.

Last but not least, my Ph.D. work would have not been possible without the help, support and collaborations of my dear friend, my office mate and my colleague, *Alexander Gazman*. Spending four years in the same office gave us enough time to ponder all the aspects of silicon photonics and talk about new papers that come out every day. Together, we developed new understandings of thermo-optic effects and pulse-width-modulation control of integrated heaters, as well as the significance of the photo-conductance effect. For all of this, I am very grateful.

I dedicate this thesis to my parents who raised me to become the person I am today.

Chapter 1

Introduction

The ever increasing growth of cloud-computing is driving the need for data compute and storage resources available at any time. Because the exact location where data is stored, computations are performed, and resources are located does not matter anymore, the trend has been to gather many servers within the same location to benefit from economies of scale [68], resulting in very large data centers with hundreds of thousands of servers. The emergence of these large scale data centers has given rise to important engineering requirements and maintenance, including the need to keep the servers up and running with minimal human intervention, the need to prevent irrecoverable data losses, and the need to adequately exhaust the heat generated by hundreds of thousands of servers.

These huge data centers also require tailored data center interconnects (DCI) that can ensure proper connectivity and stable communication of the servers with each other and with the Internet [156, 171, 193]. Because the reach of a standard 10 Gbps interconnect over copper cable is limited to ∼10 m (see Fig. 1.1), large scale data centers adopted optical transmission technologies during the transition from 1 Gbps to 10 Gbps link data rate between 2007 and 2010 [216]. In 2007, Google introduced optical communication in its data centers in the form of 10 Gbps VCSEL and Multi-Mode fiber based SFP transceivers for reaches up to 200 m (Fig. 1.1). As the intensity of traffic generated by servers doubles every year, transitions from 10 Gbps to 40 Gbps, from 40 Gbps to 100 Gbps, and from 100 Gbps to even higher rates have been predicted early on [89]. In 2017, 40 Gbps Ethernet based interconnects have already been deployed in production data centers. 100 Gbps links have been commercially available since 2014 and are currently installed in produc-
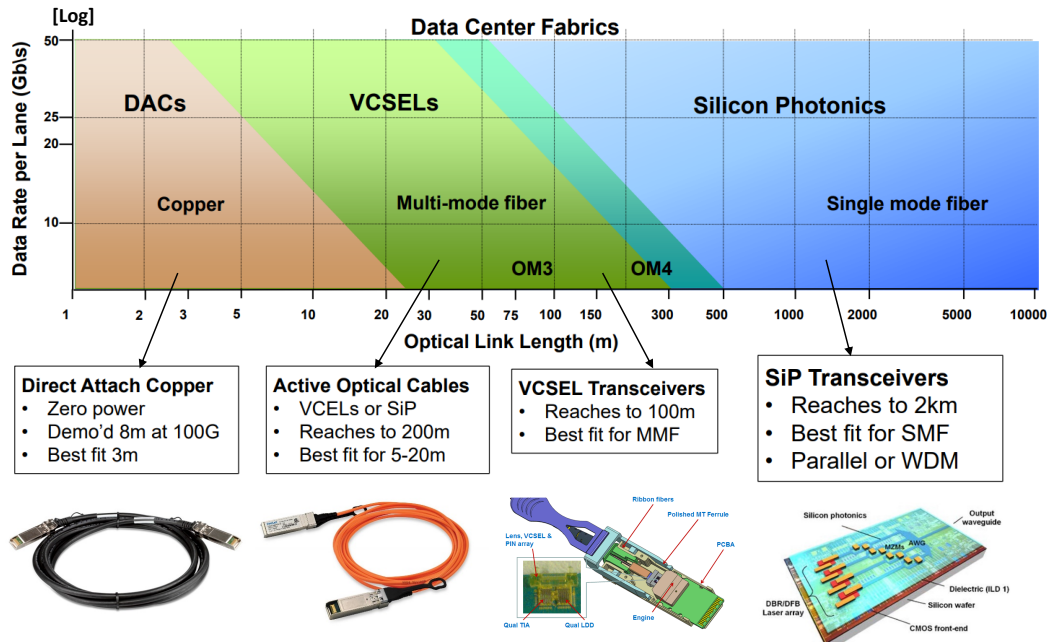
Figure 1.1: Plot of data rate (Gb/s) per lane vs. reach (m) for different data center interconnect (DCI) technologies [228] (*pictures taken from Internet*).

tion data centers. 400 Gbps equipment is expected to emerge in the near future [216]. Data center servers will require even higher bitrates to connect their compute capabilities with their peers and the external world [159], especially with the adoption of machine-learning algorithms. Application examples of these algorithms include voice assistants such as Apple *Siri*, Google *Assistant*, and Amazon *Alexa*, as well as face identification applications in social networks. In addition to expanded bandwidths, optical equipment with improved energy efficiency and compactness is also expected [43].

Data center operators are eager to see interconnects able to handle more than one hundred thousand of servers at Terabit/s bitrates. However, these operators require future high capacity DCI to comply with strict dollar and power budgets. Specifically, costs relative to bandwidth for a link will have to decrease from the current ~$5-per-Gbps down to <$1-per-Gbps to maintain the cost of the entire interconnect stable while performance is scaled up. Data center power consumption is also a matter of great concern. Not only is delivering more than 100 MW challenging for data center operators (in particular in terms

Table 1.1: Breakdown of roadmap targets for the power and cost efficiency of interconnects in data centers (Year 2016) [1].

Interconnect	Distance	Energy per bit	Cost (per Gbps)
Inter-DC	1–100 km	<10 pJ/bit	<$1000
Rack-to-rack	1 m–2 km	< 1 pJ/bit	<$100
Board-to-board	0.3–1 m	< 1 pJ/bit	<$10
Module-to-module	5–30 cm	< 0.5 pJ/bit	<$5
Chip-to-chip	1–5 cm	< 0.1 pJ/bit	<$1
Core-to-core	<1 cm	< 0.01 pJ/bit	<$0.01

of grid accessibility and reliability) [159], operators must be responsive to increased public concerns about climate change and environmental issues and the ultimate ecological footprint of data center activities. Future DCI will thus be expected to carry more data while consuming less – the energy dissipated while transmitting a single bit over a link will have to be reduced to ∼1 pJ from several tens of pJ today [1, 47, 159]. This requires better provisioning of the available communication bandwidth within a data center network [201] as well as improvements to the established technologies that enable photonic communication links in the network.

1.1   Optics for Datacom

The wide adoption of optical data movement in data centers is contingent upon transceivers that enable efficient electrical-to-optical and optical-to-electrical conversions at each node of the network, as well as optical switches that allow for redirection of optical signal flow from one node to another. Although optical transceivers have already established the path for commercialization, optical switching remains primarily a research topic. The emerging technologies that will satisfy the required performance for DCI need to first prove their value in terms of the energy efficiency (measured in pJ/bit or mW/Gbps) and the cost per bit ($/Gbps) and meet the roadmap targets. The breakdown of roadmap

targets for these two parameters for DCI are summarized in table 1.1 [1].

1.2   Optical Transceivers

Optical transceivers for datacom need to satisfy different performance requirements compared to their telecom counterparts [61, 47]. Traditional telecom systems put a high emphasis at the end points of the links (e.g. advanced modulation formats [204] at the transmitter, and adoption of DSP at the receiver) to maximize the spectral efficiency (in terms of bits/hz) of long-reach links where fiber is costly. In a data center, however, shortreach fiber is rather abundant at a low cost and spectral efficiency is sacrificed (e.g. adoption of simplest optical amplitude modulation formats) for lower power and cost effective transceivers that can provide low latency. Compatibility and interoperability of different vendors' proprietary equipment in a data center is also an important factor to keep the cost of the interconnect at bay [47]. The evolution of optical technologies for data centers can be categorized based on the way the optical interfaces are implemented as shown in Fig. 1.2. The following briefly summarizes these optical interfaces:

Pluggable Transceivers

The first type of optical interface is active optical cables (AOC) and pluggable transceivers [216, 47] (Fig. 1.2a). AOCs are used for short-range ($\sim$5–20 m) data communication and interconnect applications while pluggable transceivers are commonly designed for a longer reach ($\sim$100 m). AOC/pluggable units consist of fiber optic transceiver and control modules. Compared with direct attach copper cable (DAC) for data transmission, AOC/pluggable transceivers provide more advantages, such as lighter weight, lower power consumption, lower interconnection loss, and electromagnetic interference immunity. Figure 1.3 plots the evolution of pluggable transceivers over time [216]. The early generation of these transceivers (2007) was designed for 10 Gb/s two-level amplitude sig-
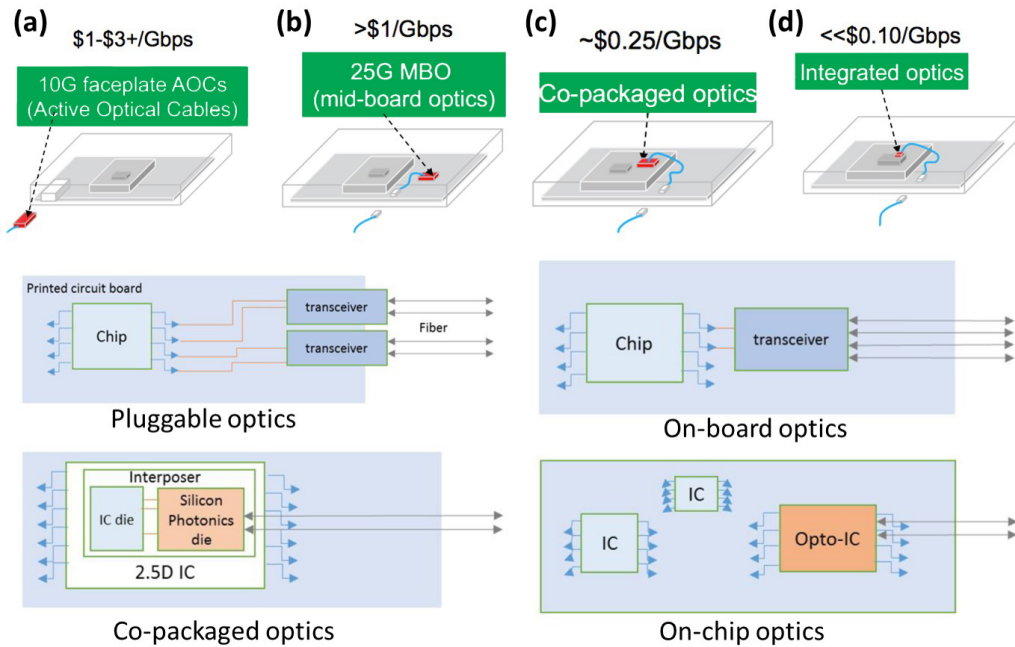
Figure 1.2: (a) Optical interface for AOCs and pluggable transceivers. (b) Optical interface for board-mounted assembly. (c) Co-packaged optics with electronics (2.5D integration on an interposer). (d) Monolithic integration of optics and electronics. (bottom) schematics of electronics and optics for each optical interface [78].

naling (on-off keying) and direct detection, utilizing only one optical carrier (850 nm for a mulit-mode fiber (MMF) and 1300 nm for a single mode fiber (SMF)). The second generation (2010) was designed for 40 Gb/s signaling in the form of QSFP (Quad Small Form-factor Pluggable). By utilizing space division multiplexing for VCSEL/MMF and coarse WDM technology for SMF, this generation introduced four optical lanes each operating at 10 Gb/s. The third generation (2014) is designed for 100 Gb/s operation in a QSFP28 packaging. It has a 4-lane, retimed 25Gb/s I/O electrical interface and supports up to 3.5 W power dissipation (35 pJ/bit) with standard cooling. The emerging generation of pluggable transceivers (2017) for data centers is envisioned to operate at 400 Gb/s by utilizing four-level amplitude signaling (PAM4) with 10 W of power dissipation (25 pJ/bit).
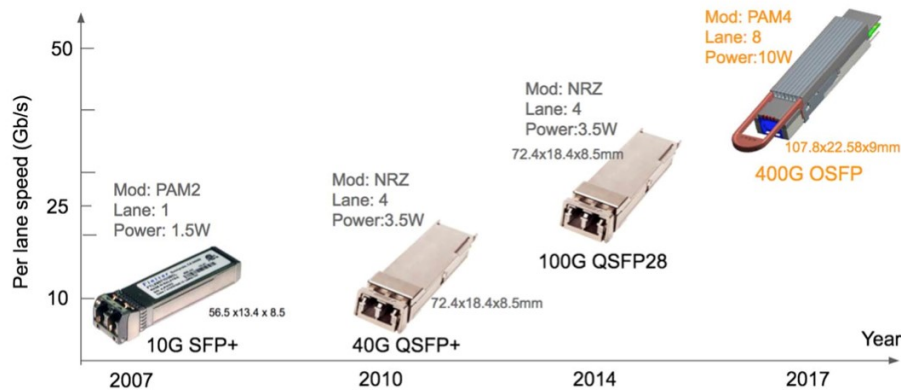
Figure 1.3: Evolution of optical pluggable transceivers over time [216] in terms of speed, power dissipation handling, and number of data lanes.

On-board Optics

The next step from pluggable transceivers is board-mounted optical assembly (Fig. 1.2b). Mid-board optics are not pluggable; they are mounted on the host printed circuit board (PCB) and are commonly used today on core routers and some switches. For this type of interface very short host PCB traces allow for low power dissipation and transfer of high-speed electrical signaling (25 Gb/s) from the electronic IC to the optics. The consortium for on-board optics (COBO) [222] has been formed by Microsoft, Cisco, Intel, Finisar, Mellanox, and other vendors to develop specifications and standards for on-board optical modules that mainly target the reduction of power requirements in network switches and server adapters.

Co-packaged Optics

In order to further bring down the cost of optical transceivers and boost their reliability, higher level of integration is required. Figure 1.2c depicts the case in which the optical interface is co-packaged with the electronic integrated circuit (IC). Commonly used methods are wire-bonding and flip-chip bonding. This method of integration is referred to as the 2.5D integration [192] in which the optical die and electronic die are separately

mounted on a chip carrier or an interposer. In the co-packaged approach, the electronic IC is typically fabricated in an advanced CMOS node (e.g. 28 nm) through CMOS foundries while the optical chip (e.g. based on silicon photonics [185]) is fabricated in a different process through silicon photonic foundries [102]. In 2013, PETRA demonstrated a silicon optical interposer with 20 Gbps data links summing up to 30 Tbps/cm^2 of bandwidth density [190, 191].

Integrated Optics

Compared to the co-packaged optics, monolithic integration of electronics and photonics promises the most cost effective solution ($\ll$\$0.1/Gbps) due to the fact that both the CMOS and silicon-based photonic devices are fabricated simultaneously through the same process and there is no need for an additional interposer. Such monolithic integration has been demonstrated by IBM in 2012 [9] in 90 nm CMOS node and in 2015 by UC Berkeley [181] in a 45 nm CMOS SOI process where all the photonic elements were designed according to the native CMOS process rules (a "zero-change" design). The latter demonstration included 70 million transistors (processor and 1 MB of memory) and 850 photonic elements.

Anatomy of a silicon photonic link

A photonic link is by definition equipped with a transmit unit (Tx) and a receive unit (Rx). The link architecture based on an integrated photonic platform such as Silicon Photonics can be envisaged as the ones presented in Fig. 1.4 [14, 18]. The Tx requires input optical power in the form of serialized/combined optical wavelengths. Each wavelength provides one channel of optical data. One convenient solution is to use a multi-wavelength laser source such as a comb laser source [38]. The optical power is then injected into the Tx using vertical grating couplers or edge couplers. This process inflicts optical loss and is
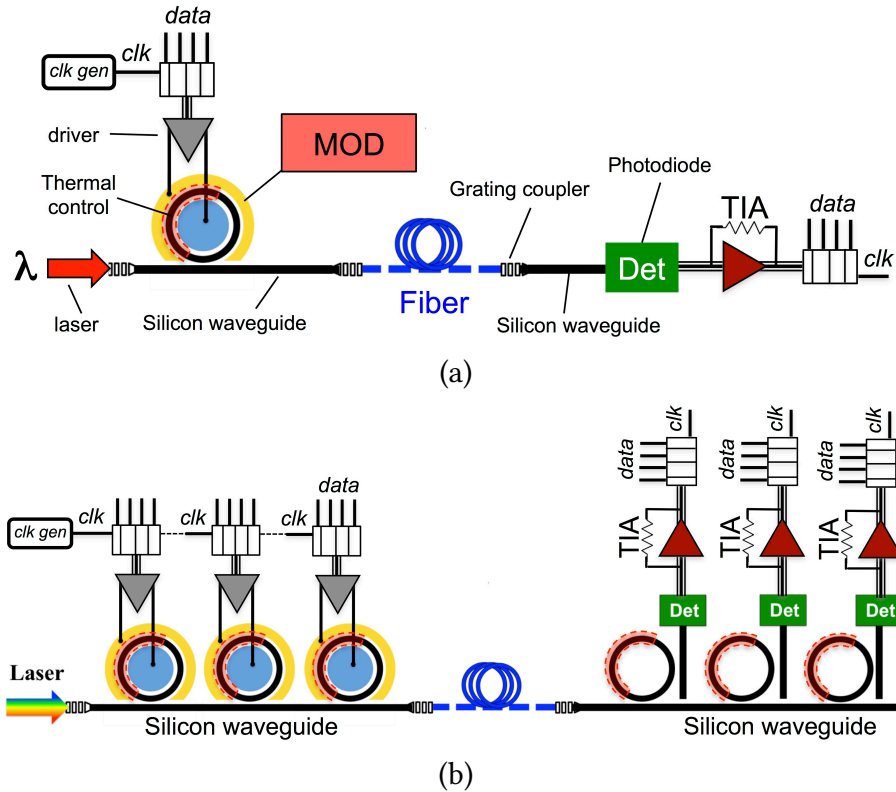
Figure 1.4: (a) Structure of a single-λ photonic link with a microring modulator at the transmitter and no spectral filtering at the receiver. (b) Anatomy of a WDM link based on microring resonators. Multiplexing of wavelengths is easily achieved due to the wavelength-selective nature of microrings [18]. Spectral filtering is required at the receiver to select individual channels.

unavoidable. Grating couplers with more than 75% coupling efficiency [195] and less than 4 dB loss [101] have already been demonstrated.

The architecture in Fig. 1.4b consists of cascaded microring modulators. Each modulator is designated to a specific wavelength through thermo-optic tuning of its resonance [22], while modulating light through fast charge-carrier-based electro-optic effects in silicon [155]. The output data streams of all modulators are combined and multiplexed inside the bus waveguide and extracted from the Tx unit into a carrier fiber through another optical coupler. As can be seen in Fig. 1.4b, multiplexing of optical channels in a transmitter architecture based on microring modulators is achieved at no additional cost because of the wavelength-selective property of such resonators.

The Rx unit is equipped with a wavelength demultiplexer since WDM scheme is used. Light from each channel, accompanied by optical crosstalk [15], is incident on a photodiode, which is then converted to electrical current. The electrical current is converted to electrical voltage signal and amplified and finally deserialized to recover each individual electrical data stream. Although inclusion of a clock-and-data recovery (CDR) module is a common choice for the Rx units, forwarded clock schemes have also been proposed for microring-based links [67].

To protect devices from reflections, or allow a fiber to be used in a bidirectional mode, isolators and circulators may be additionally inserted along the link. To compensate for low photodiode sensitivities, losses or impairments, a semiconductor optical amplifier (SOA) booster may be inserted before the photodiode.

1.3   Optical Switching

The need for high-capacity interconnects also imposes challenges in terms of switching. Currently, the majority of switching operations in data centers are being carried out by electronic packet routers. These routers, generally implemented as a custom-designed circuit (ASIC: application specific integrated circuit), receive streams of data packets over N input lanes. Among these packets, those routed to the same destination are marshaled and sent to one of the N output links. As data rates of transmission links connected to these routers increase, the switching capacity of the router must grow as well. A 32 port router accepting links at 100 Gbps must for instance be able to let up to 3.2 Tbps transit through it. A transition to 200 Gbps links obliges the router capacity to scale to 6.4 Tbps. However, with the observed end of Dennard scaling (MOSFET scaling [223]), which states that as transistors get smaller their power density stays constant (so that the power use stays in proportion with area), router chip scaling becomes more challenging. The proportion of data center power and cost budgets consumed by the electronic routers

and swtiches are thus expected to grow. To mitigate this trend, alternative switching techniques such as optical switching technologies have been investigated. The primary attraction of the optical switching is the elimination of the transceivers, allowing the data to proceed all optically through the switch fabric rather than requiring the energy and component costs of optical to electrical and then electrical to optical conversions. The use of optical switches, although very promising in terms of energy savings, requires modifications to the data center architecture. A primary challenge in the optically switched architecture is the lack of commercial or near-commercial random access optical memory and optical buffering which inevitably results in instances of packet drop due to traffic contention in the network. Optical switches thus cannot be considered as a one-to-one replacement for electronic packet switches.

Optical switching technologies

Various technologies can enable spatial or wavelength-selective optical switches, including micro-electromechanical systems (MEMS), liquid crystals on silicon (LCOS), semiconductor optical amplifiers (SOA), Mach-Zehnder interferometers (MZI) [108, 19] and micro-ring resonators (MRR) [130]. The choice of the optical switch for a particular application (e.g. in a data center) is ultimately driven by metrics such as cost-per-port and optical power penalties. The reconfiguration speed of the switch is often presented as an important metric. The following is a brief overview of optical switching technologies:

Free space MEMS optical switches

Free space MEMS-based optical switches are the most common and mature free space switching devices. The MEMS spatial switches are realized in both two-dimensional (2D) and three-dimensional (3D) configurations, as shown in Fig. 1.5 [46]. The 2D MEMS switches are implemented in the crossbar topology and operate digitally as the mirror position is bi-stable. 3D MEMS switches have been proposed to support very large-scale

optical cross-connect devices. This type of device is assembled by using 2D input and output fiber arrays with collimators. Two stages of independent 2D micro-mirror arrays are used to steer the optical beams in three dimensions, which requires micro-mirrors implemented with a two-axis tilting structure.
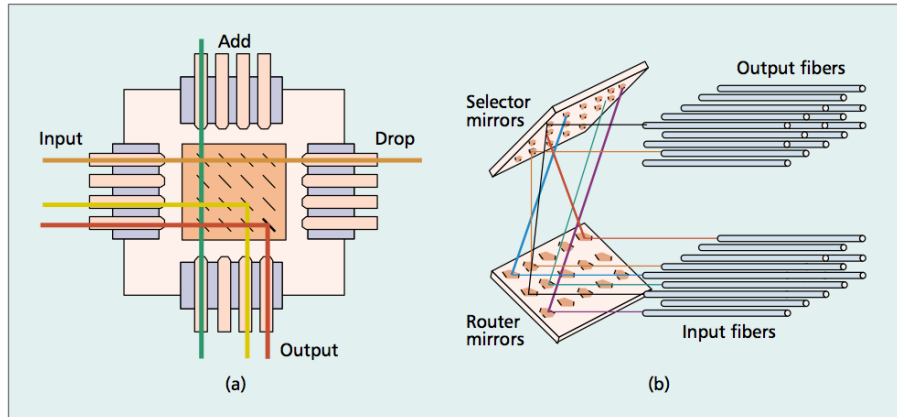
Figure 1.5: Crossbar structure of (a) 2D and (b) 3D free space MEMS [46].

LCOS-based optical switches

The light modulating properties of liquid crystal (LC) material has been explored for both amplitude modulation and phase modulation. The variation of optical properties, such as polarization and refractive index, can be adjusted by applying voltages across the LC material to change the certain relative orientation of molecules. Liquid crystal on silicon (LCOS) combines the light modulating feature of LC material and the advanced CMOS technology [25] (see Fig. 1.6).

SOA-based optical switches

Integrated switch fabrics based on semiconductor optical amplifiers (SOA) have been mainly implemented in the broadcast-and-select and wavelength-selective configurations [176]. Each input-to-output path can be gated by one SOA element that compensates for the optical losses while its high ON/OFF extinction ratio ensures excellent crosstalk sup-
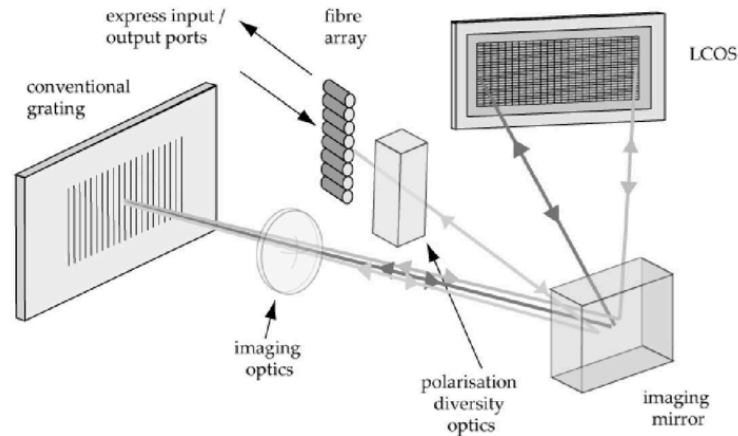
Figure 1.6: Schematic of an LCOS-based wavelength selective optical switch [25].

pression (Fig. 1.7). Broadcast-and-select topology scales with the square of the number of switch elements.
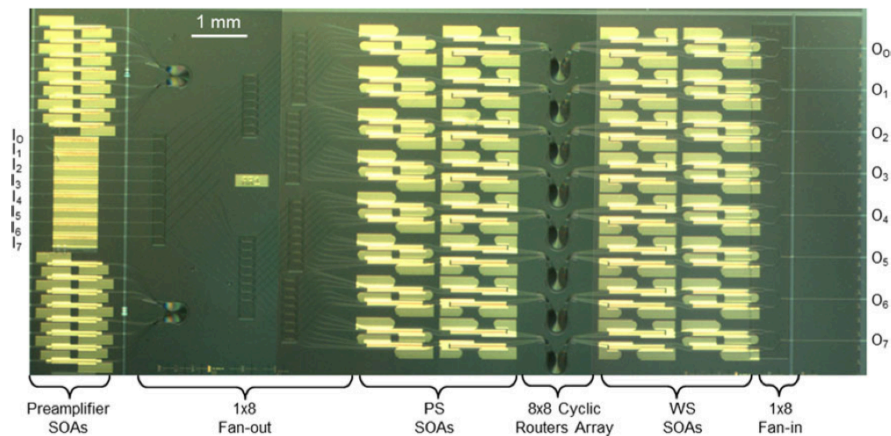
Figure 1.7: Image of an 8×8 SOA-based wavelength-selective integrated switch fabric [176].

Silicon-based optical switches

Large port count, i.e. large radix, switches in bulk optics based on MEMS are being commercialized already. However, to deploy optical switching at large scale, for example to enable bandwidth steering in a Dragonfly topology [201], more integrated, thus cheaper, optical switches are necessary. The silicon photonics platform has proven a worthy op-

tion in terms of performance, but overall in terms of integration, with a fast growth over the last decade [185]. Today, individual broadband silicon photonic switches based on Mach-Zehnder interferometers have been fabricated, tested, and shown to provide fast switching speeds (nano-second). Calo et al. [36] provided a detailed analysis of design and analysis of 2×2 silicon-based electro-optic Mach-Zehnder switch (MZS), concluding that an insertion loss of 1.1 dB, a crosstalk level of -15 dB with an optical bandwidth of 60 nm is achievable. In 2015, Dupuis et al. [59] from IBM revisited the concept of 2×2 MZS in silicon photonic platform (Fig. 1.8). Leveraging IBM's 90 nm CMOS platform, and optimizing their design for minimal loss and crosstalk, this group proved that insertion loss of 1 dB and crosstalk of -23 dB with an optical bandwidth 45 nm and 4 nsec switching time is possible. Later that year, the same group further introduced these 2×2 MZS as the enabling technologies for fast and scalable silicon photonic switches by demonstrating a 4×4 switch based on 2×2 switches [58]. Other researchers have strived to scale up the port count of larger numbers. For example, in 2016, Lu et al. [108] demonstrated a 16×16 non-blocking switch based on Benes topology with a total footprint of 40 mm^2 . The measured loss and crosstalk in the "cross" state of the 2×2 MZS were 0.1 dB and -30 dB over a wavelength range of 30nm. In the "bar" state the numbers degraded to 1 dB and -18 dB, respectively.

Microring resonators (MRRs) have also been investigated in the silicon photonic platform for the design and fabrication of wavelength selective silicon switches. Ring-based switch fabrics have mostly been exploited as cascading 1×2 or 2×2 switching elements in classic topologies, such as Cross-bar and Benes. Figure 1.9 is an example of an 8×8 Benes topology based on 2×2 MRRs. Each switch element has a bar state where both rings are dropping the signals and has a cross state in which both rings are detuning so that they are not dropping the optical signals that are passing by.

Silicon-based switch fabrics have also been designed and fabricated based on the vertical coupling of MEMS-actuated waveguides. Such a design has the potential to enable
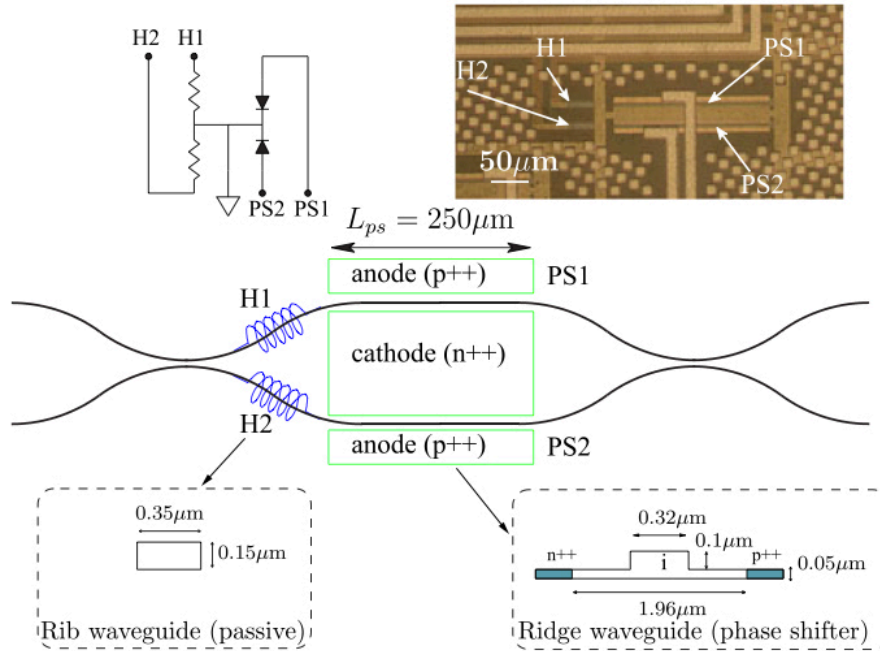
Figure 1.8: The 2×2 Mach-Zehnder switch design from IBM [59] showing the schematic of the switch with the electro-optic phase shifters and the thermo-optic phase shifters.
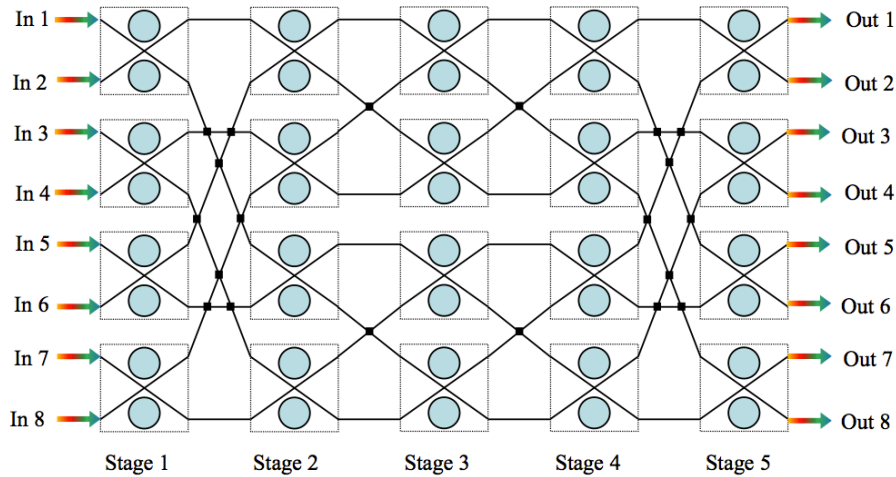
Figure 1.9: Schematic of an 8×8 microring based silicon photonic switch fabric [130]. The switch is based on Benes topology.

very low-loss crossbar switch topologies as demonstrated in Fig. 1.10a by UC Berkeley [165] for a 64×64 switch. Figure 1.10b shows one of the switching elements where the vertical coupling is in effect and the light is directed in the cross direction. Figure 1.10c
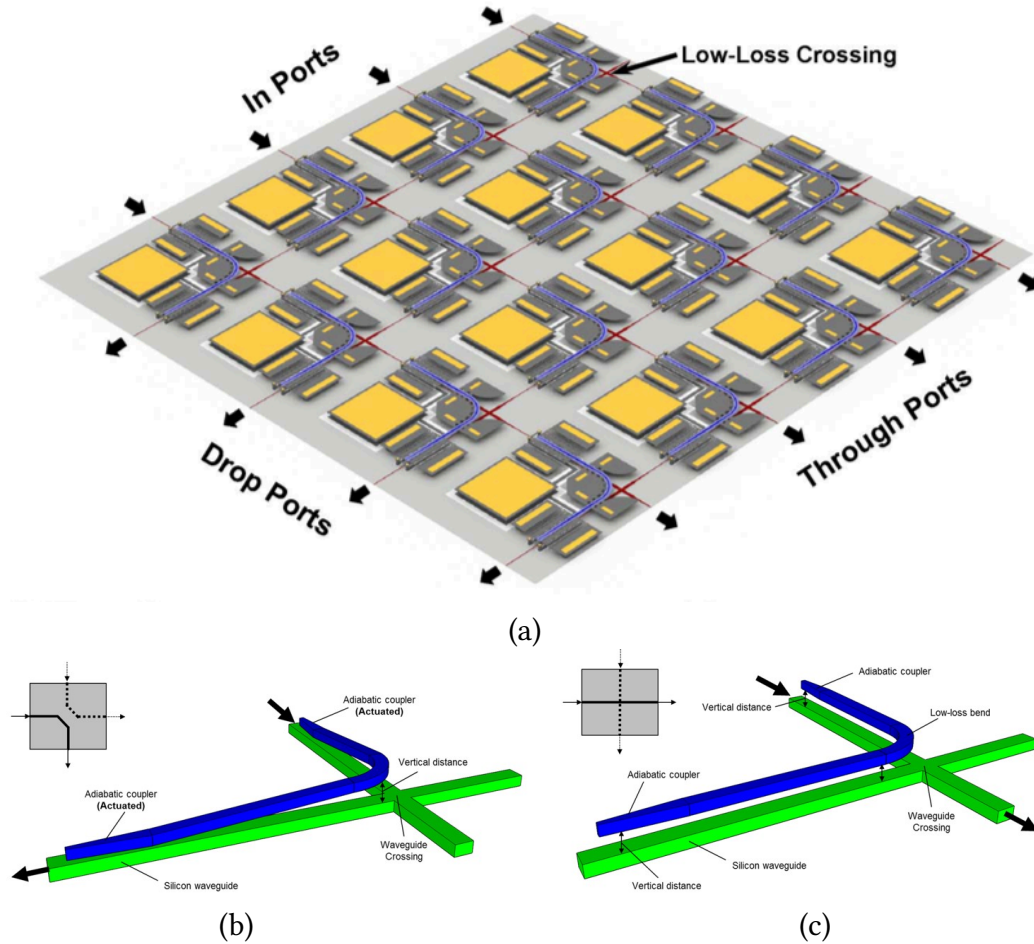
Figure 1.10: (a) 4×4 silicon-based MEMS switch based on vertically coupled waveguides [165]. (b) Schematic of the switch element as the cross connect. (c) Schematic of the switch element as the through connect.

shows the switch element where the vertical coupling is not activated and the input light goes through with very little coupling interference.

1.4   Focus of this thesis

Seeing that Silicon Photonics can provide optical interconnects in the form of optical links (transceivers) and photonic switch fabrics, this thesis aims at investigating the physical layer performance of these interconnects from the device level (e.g. optical modes, physical length, couplings) up to the interconnect level (e.g. system metrics such as loss,

crosstalk, spectral distortion of optical data).

For the photonic switch fabrics, the focus is on answering two important questions:

1. How does inserting a silicon photonic switch fabric in a network would impact the optical power budget requirements? To answer this, we build-up worst-case performance graphs for 4×4 and 8×8 switches with electro-optic Mach-Zehnder elements.

2. How should a control plane configure the photonic switch to allow specific routings of optical data while ensuring a uniform performance for all the input-output paths of the switch and avoiding worst-cases? To answer this, we propose a methodology to build up routing tables for 4×4 and 8×8 Benes switches. The answer to this questions is becoming progressively important as larger scale switches (16×16 [108] and 32×32 [149]) are being developed and demonstrated.

For a silicon photonic link, the focus is on finding two important *metrics*:

1. The maximum capacity that a link can offer based on microring resonators. In order to do so, we implement an optimization procedure based on all the possible impacts of the physical layer parameters on the power penalty of the link. To accomplish this, in various chapters of this thesis we explore the design space of microrings and investigate optical crosstalk.

2. Approximations for the energy consumption (pJ/bit) of the link.

For a network designer, the results of these two studies are the key to finding the right network architecture, i.e. the number of photonic links between nodes, and/or designing flexible networks by utilizing photonic switches [201] based on the traffic pattern and the required communication bandwidths of the network.

Throughout this thesis, a strong emphasis is put on the rigorous mathematical derivations to ensure that the results and conclusions are reasonable and sound.

Chapter 2

Silicon Photonics Waveguides

In this chapter we discuss the fundamental guiding structures in silicon photonics (SiP) platform. These are rectangular waveguides that can confine light within the sub-wavelength dimensions due to the high index contrast (HIC) that is provided in this platform [202]. The two main materials used are *Silicon* (Si) and *Silicon dioxide* (Silica - SiO_2) which typically exhibit some level of material dispersion (refractive index depends on wavelength) over the wavelength region of interest (typically from 1.2μm to 2μm).

The *linear* dispersive behavior is modeled by Lorentz harmonic oscillations. At λ = 1.55μm, the refractive index of silicon is about 3.4777 and the refractive index of silica is about 1.444. These numbers are based on the widely used Sellmier model for the material dispersion given by the following equation:

$$n^2 - 1 = \frac{A_1\ \lambda^2}{\lambda^2 - B_1^2} + \frac{A_2\ \lambda^2}{\lambda^2 - B_2^2} + \frac{A_3\ \lambda^2}{\lambda^2 - B_3^2} \qquad (2.1)$$

where λ is in μm units. For silicon material, the coefficients are A_1 = 10.6684293, A_2 = 0.0030434748, A_3 = 1.54133408, B_1 = 0.301516485 μm, B_2 = 1.13475115 μm, B_3 = 1104 μm. For the silica material, the coefficients are A_1 = 0.6961663, A_2 = 0.4079426, A_3 = 0.8974794, B_1 = 0.0684043 μm, B_2 = 0.1162414 μm, B_3 = 9.896161 μm [161]. These coefficients are also tabulated in table 2.1. Figure 2.1b and 2.1c show the material dispersion for silicon and silica in the optical wavelengths (1.25μm – 1.65μm), respectively. As can be seen on these plots, silicon exhibits 1.3% change which in most accurate designs cannot be ignored. On the contrary, silica shows only a change of 0.3% which is mostly neglected.

Table 2.1: Coefficients for Sellmier model of index of Si and SiO$_2$ around 1.55μm wavelength [161].

Parameter	Si	SiO$_2$
A_1	10.6684293	0.6961663
A_2	0.0030434748	0.4079426
A_3	1.54133408	0.8974794
B_1	0.301516485	0.0684043
B_2	1.13475115	0.1162414
B_3	1104	9.896161
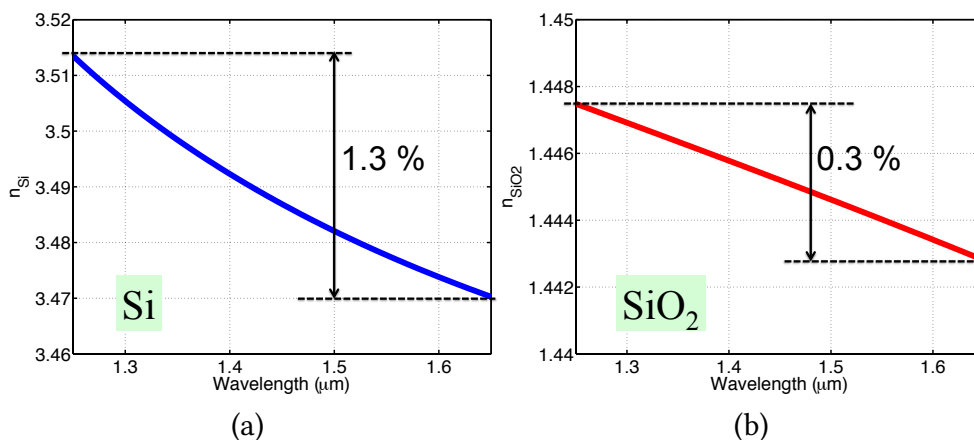
Figure 2.1: (a) Plot of index of Silicon as a function of wavelength. 1.3% variation is observed. (b) Plot of index of Silica as a function of wavelength. Only 0.3% variation is observed. Therefore, index of Silica can be considered constant.

2.1 Kramers-Kronig Relations

The complex wave-vector of a plane wave in a material is given by $K = \beta - j\alpha/2$ where β is the phase propagation constant and α is the absorption coefficient of material for the optical power. In other terms, the optical power of the wave inside the material decays according to the following equation:

$$P(z) = P(0) \times \exp(-\alpha z). \tag{2.2}$$

The complex refractive index of material is represented by $n - j\kappa$. Therefore:

$$K = \beta - j\alpha/2 = \frac{2\pi}{\lambda}(n - j\kappa) \tag{2.3}$$

and

$$\beta(f) = \frac{2\pi}{\lambda}n(f) \quad , \quad \alpha(f) = \frac{4\pi}{\lambda}\kappa(f) \tag{2.4}$$

where λ is the optical wavelength given by $\lambda = c/f$. For a physical material, the dielectric constant $\varepsilon(t)$ and the refractive index $n(t) = \sqrt{\varepsilon(t)}$ must be causal, meaning that these values should be zero for $t < 0$. It is then easy to see that we can write the following

$$n(t) = n(t) \times sign(t). \tag{2.5}$$

Taking the Fourier transform of both sides of this equation reveals that

$$n(f) - j\kappa(f) = (n(f) - j\kappa(f)) \circledast \frac{1}{j\pi f} \tag{2.6}$$

where $\circledast$ is the convolution operator. Comparing the real part of the right and left hands of this equation gives the following relations between the frequency dependent $n(f)$ and $\kappa(f)$ [163]:

$$n(f) = -\kappa(f) \circledast \frac{1}{\pi f} \quad , \quad \kappa(f) = n(f) \circledast \frac{1}{\pi f}. \tag{2.7}$$

By substituting $\alpha(f) = 4\pi/(c/f) \times \kappa(f)$, the relation between refractive index $n(f)$ and absorption coefficient $\alpha(f)$ is reached as

$$n(f) = -\frac{c}{4\pi^2}\text{PV}\left(\int_{-\infty}^{\infty} \frac{\alpha(f')}{f'} \frac{df'}{f - f'}\right) \tag{2.8}$$

where PV indicates the Cauchy principal value of the integral [122]. This integral has two singularities at $f' = 0$ and $f' = f$. The first singularity can be eliminated by breaking the

integral in two parts: $(0, \infty)$ and $(-\infty, 0)$. Since $n(t)$ is a real function of time, its Fourier transform must have a Hermition symmetry. Therefore, $n(f)$ is an even function of f [i.e. $n(f) = n(-f)$] and $\kappa(f)$ is an odd function of f [i.e. $\kappa(-f) = -\kappa(f)$]. Noting that $\alpha(f) = 4\pi f / c \times \kappa(f)$ reveals that both $n(f)$ and $\alpha(f)$ must be even functions of f. Therefore we can write

$$n(f) = \frac{n(f) + n(-f)}{2} = -\frac{c}{8\pi^2} \text{PV} \left(\int_{-\infty}^{\infty} \frac{\alpha(f')}{f'} (\frac{1}{f - f'} - \frac{1}{f + f'}) df' \right) \qquad (2.9)$$

which eventually leads to the following relation

$$n(f) = \frac{c}{4\pi^2} \text{PV} \left(\int_{-\infty}^{\infty} \frac{\alpha(f')}{f'^2 - f^2} df' \right). \qquad (2.10)$$

It is well known that the presence of charge-carriers in crystalline silicon affects its optical properties [174]. Commonly, active devices exploit this by either injecting carriers into an undoped region or by removing (depleting) charge-carriers from a doped region. It is difficult to directly measure material index changes due to injection or depletion, but there are experimental data in the literature for the absorption of doped bulk silicon wafers. Equivalence is assumed between charge-carrier absorption introduced through injection and through doping. By subtracting the absorption spectrum of undoped silicon from the absorption spectra of doped silicon, spectra for the change in absorption due to the introduction of charge-carriers can be obtained [174]. By rewriting this integral in terms of $\omega = 2\pi f$, the Kramers–Kronig coupling between Δn and $\Delta \alpha$ is

$$\Delta n(\omega) = \frac{c}{\pi} \times \text{PV} \left(\int_{0}^{\infty} \frac{\Delta\alpha(\omega') d\omega'}{\omega'^2 - \omega^2} \right) \qquad (2.11)$$

where $h\omega/2\pi$ is the photon energy. If the number of carriers introduced through doping

is ΔN, the change in absorption due to charge-carriers is

$$\Delta\alpha(\omega, \Delta N) = \alpha(\omega, \Delta N) - \alpha(\omega, 0), \tag{2.12}$$

and consequently the change in the refractive index is given by

$$\Delta n(\omega) = n(\omega, \Delta N) - n(\omega, 0). \tag{2.13}$$

2.2   Types of 3D Silicon Waveguides

The silicon waveguides are the fundamental elements to be studied. The most important property that we are interested in is the optical mode and the effective index of that mode. The dispersion of the waveguide is related to the dependence of the effective index of the waveguide mode on the wavelength. Figure 2.2 shows the two most common types of the silicon waveguides based on the silicon-on-insulator (SOI) technology. Fig. 2.2a is usually referred to as the strip waveguide. The core material is silicon and the surrounding is the oxide (SiO$_2$). The width of the waveguide is denoted by w and the height is denoted by h. Fig. 2.2b is the curved version of the strip waveguide with a radius of curvature denoted by R. Due to the high index contrast between core and its surrounding, the bent radius can be as small as 1.5μm with small optical loss [208]. Fig. 2.2c is referred to as the rib (ridge) waveguide. This structure consists of a rectangular geometry sitting on top of a thin silicon slab. The height of the slab is denoted by h_{slab} which is typically in the range of 50 nm–100 nm. Both of these waveguide designs can confine light within the boundaries of the silicon region. Note that since the index contrast between the silicon and silica is high, the confinement is possible for such small dimensions. Typical dimensions of these waveguides are 400–500 nm (width) $\times$ 220 nm (height). Due to the standard fabrication processes that have been developed by silicon foundries for silicon photonics technology [102], the height of the waveguide is kept at a constant value such as 220 nm and the
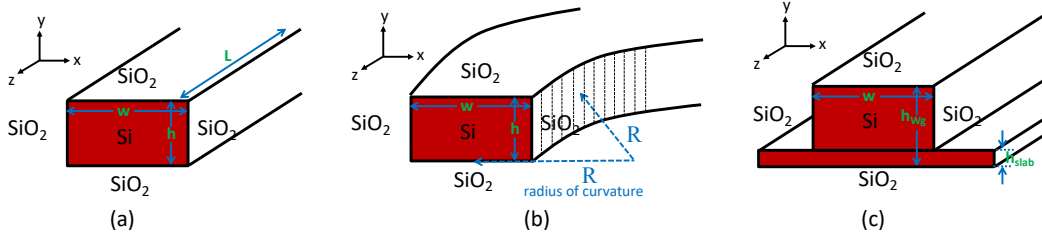
Figure 2.2: Three structures of rectangular waveguides in SiP platform. (a) Rectangular Strip waveguide. The width is denoted by w and the height of the waveguide is denoted by h. (b) Bent (curved) strip waveguide. Radius of curvature is denoted by R and could be as small as 4 μm with not much radiation loss. (c) Rib waveguide. The height of the waveguide is denoted by h_{wg} and the height of the slab is denoted by h_{slab}.

designer does not have the freedom to change it arbitrarily.

A *mode* of the waveguide is defined as a non-zero (non-trivial) solution of Maxwell's equations that has an exponential decay outside the boundaries of the rectangular silicon waveguide. Each mode is identified by its complex propagation constant defined as

$$K_{mode} = \frac{2\pi}{\lambda}(n_{\text{eff}} - j\kappa_{\text{eff}})$$
(2.14)

where n_{eff} is the effective index of the optical mode of the waveguide and κ_{eff} accounts for the propagation extinction (related to the optical loss) along the waveguide. If a plane wave propagates in free space, then $n_{\text{eff}} = 1$ and $\kappa_{\text{eff}} \approx 0$. However, the waveguide boundaries and the wavelength-dependence of the refractive index of both silicon and silica cause the effective index of a waveguide mode to be higher than the free space. Since the waveguide has two dimensions in its cross-section, w for width and h for height, we are interested in seeing how changing these dimensions will affect the effective index of the mode (waveguide dispersion). The height of the waveguide is typically constrained by the fabrication to 220 nm (Si on top of SiO$_2$ – SOI platform), hence we only examine the effect of w on the effective index. In order to do this, we use COMSOL multiphysics software and perform mode solver analysis. We draw the cross-section of the waveguide as shown in Fig. 2.3a and define the Si and SiO$_2$ regions. Then we need to define the meshing
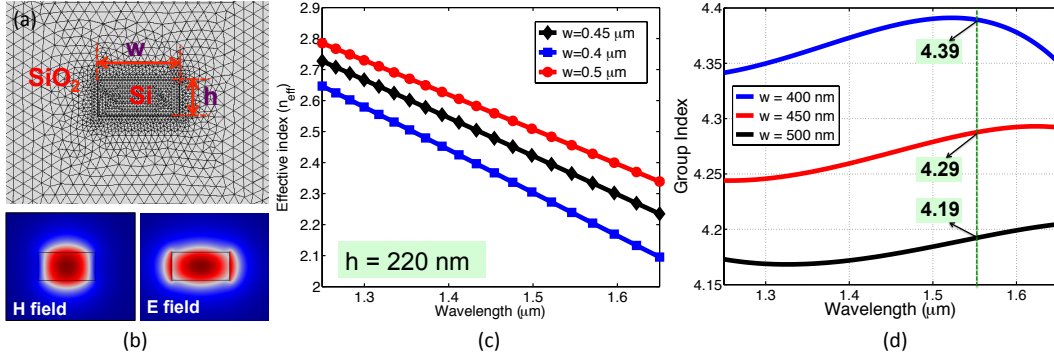
Figure 2.3: COMSOL mode analysis of silicon strip waveguide. (a) Physical structure with fine mesh for the core. (b) Mode profiles of the fundamental mode (TE_{00}). (c) Effective index of the fundamental mode for strip waveguide with h = 220 nm and w = 400, 450, 500 nm. (d) Group index of the optical mode for three different cases.

condition. Inasmuch as we know the waveguide mode is mostly within the boundaries of the waveguide core, we define an extremely fine mesh for inside the waveguide and a coarse mesh for outside the waveguide. Then the mode solver is setup to find the fundamental mode (the mode with the highest eigen value, i.e. effective index). The electric and magnetic profiles of this mode are shown in Fig. 2.3b. The electric field has a long tail outside the waveguide in the horizontal direction, whereas the magnetic field has a tail in the vertical direction.

Due to the dispersion of silicon, the effective index has a dependence on the wavelength. We calculate and plot the effective index of the fundamental mode of the waveguide for w = 400 nm, 450 nm, 500 nm in Fig. 2.3c. As can be observed, increasing the width of the waveguide will increase the effective index, and increasing the wavelength will decrease the effective index. Due to the time consuming nature of the Finite Element simulations and to build a standard model for the waveguide, we choose the 450 nm×220 nm cross-section as the standard dimension for the strip waveguide and fit a forth-order polynomial to the calculated effective index of the fundamental mode over the wavelength range of 1.5 μm−1.6 μm. The polynomial is given by

$$n_{\text{eff}}(\lambda_{\mu m}) = A_4(\lambda_{\mu m})^4 + A_3(\lambda_{\mu m})^3 + A_2(\lambda_{\mu m})^2 + A_1(\lambda_{\mu m}) + A_0 \qquad (2.15)$$

where A_4 = 0.3391437655, A_3 = -1.954733, A_2 = 4.155727848, A_1 = -5.089414343, A_0 = 5.585688163 and the wavelength is in μm unit.

Next, we use the calculated effective index to find the group index of the optical mode inside the waveguide. *Group index* is a relatively important parameter that describes the speed at which the envelop of light propagates inside the waveguide. This parameter is denoted by n_g and calculated as

$$n_g = \frac{d}{d\omega}(\omega \times n_{\text{eff}}(\omega)) = n_{\text{eff}}(\omega) + \omega\frac{dn_{\text{eff}}(\omega)}{d\omega}, \tag{2.16}$$

and considering the equivalence of $\omega/d\omega = -\lambda/d\lambda$, we can rewrite the above equation in terms of wavelength:

$$n_g = n_{\text{eff}}(\lambda) - \lambda\,\frac{dn_{\text{eff}}(\lambda)}{d\lambda}. \tag{2.17}$$

Based on the fourth-order polynomial expansion for the effective index, the group index is calculated as:

$$n_g(\lambda_{\mu m}) = A_0 - A_2(\lambda_{\mu m})^2 - 2A_3(\lambda_{\mu m})^3 - 3A_4(\lambda_{\mu m})^4. \tag{2.18}$$

We evaluate the group index for the three waveguide widths presented in Fig. 2.3c and plot them in Fig. 2.3d. Note that for the fundamental optical mode of the waveguide $dn_{\text{eff}}/d\lambda <$ 0. At the telecomm wavelength 1.55μm, the group index for the three cases is about 4.39, 4.29, and 4.19, respectively. As can be seen, because of the high index contrast and high confinement of the optical mode, the group index is much higher than the effective index of the optical mode. This means that using n_{eff} in place of n_g for approximation purposes may lead to large errors and better to be avoided as much as possible [105].

To further explain the implications of the group velocity v_g and group index n_g, consider the following cases:

- **An electric field is launched at the input of the waveguide (defined as $z = 0$ along**

the z-axis of the waveguide).

Assume that this electric field only excites the fundamental mode of the waveguide and for simplicity we ignore the role of mode profile on the propagation behavior. At the input of the waveguide the field is characterized as:

$$E_1(0,t) = E_0 e^{j\omega_1 t} + C.C. \tag{2.19}$$

where C.C. denotes the complex conjugate term. The response of the Maxwell's equations to this harmonic excitation at frequency ω is a propagating field with a propagation constant denoted by $\beta(\omega) = n_{\text{eff}}(\omega) \times 2\pi/\lambda$. At each location inside the waveguide, the field is therefore given by:

$$E_1(z,t) = \xi E_0 e^{-j\beta(\omega_1)z} e^{j\omega_1 t} + C.C. \tag{2.20}$$

where ξ denotes the linear attenuation of the field along the waveguide. In other words, the relation between the input and output fields can be given by:

$$E_1(z,t) = E_1(0, t - \frac{\omega_1}{\beta(\omega_1)}) \tag{2.21}$$

where $\omega/\beta(\omega)$ is defined as the phase velocity of wave.

$$v_{phase} = \frac{\omega}{\beta(\omega)} = \frac{3 \times 10^8 m/s}{n_{\text{eff}}(\omega)}. \tag{2.22}$$

- **A narrow-band pulse is launched at the input of the waveguide (defined as $z = 0$).** This case is shown in Fig. 2.4a.

In this case, the time dependence of the electric field at the input of the waveguide can be
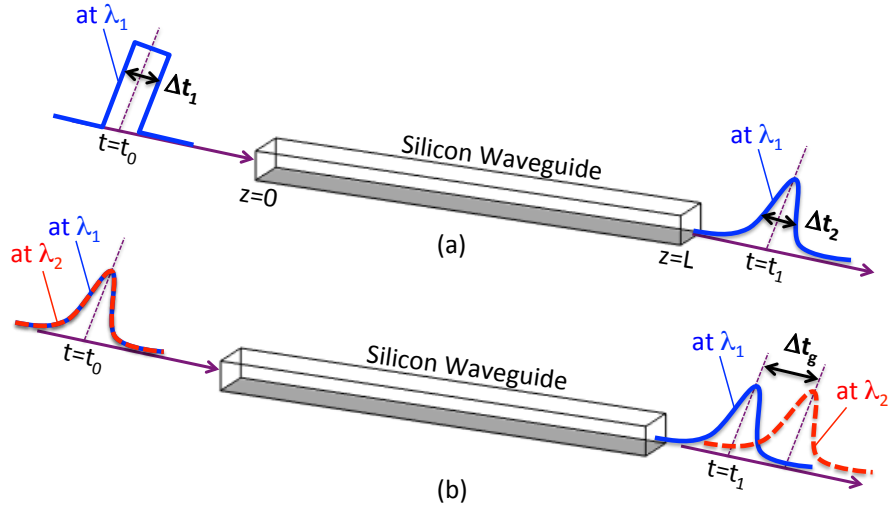
Figure 2.4: (a) Schematic of propagation of a single pulse in a silicon waveguide. The delay in the envelope of electric field and the broadening of the pulse in the time domain are described by the group velocity and dispersion of the waveguide. (b) Schematic of propagation of two optical pulses at different carrier wavelengths. The relative delay Δt_g is related to the group velocity dispersion of the waveguide.

described as the amplitude modulation of the carrier frequency:

$$E_2(0,t) = A(0,t)e^{j\omega_1 t} + C.C. \tag{2.23}$$

Since the input excitation consists of multiple frequencies, a Fourier decomposition is necessary to analyze the output response:

$$E_2(0,t) = \int_{-\infty}^{\infty} \tilde{A}(0,\omega)e^{j\omega t}e^{j\omega_1 t}d\omega/2\pi + C.C. \tag{2.24}$$

where we note that due the narrow-band nature of the amplitude $|\omega| \ll \omega_1$. The linear response of the waveguide propagation can then be written as

$$E_2(z,t) = \int_{-\infty}^{\infty} \tilde{A}(0,\omega)e^{-j\beta(\omega+\omega_1)z}e^{j(\omega+\omega_1)t}d\omega/2\pi + C.C. \tag{2.25}$$

and to the first order approximation, $\beta(\omega+\omega_1)$ can be expanded by a Taylor series. There-

fore

$$\beta(\omega + \omega_1) \approx \beta(\omega_1) + \frac{d\beta}{d\omega} \times \omega = \beta(\omega_1) + \frac{1}{v_g(\omega_1)} \times \omega \tag{2.26}$$

where v_g is called the group velocity. By substituting this approximation, the electric field at the output is given by

$$E_2(z, t) \approx A(0, t - \frac{z}{v_g}) \times e^{j\omega_1(t - z/v_{phase})} + C.C. \tag{2.27}$$

The interpretation of this result is that the amplitude modulation of the input gets a delay equal to z/v_g and the phase of the carrier gets a delay equal to z/v_{phase}. In other words, the envelope of the mode travels at the group velocity, while the carrier travels at the phase velocity. Therefore, as shown in Fig. 2.4a, the difference between t_0 and t_1 is equal to the group delay. As shown in Fig. 2.4b, two pulses with different wavelengths come out with a delay due to the wavelength-dependence of group velocity.

Note that this equation for the propagation of narrow-band pulse only holds as long as the group velocity is almost constant (frequency independent) over the bandwidth of the pulse. If the bandwidth of excitation pulse is large such that this assumption is no longer valid, we can add one more correction term to the Taylor expansion of the propagation constant $\beta(\omega)$:

$$\beta(\omega + \omega_1) \approx \beta_0 + \beta_1 \times \omega + 1/2 \times \beta_2 \times \omega^2 \tag{2.28}$$

where $\beta_0 = \beta(\omega_1)$, $\beta_1 = d\beta/d\omega$, and $\beta_2 = d^2\beta/d\omega^2$ calculated at ω_1. Using this approximation, the electric field at a position along the waveguide is estimated as:

$$E_t(z, t) = e^{j\omega(t - z/v_{phase})} \int_{-\infty}^{\infty} \tilde{A}(0, \omega) e^{j\omega(t - z/v_g(\omega))} d\omega/2\pi + C.C. \tag{2.29}$$

where the phase velocity is $v_{phase} = \omega/\beta_0$, and the frequency dependent group velocity

is defined as

$$v_g(\omega) = \frac{1}{\beta_1 + 1/2 \times \beta_2 \times \omega}. \tag{2.30}$$

Note that the unit of β_1 is s/m and unit of β_2 is s^2/m. This expression implies that different frequency components of the input pulse travel at different velocities. This phenomenon is called *group velocity dispersion* and results in the broadening of the pulse. For example as shown in Fig. 2.4a, if the input pulse has a width of Δt_1, the output pulse has a width of $\Delta t_2 > \Delta t_1$. To put this under more perspective, we approximate the input pulse as an Gaussian pulse which is mathematically expressed as

$$A(0,t) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(t-t_0)^2/2\sigma^2} \Rightarrow \tilde{A}(0,\omega) = e^{-j\omega t_0} \times e^{-\sigma^2\omega^2/2}. \tag{2.31}$$

which shows that the Fourier transform of the Gaussian pulse is a Gaussian distribution. The 90% width in the time domain and the frequency domain are given by

$$\Delta t_1(90\%) \approx \sigma \quad , \quad \Delta\omega_1(90\%) \approx \frac{1}{\sigma}. \tag{2.32}$$

The result of the propagation integral for the Gaussian pulse can then be straightforwardly calculated as

$$A(z,t) \approx \frac{1}{2\sqrt{\pi}\sqrt{j\beta_2 z/2 + \sigma^2/2}}e^{-\frac{(t-t_0-\beta_1 z)^2}{4(j\beta_2 z/2+\sigma^2/2)}} \tag{2.33}$$

If we assume that $\beta_2 z << \sigma^2$, the pulse is further approximated as

$$|A(z,t)| \approx \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(t-t_0-\beta_1 z)^2}{2\sigma^2(1+(\beta_2 z/\sigma^2)^2)}}. \tag{2.34}$$

Based on this equation, we see that the peak of the pulse gets a delay equal to $\beta_1 z$ or equivalently we can say that the peak comes out of the waveguide at the time instance $t_2 = t_1 + \beta_1 z$. The other observation is that the 90% width of the pulse has been increased

to

$$\Delta t_2(90\%) = \sigma' \approx \sigma \left[1 + \left(\frac{\beta_2 z}{\sigma^2} \right)^2 \right]^{1/2}.$$ (2.35)

The interpretation of this equation is that the width of the pulse has increased due to the β_2 term. Note that the sign of β_2 does not matter and whether it is positive or negative, the width of the pulse will increase. The approximate amount of the increase in the pulse width is given by

$$\Delta t_2(90\%) - \Delta t_1(90\%) \approx \sigma \left(\left[1 + \left(\frac{\beta_2 z}{\sigma^2} \right)^2 \right]^{1/2} - 1 \right)$$ (2.36)

For a case where the dispersion is severe such that $|\beta_2 z| >> \sigma^2$, the pulse broadening is then approximated as

$$\Delta t_2(90\%) - \Delta t_1(90\%) \approx |\beta_2| z \times \Delta \omega_1(90\%).$$ (2.37)

The dispersion coefficient of the waveguide [211] is defined based on this equation

$$D_\omega = \beta_2(\omega) = \frac{d}{d\omega} \left(\frac{1}{v_g(\omega)} \right) = \frac{1}{c} \times \frac{d}{d\omega} n_g(\omega)$$ (2.38)

and using the equivalence of $\omega/d\omega = -\lambda/d\lambda$, we can express the D_ω in terms of wavelength

$$D_\lambda = \frac{d}{d\lambda} \left(\frac{1}{v_g(\omega)} \right), D_\omega = \frac{1}{c\omega} \times \omega \frac{d}{d\omega} n_g(\omega) = -\frac{\lambda^2}{2\pi c^2} \times \frac{d}{d\lambda} n_g(\lambda)$$ (2.39)

and finally, the relation between D_λ and D_ω is given by

$$D_\lambda = -\frac{2\pi c}{\lambda^2} D_\omega = -\frac{2\pi c}{\lambda^2} \beta_2$$ (2.40)

and the following equality holds:

$$\Delta t_2 - \Delta t_1 \approx |D_\lambda \times \Delta\lambda| \approx |D_\omega \times \Delta\omega| \approx |D_f \times \Delta f| \tag{2.41}$$

and the amount of pulse broadening can be calculated based on wavelength or frequency bandwidth of the pulse.

2.3 2D Dielectric Slab Waveguides

Although two-dimensional waveguides are not practical, they hold a special place in understanding the modes of three-dimensional waveguides. Therefore, we start by analyzing the modes of 2D waveguides. The cross-section of a two dimensional dielectric slab waveguide is shown in Fig. 2.5a. This waveguide consists of three layers: the lower layer is called the substrate layer. The middle layer is called the core layer in which most of the light is propagating all the time. The upper layer is called the cladding layer. In order to guide light through the core in the z-direction, the index of the core should be larger than the substrate and the cladding layer. The ray optics perspective is that the rays of light must undergo total internal reflection in the core-substrate and core-cladding interfaces in order to remain in the core layer all the time. The typical index profile of the dielectric slab waveguide is given by

$$n(x) = \begin{cases} n_{\text{cladding}} & \text{for} \quad x > w \\ n_{\text{core}} & \text{for} \quad 0 < x < w \\ n_{\text{substrate}} & \text{for} \quad x < 0 \end{cases} \tag{2.42}$$

which is schematically plotted in Fig. 2.5b. Note that for this type of 2D waveguide, the propagation is assumed to be in the z direction, and the material interfaces are in the x direction. The structure is assumed to be exactly the same in the y and z direction.
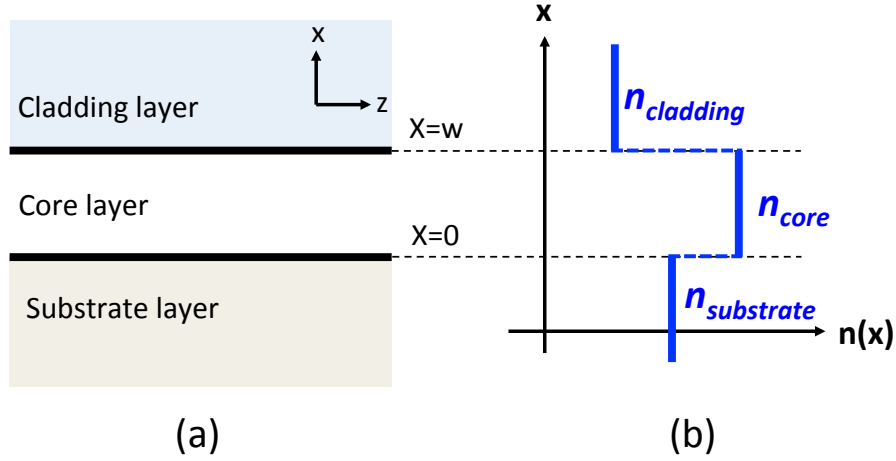
Figure 2.5: (a) Cross-section of the 2D dielectric slab waveguide. The width of the waveguide is assumed to be w and the propagation of the wave is assumed to be in the z direction. (b) Index profile of the dielectric slab waveguide.

Therefore, if a wave is propagating in the yz plane, we can always rotate the coordinate axes such that the direction of propagation is aligned with the z direction.

The classic way to analyze the wave propagation in the dielectric slab waveguide is through the source-free Maxwell's equations:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.43}$$

$$\nabla \cdot \mathbf{D} = 0 \tag{2.44}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \tag{2.45}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.46}$$

where the material is assumed non-magnetic $\mathbf{B} = \mu_0\mathbf{H}$ and the material response is reflected into the polarization dipole $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P} = \epsilon_0 n^2(x)\mathbf{E}$. Combining the curl equations will lead to the well-known wave equation for the electric field

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2\mathbf{E} = -\mu_0\epsilon_0 n^2(x)\frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{2.47}$$

which applies in each of the three layers separately with the simplification of $\nabla \cdot \mathbf{E} = 0$. Therefore, the wave equation in each layer boils down to the following:

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 n^2 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{2.48}$$

Another aspect of the 2D structure is the symmetry in the y and z direction, therefore we can conveniently assume that $\partial/\partial y = 0$. With this assumption, the first curl equation can be decomposed component-wise as

$$\frac{\partial E_y}{\partial z} = \mu_0 \frac{\partial H_x}{\partial t} \tag{2.49}$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -\mu_0 \frac{\partial H_y}{\partial t} \tag{2.50}$$

$$\frac{\partial E_y}{\partial x} = -\mu_0 \frac{\partial H_z}{\partial t} \tag{2.51}$$

and the second curl equation is decomposed as

$$\frac{\partial H_y}{\partial z} = -\epsilon_0 n^2 \frac{\partial E_x}{\partial t} \tag{2.52}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = \epsilon_0 n^2 \frac{\partial E_y}{\partial t} \tag{2.53}$$

$$\frac{\partial H_y}{\partial x} = \epsilon_0 n^2 \frac{\partial E_z}{\partial t} \tag{2.54}$$

A close inspection of these two sets of equations reveals that the coupling of electric and magnetic fields' components can be described by two distinct set of equations:

- **Transverse Electric (TE^z) modes**

These are the modes for which the electric field has no component in the direction of propagation, i.e. $E_z = 0$. For these modes only the field components E_y, H_x and H_z are

nonzero and connected together via the following equations:

$$\frac{\partial E_y}{\partial z} = \mu_0 \frac{\partial H_x}{\partial t} \tag{2.55}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = \epsilon_0 n^2 \frac{\partial E_y}{\partial t} \tag{2.56}$$

$$\frac{\partial E_y}{\partial x} = -\mu_0 \frac{\partial H_z}{\partial t} \tag{2.57}$$

Now we consider the harmonic solutions such as

$$E_y(x, z, t) = E(x)\ e^{-j\beta z}\ e^{j\omega t} + C.C. \tag{2.58}$$

where $E(x)$ is the field mode profile along the x direction due to the index discontinuity at the interfaces at $x = 0$ and $x = w$. The wave equation for the E_y component then leads to

$$\frac{\partial^2}{\partial x^2} E(x) + (k_0^2 n^2 - \beta^2) E(x) = 0 \tag{2.59}$$

where $k_0 = \omega\sqrt{\mu_0 \epsilon_0} = \omega/c_0 = 2\pi/\lambda_0$ is the free-space wavenumber and $\beta = k_0 n_{\text{eff}}$ is the propagation constant of the mode. The x component of the k-vector can be defined as

$$k_x = \sqrt{k_0^2 n^2 - \beta^2} \tag{2.60}$$

if $k_0 n > \beta$ and defined as

$$k_x = -j\sqrt{\beta^2 - k_0^2 n^2} \tag{2.61}$$

if $k_0 n < \beta$. The solution for $E(x)$ would be a possible combination of two propagating waves in $\pm\hat{x}$ directions:

$$E(x) = E_+\ e^{-jk_x x} + E_-\ e^{jk_x x}. \tag{2.62}$$

The *transverse impedance of the mode* is defined as the ratio between the transverse com-

ponents of the electric and magnetic fields:

$$Z_{\text{trans}}^{TE} = -\frac{E_y}{H_x} = \frac{\omega \mu_0}{\beta} = \frac{\eta_0}{n_{\text{eff}}}, \tag{2.63}$$

and the *longitudinal impedance of the mode* is defined as the ratio between the electric field and longitudinal component of the magnetic field:

$$Z_{\text{long}}^{TE} = \frac{E_y}{H_z} = \frac{\omega \mu_0}{k_x} = \frac{\eta_0}{n_x}, \tag{2.64}$$

where $n_{\text{eff}} = \beta/k_0$ and $n_x = k_x/k_0$.

A mode of the waveguide can be viewed as a resonance condition of the electric and magnetic fields that are transverse (tangential) to the discontinuity interfaces. For the TE^z modes, these components are E_y and H_z. Since the boundary conditions associated with Maxwell's equations require that these two components be continuous along each interface, we can construct a resonant transmission line model for dielectric slab waveguide [see Fig. 2.6a] based on the longitudinal components of the k-vector and the longitudinal impedance as depicted in Fig. 2.6b. From the circuit theory, the resonant condition is satisfied when the reactive power is balanced in the system:

$$Im\left(Z_{\text{in}} + Z_{\text{long}}^{subs}\right) = 0 \tag{2.65}$$

where the input impedance can be calculated by referring the clad impedance through the core layer:

$$Z_{\text{in}} = Z_{\text{long}}^{core} \times \frac{Z_{\text{long}}^{clad} + jZ_{\text{long}}^{core} \times \tan(k_x^{core}w)}{Z_{\text{long}}^{core} + jZ_{\text{long}}^{clad} \times \tan(k_x^{core}w)}. \tag{2.66}$$

By defining the new variables

$$D = \frac{Z^{subs}}{Z^{core}} \quad , \quad U = \frac{Z^{clad}}{Z^{core}} \tag{2.67}$$
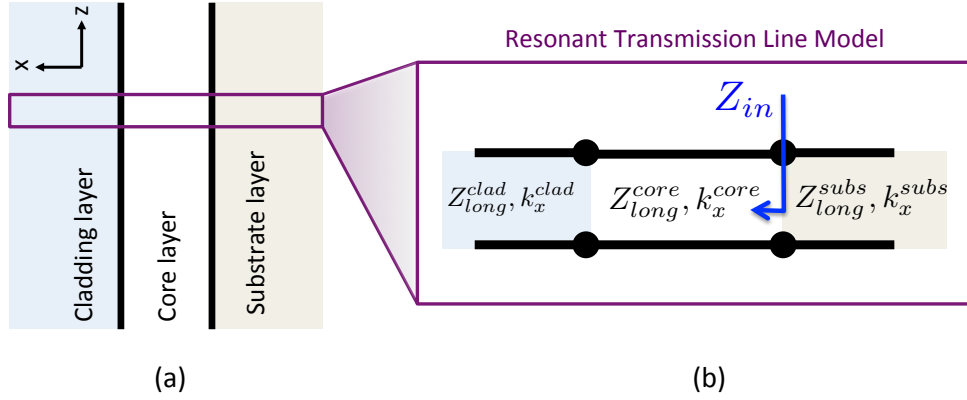
Figure 2.6: (a) 2D dielectric slab waveguide with the discontinuity in the x direction. (b) Equivalent resonant longitudinal transmission line model for the modes of the slab waveguide. k_x and Z_{long} can be calculated for both TE^z and TM^z modes.

the mode condition is written as

$$Im\left(D + \frac{U + j\tan(k_x^{core}w)}{1 + jU\tan(k_x^{core}w)}\right) = 0. \tag{2.68}$$

Since both D and U are pure imaginary and k_x^{core} is real, the mode equation becomes:

$$\tan(k_x^{core}w) = j\frac{D + U}{1 + DU} \tag{2.69}$$

which in general applies to both TE and TM modes.

For TE^z modes, if we define

$$k_x^{subs} = -j\gamma_{subs} = -jk_0\sqrt{n_{\text{eff}}^2 - n_{subs}^2} \tag{2.70}$$

$$k_x^{clad} = -j\gamma_{clad} = -jk_0\sqrt{n_{\text{eff}}^2 - n_{clad}^2} \tag{2.71}$$

the mode equation becomes [128]:

$$\text{TE modes:} \quad \tan(k_x^{core}w) = \frac{k_x^{core}(\gamma_{subs} + \gamma_{clad})}{(k_x^{core})^2 - \gamma_{subs} \times \gamma_{clad}}, \tag{2.72}$$
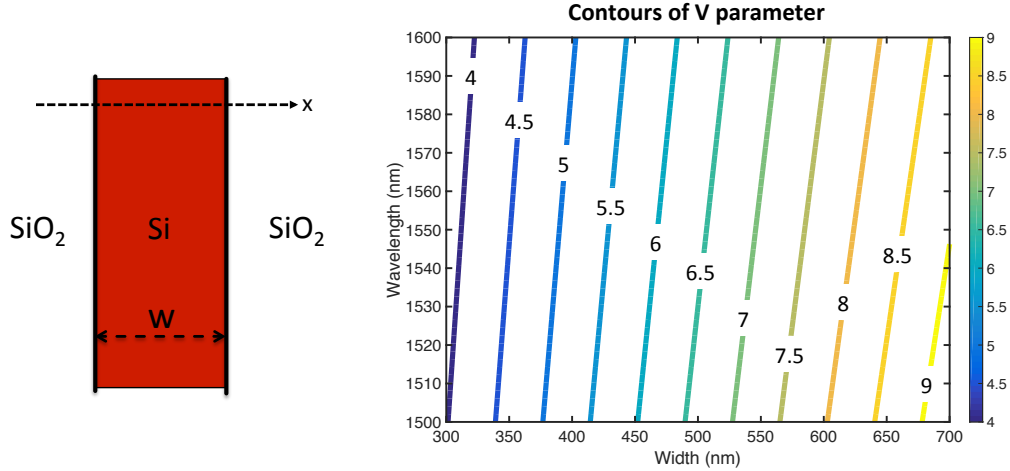
Figure 2.7: (Left) Cross-section of the 2D silicon slab waveguide. (Right) Contours of the V parameters as a function of width of the slab and the wavelength over the 1500nm–1600nm range. V parameter varies from 4 to 9.

or equivalently [140]:

$$k_x^{core} w = m\pi + \tan^{-1}\left(\frac{\gamma_{subs}}{k_x^{core}}\right) + \tan^{-1}\left(\frac{\gamma_{clad}}{k_x^{core}}\right). \tag{2.73}$$

where $m = 0, 1, 2, \ldots$ indicates the mode order. The unknown parameter in this equation is the effective index n_{eff} and all the parameters can be expressed in terms of the index of materials and the effective index. This equation has discrete solutions for each combination of wavelength and width of the waveguide. These solutions refer to the fundamental mode TE_0 and higher order modes (TE_1, TE_2, ...). It can be shown that the modes can be described based on the normalized frequency number V and the normalized effective index parameter b:

$$V = 2\pi \frac{w}{\lambda} \sqrt{n_{core}^2 - n_{high}^2} \quad , \quad b = \frac{n_{\text{eff}}^2 - n_{high}^2}{n_{core}^2 - n_{high}^2} \tag{2.74}$$

where n_{high} is the higher index between the substrate and the cladding material. Figure 2.7 shows the silicon slab and plots the contours of V as a function of wavelength and width. It is almost always assumed that the index of substrate is higher than the index
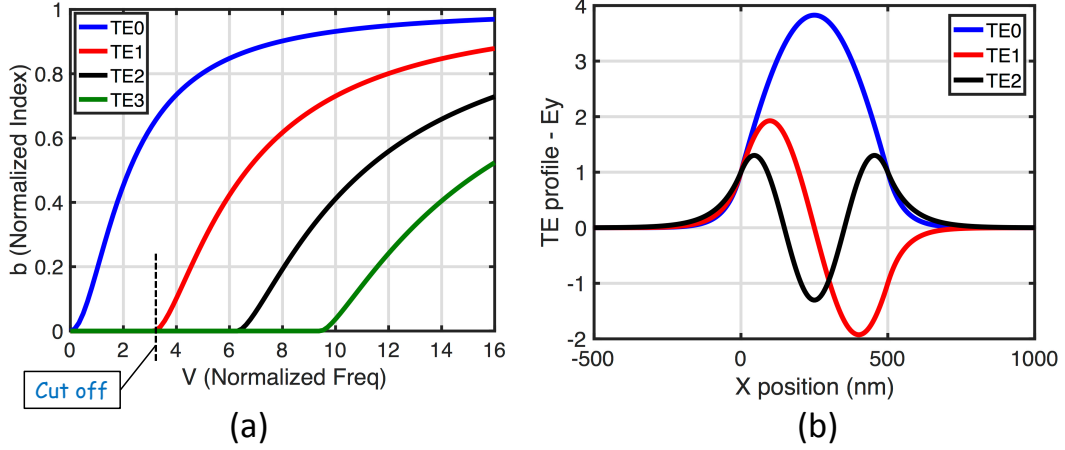
Figure 2.8: Analysis of the 2D dielectric slab waveguide for TE modes with symmetric profile. (a) Normalized effective index as a function of V. The cutoff point of the second mode is given by $V = \pi$. (b) Field profile for the first three TE modes.

of cladding. Since in general the substrate material and the cladding material are not the same, we can also define a normalized asymmetry coefficient for the TE modes:

$$a^{TE} = \frac{n_{subs}^2 - n_{clad}^2}{n_{core}^2 - n_{subs}^2}. \tag{2.75}$$

Based on the normalized parameters, it is easy to show that:

$$k_x^{core} w = V\sqrt{1-b} \tag{2.76}$$

$$\gamma_{subs} w = V\sqrt{b} \tag{2.77}$$

$$\gamma_{clad} w = V\sqrt{b + a^{TE}} \tag{2.78}$$

and therefore the TE mode equation is written as

$$V\sqrt{1-b} = m\pi + \tan^{-1}\left(\sqrt{\frac{b}{1-b}}\right) + \tan^{-1}\left(\sqrt{\frac{b + a^{TE}}{1-b}}\right) \tag{2.79}$$

37
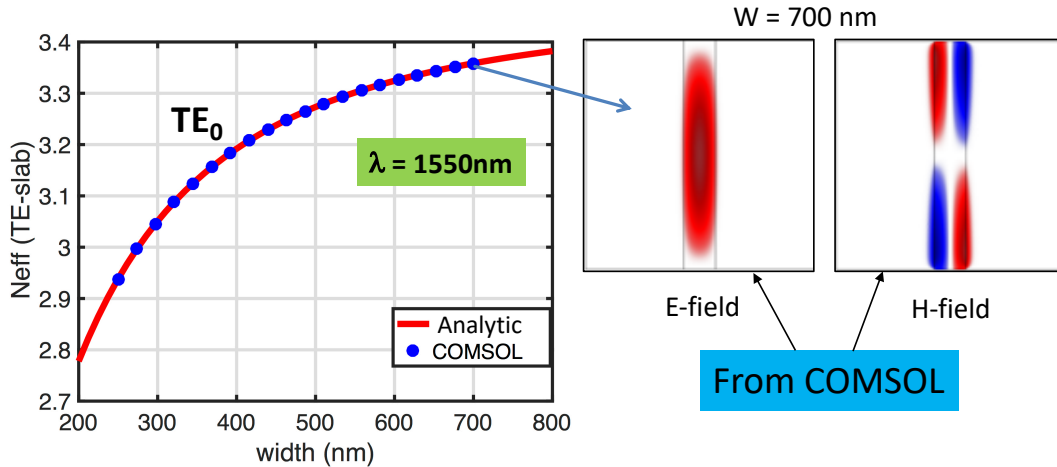
Figure 2.9: Effective index of the fundamental TE mode of the 2D dielectric slab waveguide $SiO_2|Si|SiO_2$ calculated by COMSOL (circles) and analytic equation (solid) at 1550nm wavelength as a function of slab width. The field profiles are given for w = 700 nm.

and the cutoff of the mth order mode is given by

$$V_{\text{cutoff}} = m\pi + \tan^{-1}\left(\sqrt{a^{TE}}\right).$$

(2.80)

A mode solver was then implemented to solve the TE mode equation of the slab waveguide. This mode solver can find all the TE modes and calculate effective index and normalized mode index. The results of the mode analysis for TE_0, TE_1, TE_2 and TE_3 modes of a 500 nm wide waveguide whose core is silicon and substrate and cladding are silica are shown in Fig. 2.8a. The transverse component of the electric field is calculated for the first three modes and plotted in Fig. 2.8b. The effective index of the slab waveguide as a function of its width is calculated in Fig. 2.9 and compared against COMSOL mode solver, showing the exact agreement between the two. As expected, a wider waveguide results in a stronger confinement and a higher effective index.

- **Transverse Magnetic (TM^z) modes**

These are the modes for which the magnetic field has no component in the direction of propagation, i.e. $H_z = 0$. For these modes only the field components H_y, E_x and E_z are

nonzero and connected together via the following equations:

$$\frac{\partial H_y}{\partial z} = -\epsilon_0 n^2 \frac{\partial E_x}{\partial t} \tag{2.81}$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -\mu_0 \frac{\partial H_y}{\partial t} \tag{2.82}$$

$$\frac{\partial H_y}{\partial x} = \epsilon_0 n^2 \frac{\partial E_z}{\partial t}. \tag{2.83}$$

Similar to the TE^z modes, we can define the transverse impedance of the harmonic mode as the relation between the transverse components of the electric and magnetic fields:

$$Z_{\text{trans}}^{TM} = \frac{E_x}{H_y} = \frac{\beta}{\omega \epsilon_0 n^2} = \eta_0 \frac{n_{\text{eff}}}{n^2} \tag{2.84}$$

and the longitudinal impedance is defined as

$$Z_{\text{long}}^{TM} = \frac{E_z}{H_y} = \frac{k_x}{\omega \epsilon_0 n^2} = \eta_0 \frac{n_x}{n^2}. \tag{2.85}$$

As can be seen, in the case of TM modes, the index of materials also directly appears in the impedance equations. The resonant mode condition can be written from the general mode equation based on D and U parameters:

$$D = \frac{Z_{\text{long}}^{subs}}{Z_{\text{long}}^{core}} = -j \frac{n_{core}^2}{n_{subs}^2} \frac{\gamma_{subs}}{k_x^{core}} \quad , \quad U = \frac{Z_{\text{long}}^{clad}}{Z_{\text{long}}^{core}} = -j \frac{n_{core}^2}{n_{clad}^2} \frac{\gamma_{clad}}{k_x^{core}} \tag{2.86}$$

and therefore the TM mode equation is

$$\tan(k_x^{core} w) = \frac{k_x^{core}/n_{core}^2 (\gamma_{subs}/n_{subs}^2 + \gamma_{clad}/n_{clad}^2)}{(k_x^{core}/n_{core}^2)^2 - \gamma_{subs}/n_{subs}^2 \times \gamma_{clad}/n_{clad}^2} \tag{2.87}$$

or equivalently:

$$k_x^{core} w = m\pi + \tan^{-1}\left(\frac{\gamma_{subs}/n_{subs}^2}{k_x^{core}/n_{core}^2}\right) + \tan^{-1}\left(\frac{\gamma_{clad}/n_{clad}^2}{k_x^{core}/n_{core}^2}\right) \tag{2.88}$$
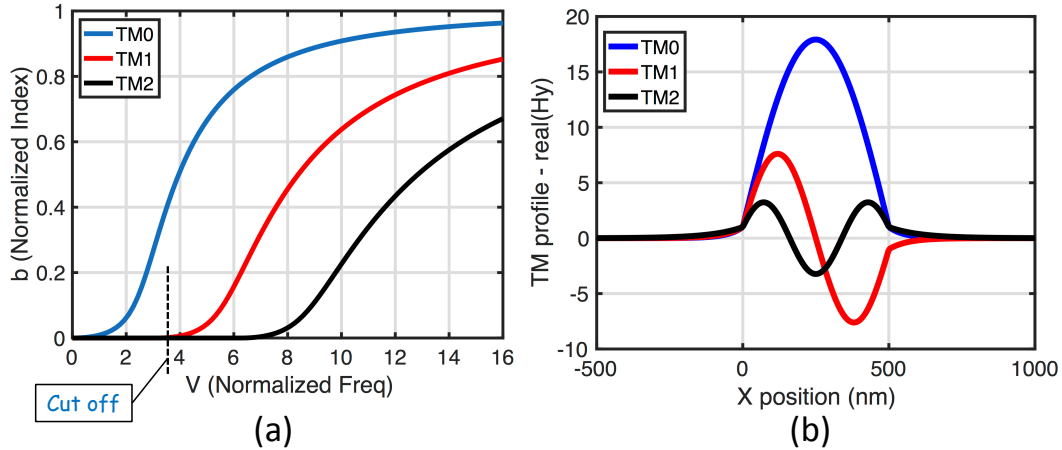
Figure 2.10: Analysis of the 2D dielectric slab waveguide for TM modes. (a) Normalized effective index as a function of V. (b) Field profile for the first three TM modes.

where $m = 0, 1, 2, \ldots$ indicates the mode order. The asymmetry factor for the TM modes is defined as

$$a^{TM} = \frac{n_{core}^4}{n_{clad}^4} \times \frac{n_{subs}^2 - n_{clad}^2}{n_{core}^2 - n_{subs}^2}. \tag{2.89}$$

A mode solver was implemented to find all the TM modes and calculate the effective index and normalized mode index. The results of the mode analysis for TM_0, TM_1, and TM_2 modes of a 500 nm wide waveguide whose core is silicon and substrate and cladding are silica are shown in Fig. 2.10a. The transverse component of the magnetic field, H_y, is calculated for the first three modes and plotted in Fig. 2.10b.

Formulation of the Transfer Matrix Method

Although the equivalent resonant transmission line circuit was used to analyze the modes of the 2D dielectric slab waveguide, another convenient way of dealing with propagation of light in layered media (especially for the case of two coupled slab waveguides) is the use of transfer matrix method. Such transfer matrix describes the evolution of the amplitude of tangential components of the electric field and magnetic field to the material interface. Consider the situation of a material interface at $x = X$ as shown in Fig. 2.11. Since the
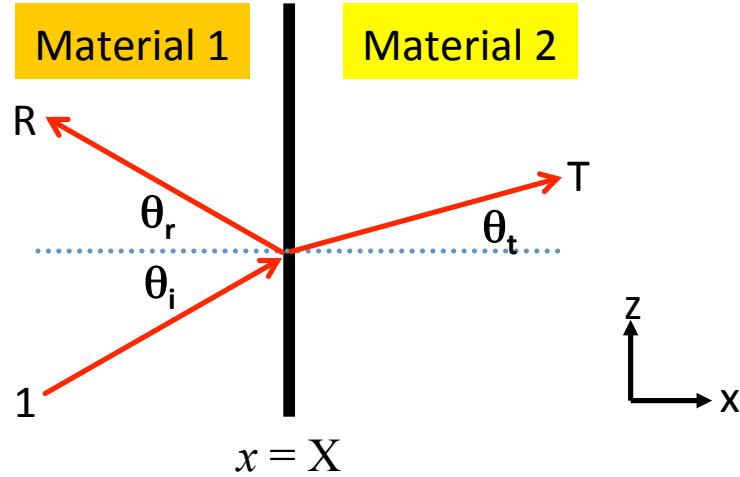
Figure 2.11: The interface of two materials. The normal direction of the interface is assumed to be x.

interface is in the x direction, we can freely choose a coordinate system such that the wave vector $\mathbf{k}$ of the light lies within the xz plane. Therefore, any dependence and variation in the y direction can be ignored, i.e. $\partial/\partial y = 0$. As we showed for the dielectric waveguide, this simplifying assumption leads to two distinct set of modes that relate components of the electric and magnetic fields together, namely TE modes and TM modes. The transfer matrix is defined to relate the amplitudes of forward (propagating in $+\hat{x}$ direction) and backward (propagating in $-\hat{x}$ direction) tangential components of electric or magnetic fields at the interface, i.e.

$$
\begin{pmatrix} \text{amp-forward}(Field)_{t,2} \\ \text{amp-backward}(Field)_{t,2} \end{pmatrix} = \mathbf{Q}_{1\rightarrow 2} \begin{pmatrix} \text{amp-forward}(Field)_{t,1} \\ \text{amp-backward}(Field)_{t,1} \end{pmatrix} \tag{2.90}
$$

- **TE modes of the interface**

The equations for the TE modes involve the field components (E_y, H_x, H_z) and are given

by

$$\frac{\partial E_y}{\partial z} = \mu_0 \frac{\partial H_x}{\partial t} \tag{2.91}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = \epsilon_0 n^2 \frac{\partial E_y}{\partial t} \tag{2.92}$$

$$\frac{\partial E_y}{\partial x} = -\mu_0 \frac{\partial H_z}{\partial t}. \tag{2.93}$$

The solution for E_y includes a propagating component in the z direction and two possible propagating components in the $+x$ and $-x$ directions, therefore:

$$E_y(x, z) = (A_+ e^{-jk_x x} + A_- e^{jk_x x}) e^{-jk_z z} \tag{2.94}$$

where $k_x^2 + k_z^2 = k_0^2 n^2$ in each side of the interface. The tangential component of the magnetic field H_z is then given by:

$$H_z(x, z) = -\frac{1}{j\omega\mu_0} \frac{\partial E_y}{\partial x} \tag{2.95}$$

or more explicitly:

$$H_z(x, z) = -\frac{1}{j\omega\mu_0} \left((-jk_x) A_+ e^{-jk_x x} + (jk_x) A_- e^{jk_x x} \right) e^{-jk_z z}. \tag{2.96}$$

In the matrix terms, we can write the tangential components together as

$$\begin{pmatrix} E_y \\ H_z \end{pmatrix} = \mathbf{M}^{\text{TE}}(x) \times \begin{pmatrix} A_+ \\ A_- \end{pmatrix} \times e^{-jk_z z} \tag{2.97}$$

where $\mathbf{M}^{\text{TE}}(x)$ is defined as

$$\mathbf{M}^{\text{TE}}(x) = \begin{bmatrix} e^{-jk_x x} & e^{jk_x x} \\ \frac{k_x}{\omega\mu_0} e^{-jk_x x} & -\frac{k_x}{\omega\mu_0} e^{jk_x x} \end{bmatrix} \tag{2.98}$$

and we notice that

$$\det(\mathbf{M}) = -\frac{2k_x}{\omega\mu_0}. \tag{2.99}$$

Note that $k_z = k_0 n \sin\theta$ should be the same on both sides of the interface for all the forward and backward waves due to the boundary conditions of Maxwell's equations. This leads to the famous Snell's law for reflection and refraction angles:

$$\text{Material 1:} \quad \sin(\theta_i) = \sin(\theta_r) \rightarrow \theta_i = \theta_r \tag{2.100}$$

$$\text{Material 2:} \quad n_1 \sin(\theta_i) = n_2 \sin(\theta_t) \tag{2.101}$$

Now it is easy to see that the relation between amplitudes of the forward and backward waves at the interface is given by

$$\begin{pmatrix} A_+^{(2)} \\ A_-^{(2)} \end{pmatrix} = \left(\mathbf{M^{TE}}_{(2)}(X)\right)^{-1} \times \mathbf{M^{TE}}_{(1)}(X) \times \begin{pmatrix} A_+^{(1)} \\ A_-^{(1)} \end{pmatrix} \tag{2.102}$$

and therefore

$$\mathbf{Q}_{1\rightarrow 2}(\mathbf{x} = \mathbf{X}) = \left(\mathbf{M}_{(2)}(\mathbf{X})\right)^{-1} \times \mathbf{M}_{(1)}(\mathbf{X}) \tag{2.103}$$

and in the same way, we can write:

$$\mathbf{Q}_{2\rightarrow 1}(\mathbf{x} = \mathbf{X}) = \left(\mathbf{M}_{(1)}(\mathbf{X})\right)^{-1} \times \mathbf{M}_{(2)}(\mathbf{X}). \tag{2.104}$$

The transfer matrix can be utilized to investigate the reflection and transmission of light at the interface. The angles of reflection and refraction obey Snell's law. The amount of power that has been reflected, i.e. *reflectance*, and has been transmitted, i.e. *transmittance* can be calculated based on the following steps:

- **Step1**: Writing the matrix equation for the forward and backward waves in both materials at the interface:
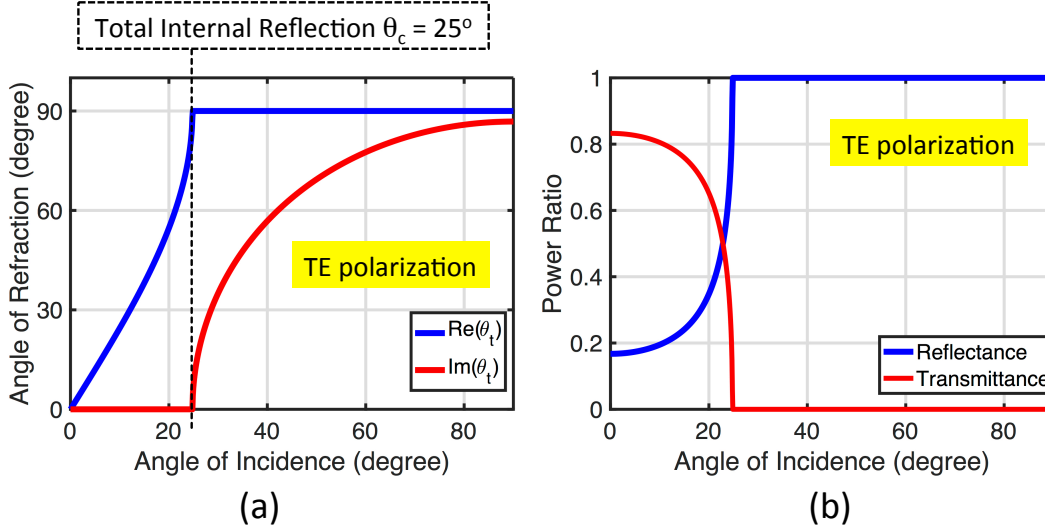
Figure 2.12: Analysis of reflection and refraction at the Si/SiO₂ interface. It is assumed that the light is incident on the interface inside the Si material. (a) Angle of refraction, θ_t as a function of angle of incidence, θ_i. Total internal reflection happens at 25° degrees. (b) Reflectance and transmittance as a function of θ_i.

$$\begin{pmatrix} T \\ 0 \end{pmatrix} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \times \begin{pmatrix} 1 \\ R \end{pmatrix} \tag{2.105}$$

- **Step2**: Finding the reflection based on the knowledge that there is no backward wave in the second material:

$$R = \frac{A_-^{(1)}}{A_+^{(1)}} = -\frac{q_{21}}{q_{22}} \Rightarrow |R_{1\to2}|^2 = \left| -\frac{q_{21}}{q_{22}} \right|^2 \tag{2.106}$$

- **Step 3**: Calculating transmittance:

$$T = q_{11} + q_{12}R = \frac{q_{11}q_{22} - q_{12}q_{21}}{q_{22}} \Rightarrow |T_{1\to2}|^2 = \left| \frac{\det(\mathbf{Q}_{1\to2})}{q_{22}} \right|^2 \tag{2.107}$$

It can be easily shown that the reflectance and transmittance coefficients obtained above are identical to the ones given by Fresnel's equations [161], when the substitution $k_x = k\cos(\theta)$ and $k_z = k\sin(\theta)$ is used. To prove this, we explicitly calculate the transfer

matrix:

$$\mathbf{Q}_{1\to 2}(x) = \begin{bmatrix} \frac{\xi_2+\xi_1}{2\xi_2}\exp(j\Delta k_x x) & \frac{\xi_2-\xi_1}{2\xi_2}\exp(j\Sigma k_x x) \\ \\ \frac{\xi_2-\xi_1}{2\xi_2}\exp(-j\Sigma k_x x) & \frac{\xi_2+\xi_1}{2\xi_2}\exp(-j\Delta k_x x) \end{bmatrix} \tag{2.108}$$

where $\Delta k_x = k_x^{(2)} - k_x^{(1)}$, $\Sigma k_x = k_x^{(2)} + k_x^{(1)}$ and

$$\xi_1^{TE} = \frac{k_x^{(1)}}{\omega\mu_0} \quad , \quad \xi_2^{TE} = \frac{k_x^{(2)}}{\omega\mu_0} \tag{2.109}$$

and note that

$$\det(\mathbf{Q}) = \frac{\xi_1}{\xi_2}. \tag{2.110}$$

We see that ξ coefficients are basically the inverse of the longitudinal impedance of the waves. Now the reflection coefficient is given by

$$R_{1\to 2}(x = X) = -\frac{q_{21}}{q_{22}} = \frac{\xi_1 - \xi_2}{\xi_1 + \xi_2} \times e^{-j2k_x^{(1)}X} \tag{2.111}$$

or equivalently

$$|R_{1\to 2}|^2 = \left| \frac{k_x^{(1)} - k_x^{(2)}}{k_x^{(1)} + k_x^{(2)}} \right|^2 = \left| \frac{n_1\cos(\theta_i) - n_2\cos(\theta_t)}{n_1\cos(\theta_i) + n_2\cos(\theta_t)} \right|^2. \tag{2.112}$$

The transmission coefficient is given by

$$T_{1\to 2}(x = X) = \frac{A_+^{(2)}}{A_+^{(1)}} = \frac{\det(\mathbf{Q}_{1\to 2})}{q_{22}} = \frac{2\xi_1}{\xi_1 + \xi_2} \times e^{j\Delta k_x X}. \tag{2.113}$$

As an example, we use the transfer matrix to analyze the Si/SiO$_2$ interface. If the light is incident inside the Si material, at some angle we expect to see the total internal reflection. This critical angle can be calculated as

$$\theta_c = \sin^{-1}\left(\frac{n_{sio_2}}{n_{si}}\right) = \sin^{-1}(\frac{1.445}{3.48}) \approx 25°. \tag{2.114}$$

Fig. 2.12a shows the transmitted angle as a function of incidence angle. As can be seen, after the total internal point, the transmitted angle becomes complex. Fig. 2.12b shows the calculated reflectance and transmittance as a function of the incidence angle.

- **TM modes of the interface**

The analysis of the interaction of TM modes at the interface can be performed very similar to the TE modes. The equations for the TM modes are

$$\frac{\partial H_y}{\partial z} = -\epsilon_0 n^2 \frac{\partial E_x}{\partial t} \tag{2.115}$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -\mu_0 \frac{\partial H_y}{\partial t} \tag{2.116}$$

$$\frac{\partial H_y}{\partial x} = \epsilon_0 n^2 \frac{\partial E_z}{\partial t} \tag{2.117}$$

and therefore, the magnetic field component H_y is the one that the transfer matrix is defined for. In this case,

$$H_y(z,t) = \left(A_+ e^{-jk_x x} + A_- e^{jk_x x}\right) e^{-jk_z z} \tag{2.118}$$

and the tangential component of the electric field is given by

$$E_z(x,z) = \frac{1}{j\omega\epsilon_0 n^2} \left((-jk_x)A_+ e^{-jk_x x} + (jk_x)A_- e^{jk_x x}\right) e^{-jk_z z} \tag{2.119}$$

and thus the **M** matrix is defined to relate the tangential components of the fields in each material:

$$\begin{pmatrix} H_y \\ E_z \end{pmatrix} = \mathbf{M}^{\mathbf{TM}}(x) \times \begin{pmatrix} A_+ \\ A_- \end{pmatrix} \times e^{-jk_z z} \tag{2.120}$$

where
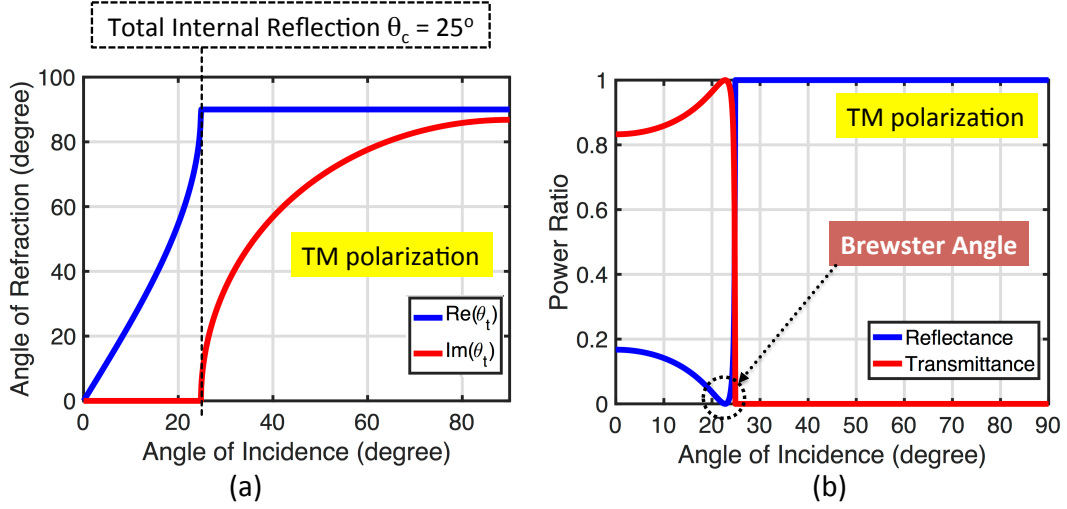
Figure 2.13: Transfer matrix method for TM modes. (a) Calculation of angle of refraction as a function of incidence angle. (b) Calculation of reflectance and transmittance as a function of incidence angle. The Brewster angle is about $22°$ and the critical angle is about $25°$.

$$\mathbf{M}^{\mathbf{TM}}(x) = \begin{bmatrix} e^{-jk_x x} & e^{jk_x x} \\ -\dfrac{k_x}{\omega\epsilon_0 n^2} e^{-jk_x x} & \dfrac{k_x}{\omega\epsilon_0 n^2} e^{jk_x x} \end{bmatrix} \tag{2.121}$$

and

$$\det(\mathbf{M}^{\mathbf{TM}}) = \frac{2k_x}{\omega\epsilon_0 n^2}. \tag{2.122}$$

The transfer matrix $\mathbf{Q}$ is given by

$$\mathbf{Q}_{1\to2}(\mathbf{x} = \mathbf{X}) = \left(\mathbf{M}_{(2)}(\mathbf{X})\right)^{-1} \times \mathbf{M}_{(1)}(\mathbf{X}) \tag{2.123}$$

and is simplified to

$$\mathbf{Q}_{1\to2}(x) = \begin{bmatrix} \dfrac{\xi_2+\xi_1}{2\xi_2}\exp(j\Delta k_x x) & \dfrac{\xi_2-\xi_1}{2\xi_2}\exp(j\Sigma k_x x) \\ \dfrac{\xi_2-\xi_1}{2\xi_2}\exp(-j\Sigma k_x x) & \dfrac{\xi_2+\xi_1}{2\xi_2}\exp(-j\Delta k_x x) \end{bmatrix} \tag{2.124}$$

where $\Delta k_x = k_x^{(2)} - k_x^{(1)}$, $\Sigma k_x = k_x^{(2)} + k_x^{(1)}$ and

$$\xi_1^{TM} = \frac{k_x^{(1)}}{\omega \epsilon_0 n_1^2} \quad , \quad \xi_2^{TM} = \frac{k_x^{(2)}}{\omega \epsilon_0 n_2^2}. \tag{2.125}$$

As can be seen the format of the **Q** matrix is exactly the same for both the TE and TM modes with the right definition of ξ factors. The reflection coefficient of the magnetic field is then given by

$$|R_{1\to2}|^2 = \left| \frac{\xi_1^{TM} - \xi_2^{TM}}{\xi_1^{TM} + \xi_2^{TM}} \right|^2 = \left| \frac{n_1 \cos(\theta_t) - n_2 \cos(\theta_i)}{n_1 \cos(\theta_t) + n_2 \cos(\theta_i)} \right|^2 \tag{2.126}$$

which exhibits a zero at an incidence angle called the *Brewster angle* calculated as

$$\theta_B + \theta_t = \frac{\pi}{2} \Rightarrow \theta_B = \tan^{-1}\left(\frac{n_2}{n_1}\right) \tag{2.127}$$

The Brewster angle always occurs before the total internal reflection. As shown in Fig. 2.13a for Si/SiO$_2$ interface at 1550 nm wavelength, the Snell's law is still exactly the same for the TM modes just as the TE modes. In other words, Snell's law does not depends on the polarization. Fig. 2.13b shows the reflectance and transmittance for the TM modes. The incidence angle at which the reflection becomes zero is the Brewster angle ($\approx 22°$).

2.4 Analytic Method for 3D Waveguides

In practice, the structure of a fabricated silicon waveguide has boundaries in both x and y directions as shown in Fig. 2.14a. Therefore, a real waveguide cannot be simply treated as a 2D structure. However, an approximation can be made if the index profile of the waveguide is viewed as the one shown in Fig. 2.14b [128]. The four regions at the corners are ruled out of the index profile with the assumption that the light does not exist in those regions. This leads to the analysis of the *effective index method* [202] in which the

Helmholtz equation is separately solved for both x and y direction with the solutions of the following type:

$$\text{Region I:} \quad \cos(k_x x)\cos(k_y y) \tag{2.128}$$

$$\text{Region II:} \quad \exp(\gamma_l x)\cos(k_y y) \tag{2.129}$$

$$\text{Region III:} \quad \exp(-\gamma_r x)\cos(k_y y) \tag{2.130}$$

$$\text{Region IV:} \quad \cos(k_x x)\exp(-\gamma_{cl} y) \tag{2.131}$$

$$\text{Region V:} \quad \cos(k_x x)\exp(\gamma_s y) \tag{2.132}$$

However, such solutions cannot satisfy the boundary conditions at all the four boundaries for both electric field and magnetic field. Therefore, the boundary condition for either E field or H field is satisfied in order to find the k_x and k_y values. The effective index can then be calculated based on the propagation condition in the core region:

$$k_0^2 n_{core}^2 = k_x^2 + k_y^2 + (k_0 n_{\text{eff}})^2. \tag{2.133}$$

The k_x factor can be approximated by analyzing the 2D slab waveguide whose boundaries are in the x direction:

$$k_0^2 n_{core}^2 = (k_0 n_{\text{eff, x}})^2 + k_x^2 \tag{2.134}$$

and similarly k_y can be approximated by analyzing the 2D slab waveguide whose boundaries are in the y direction:

$$k_0^2 n_{core}^2 = (k_0 n_{\text{eff, y}})^2 + k_y^2. \tag{2.135}$$

By combining these equations, the effective index of the 3D waveguide can be approximated based on the effective indices of two 2D dielectric slabs in the x and y directions:
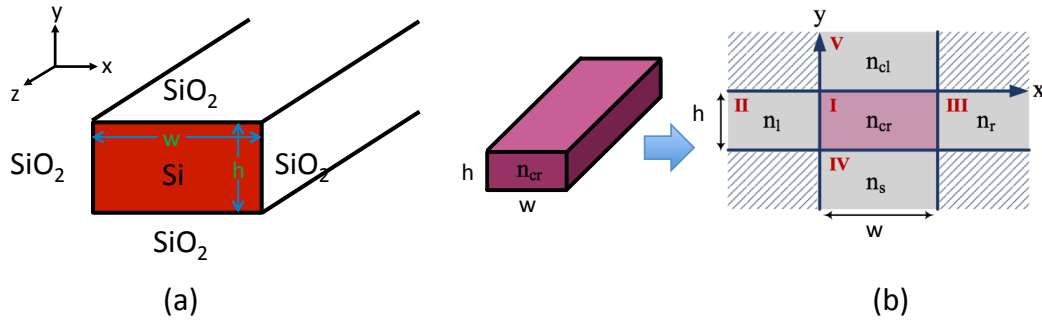
Figure 2.14: Schematic of 3D strip waveguide. (a) Core is Si and surrounding is SiO$_2$. (b) Decomposition of index profile into x and y directions for the effective index [128].

$$n_{\text{eff}} \approx \sqrt{n_{\text{eff, x}}^2 + n_{\text{eff, y}}^2 - n_{core}^2}. \tag{2.136}$$

Throughout the rest of this section, we always assume that the core material is Si and the surrounding is SiO$_2$.

3D Strip Waveguide

The strip waveguide is a rectangular waveguide whose core is Si which is completely surrounded by SiO$_2$. The schematic of the 3D strip waveguide is shown in Fig. 2.14a. Figure 2.14b demonstrates the decomposition of the index profile in both x and y direction [128].

- **quasi-TE^z Modes of the 3D Strip Waveguide**

Since the boundary conditions at all four boundaries cannot be satisfied with either TE^z or TM^z waves, in general the modes of a strip waveguides are hybrid. However, for a class of modes the E_z component is much weaker than the E_x and E_y components, while the H_z component is comparable to H_x or H_y. This class of modes is called quasi-TE^z since these modes resemble TE modes. As shown in Fig. 2.15, the process of estimating the effective index of the optical mode for the strip waveguide consists of breaking the strip waveguide into two slab waveguides in the x and y directions. For the case of TE^z modes, it can be
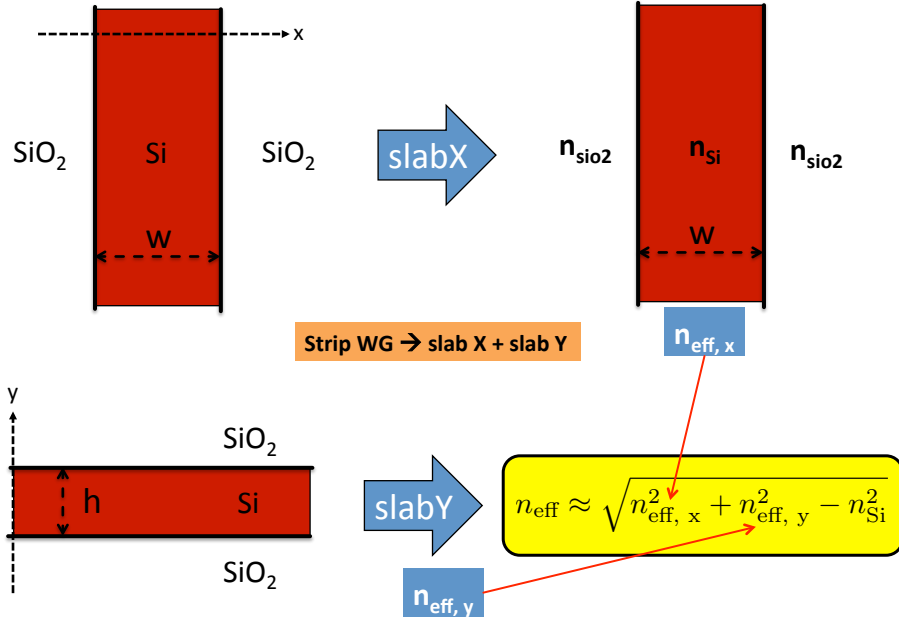
Figure 2.15: Process of the effective index method for the 3D strip waveguide. The 2D index profile of the cross-section is approximately decomposed into two 1D index profiles. SlabX has its boundaries in the x direction, while SlabY has its boundaries in the y direction.

shown that satisfying the boundary conditions of E-field in the x direction results in an eigen mode equation for the TM modes of a slab of width w, and satisfying the boundary conditions in the y direction results in an eigen mode equation for the TE modes of a slab of width h [202]. By finding the $n_{\text{eff, x}}$ and $n_{\text{eff, y}}$ in both directions, the effective index of the mode can be estimated.

$$\text{SlabX} \Rightarrow \text{TM mode} \rightarrow n_{\text{eff, x}} \rightarrow k_x \tag{2.137}$$

$$\text{SlabY} \Rightarrow \text{TE mode} \rightarrow n_{\text{eff, y}} \rightarrow k_y \tag{2.138}$$

We performed a finite element solution of the fundamental TE^z mode of the standard strip waveguide with dimensions 450 nm $\times$ 220 nm in COMSOL and the comparison with the approximate analytic mode solver are shown in Fig. 2.16a. As can be observed, the results of our analytic method are very close to the results of finite element analysis. In fact, the error is less than 0.5%. Next, the wavelength was set to 1550 nm and the

width of the waveguide was varied from 300 nm up to 1000 nm while monitoring the effective index of the fundamental TE_{00} mode. The results of comparison between the analytic mode solver and COMSOL are presented in Fig. 2.16b. A very good agreement is observed, further confirming that the underlying assumption of the mode being quasi-TE is correct. The mode analysis of the rectangular waveguide was also performed in Lumerical MODE software with 5 nm grid resolution. Figure 2.16c compares the results of Lumerical with the analytic approach for the fundamental TE_{00} mode. A good agreement is again observed. Figure 2.16d shows the comparison of group index.

Next, we sweep the width and the height of the waveguide over 300 nm–1000 nm and 100 nm–300 nm, respectively and find the effective index. Fig. 2.17 shows the contours as a result of this sweeping. The highlighted area is the region of operation for the typical silicon waveguides whose effective index is between 2.2 and 2.5. Figure 2.18 shows the developed graphical interface for the implemented analytic mode solver fo silicon strip waveguides. The effective index and group index are calculated based on the wavelength, width, and height of the waveguide. Finally, Fig. 2.19a shows the electric and magnetic field profiles of the TE_{00} mode of the rectangular strip waveguide.

- **quasi-TM^z Modes of the 3D Strip Waveguide**

For this class of modes, the H_z component is ignored compared to H_x and H_y components, while E_z cannot be ignored compared to either E_x or E_y. The approximate analytic mode analysis results in the breakdown of the strip waveguide in SlabX and SlabY. In this case, SlabX has a width of w and $n_{\text{eff, x}}$ comes from the TE modes of the SlabX, while SlabY has a width of h and $n_{\text{eff, y}}$ comes from the TM modes of the SlabY. The effective index of the quasi-TM^z modes can then be calculated based on these two parameters.

$$\text{SlabX} \Rightarrow \text{TE mode} \rightarrow n_{\text{eff, x}} \rightarrow k_x \tag{2.139}$$

$$\text{SlabY} \Rightarrow \text{TM mode} \rightarrow n_{\text{eff, y}} \rightarrow k_y \tag{2.140}$$
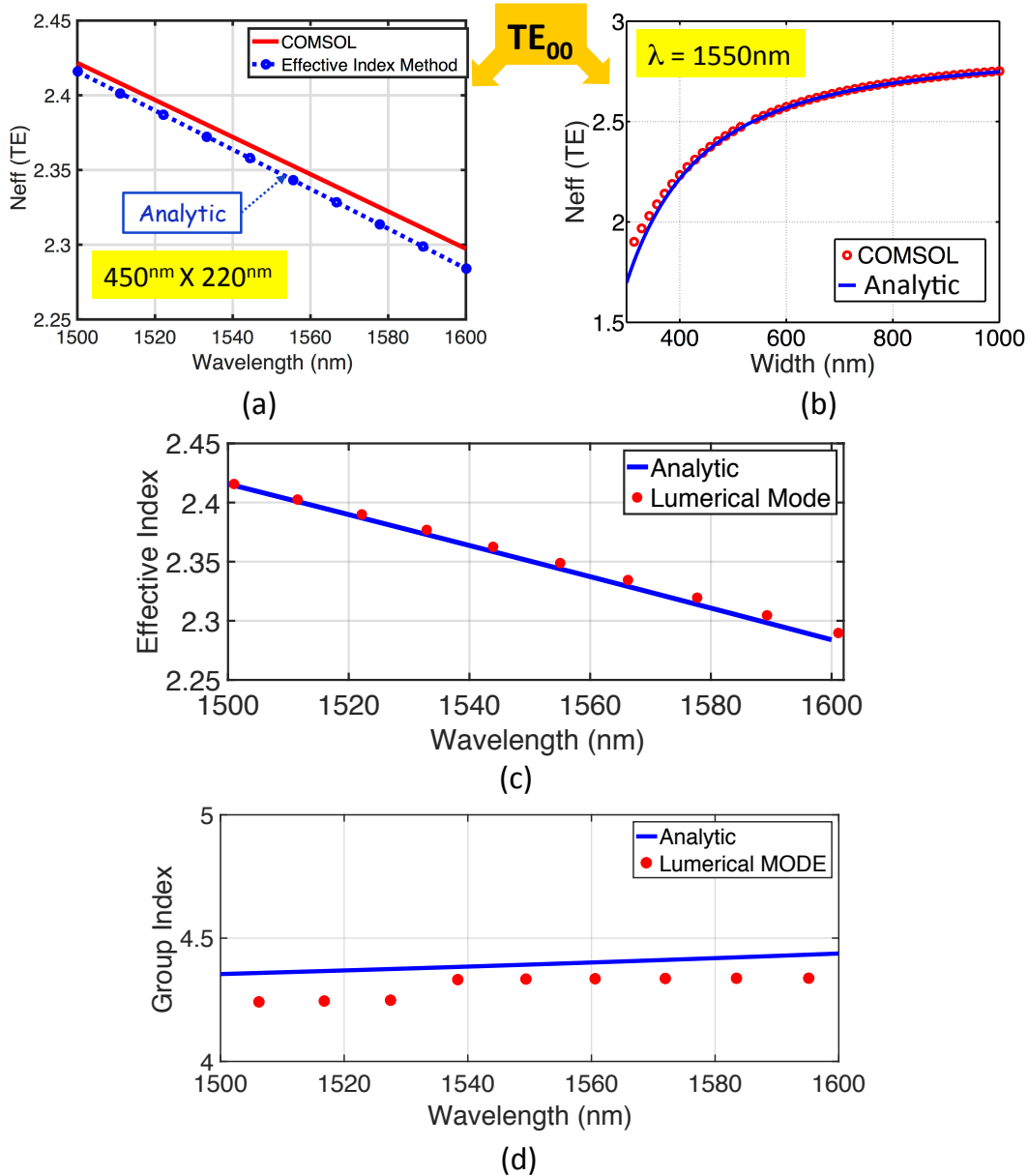
Figure 2.16: Analysis of quasi-TE mode of the Strip Waveguide. (a) Effective index as a function of wavelength. Good agreement between COMSOL and analytic mode solver is observed. (b) n_{eff} as a function of waveguide width w at 1550nm wavelength. Very good agreement is observed. (c) Comparison of the effective index analysis from Lumerical MODE software (circles) and the analytic approach (solid line). (d) Comparison of group index. The error rate is higher than the effective index.

Figure 2.19a shows the transverse components of the electric and magnetic fields (E_x, E_y, H_x, H_y) for the fundamental quasi-TE mode (TE_{00}) for the strip waveguide. It is observed that the optical power is mainly carried by E_x and H_y components. Figure 2.19b shows
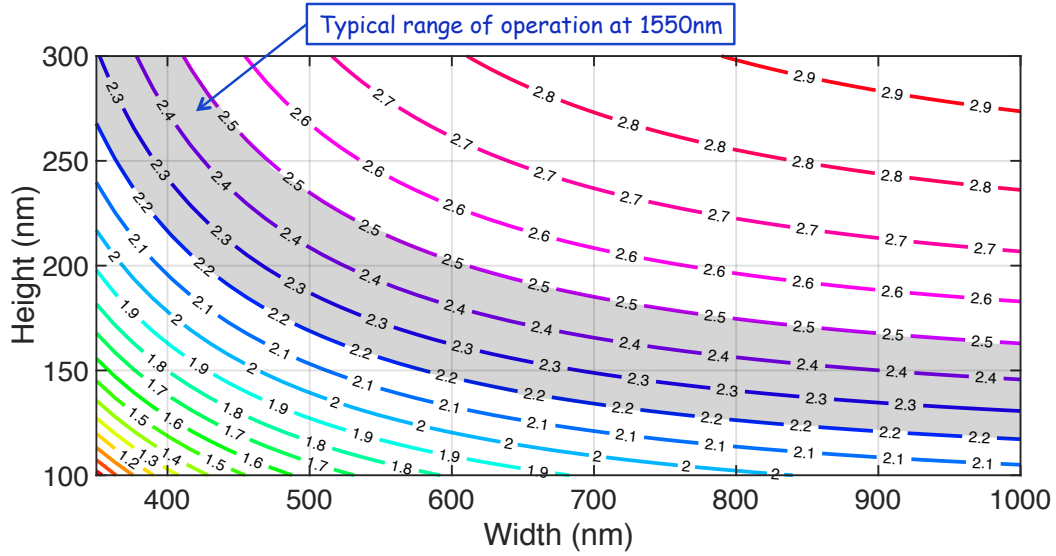
Figure 2.17: Contours of the effective index of the fundamental TE$_{00}$ mode of the strip waveguide as a function of width and height.
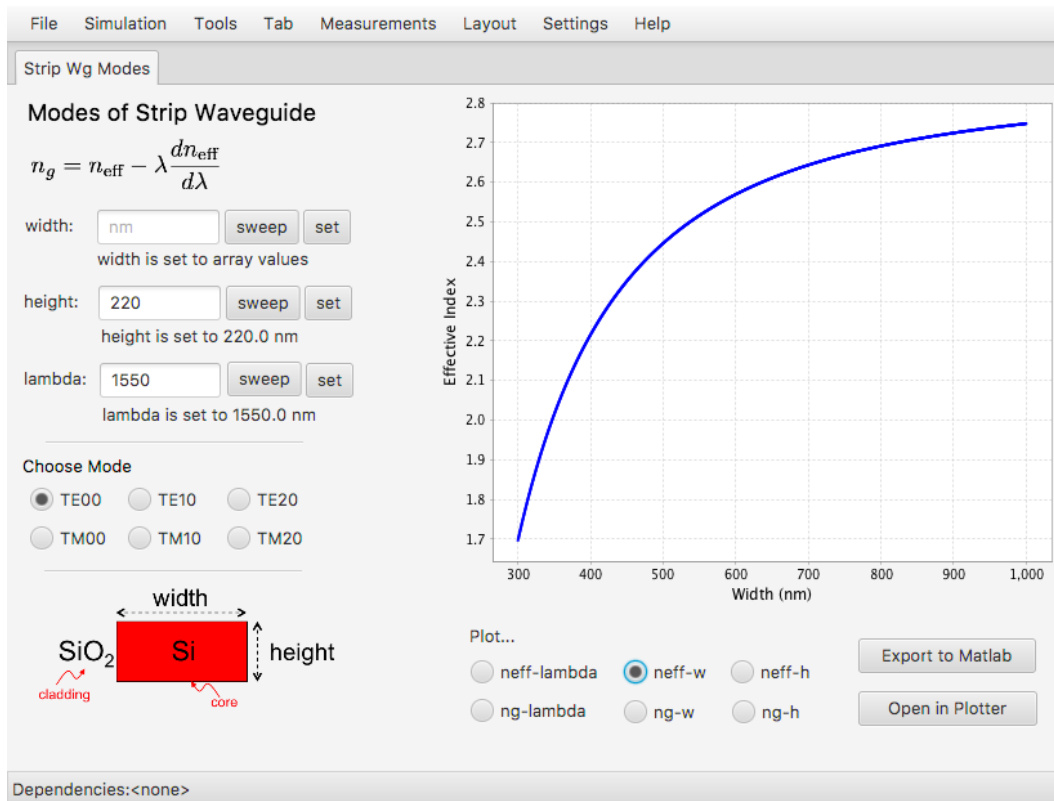
Figure 2.18: Developed graphical interface for calculating the effective index and group index of the silicon strip waveguides.
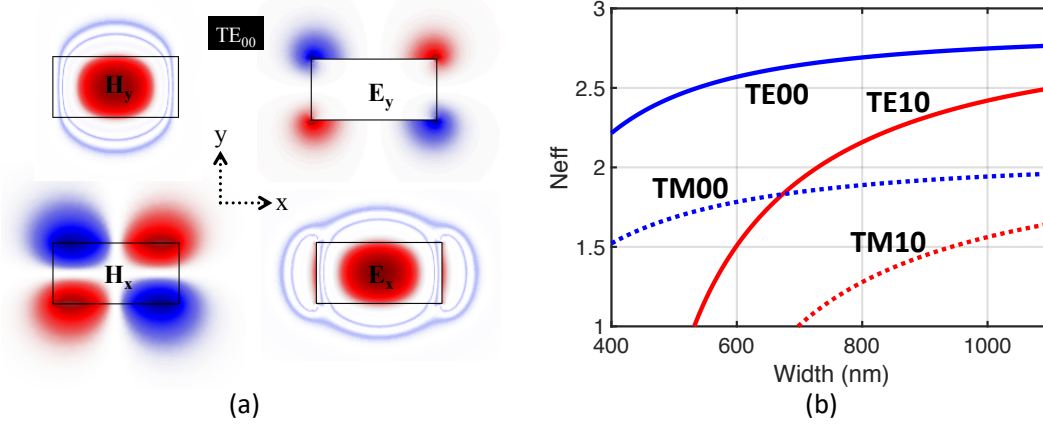
Figure 2.19: (a) field profile for the fundamental quasi-TE$_{00}$ mode of the strip waveguide. Power is carried by E_x and H_y components. (b) Effective index of the first two TE and TM modes of the strip waveguide as a function of waveguide width. More details of the mode cross-over point is presented in [202].

the effective index of TE$_{00}$, TE$_{10}$, TM$_{00}$ and TM$_{10}$ as a function of width of the waveguide for $h = 220$ nm. These curves are calculated by the analytic approach.

Sensitivity Parameters of Strip Waveguide

It can be conferred from the presented analytic method that any change in the width of the waveguide, w, or height of the waveguide, h, or index of silicon n_{si}, or index of silica, n_{sio_2}, can cause a change in the effective index of the mode inside the waveguide at a particular wavelength. Assume that χ is an external factor (e.g. temperature) that somehow affects these four parameters. The change in the effective index is then estimated as [202]:

$$\frac{dn_{\text{eff}}}{d\chi} = \frac{\partial n_{\text{eff}}}{\partial n_{si}} \times \frac{dn_{si}}{d\chi} + \frac{\partial n_{\text{eff}}}{\partial n_{sio_2}} \times \frac{dn_{sio_2}}{d\chi} + \frac{\partial n_{\text{eff}}}{\partial w} \times \frac{dw}{d\chi} + \frac{\partial n_{\text{eff}}}{\partial h} \times \frac{dh}{d\chi} \qquad (2.141)$$

where the first factor of each term on the right side of this equation determines the sensitivity of the optical mode to that parameter. For the standard strip waveguide with the dimensions 450 nm $\times$ 220 nm, we can use the analytic method and COMSOL to estimate these factors for the fundamental quasi-TE mode. The results are plotted in Fig. 2.20.
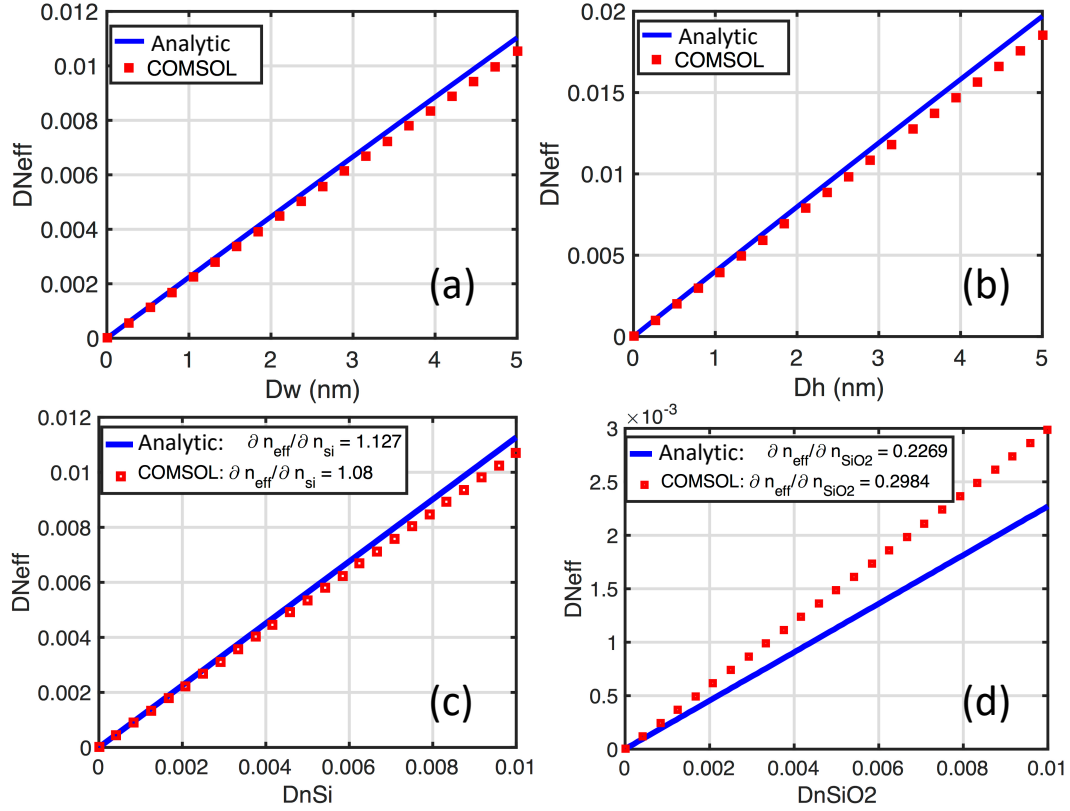
Figure 2.20: Sensitivity of the fundamental quasi-TE$_{00}$ mode of the strip waveguide. (a) Sensitivity to the change in width. (b) Sensitivity to the change in height. (c) Sensitivity to the change in silicon index. (d) Sensitivity to the change in Silica index.

As can be seen, a linear behavior is observed with good agreement between the analytic method and COMSOL simulations. These sensitivity factors are the slopes of each plot, given by:

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial n_{si}} \approx 1.127 \tag{2.142}$$

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial n_{sio_2}} \approx 0.2269 \tag{2.143}$$

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial w} \approx 0.0022 \ nm^{-1} \tag{2.144}$$

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial h} \approx 0.0039 \ nm^{-1} \tag{2.145}$$

which clearly shows that sensitivity of the optical mode is highly dependent on the index of core (silicon). As will be discussed in later chapters, this gives the motivation to alter

the properties of the optical signal that is traveling down the waveguide by manipulating its core index. Thermo-optic (TO) effects and electro-optic (EO) effects will be discussed separately. Inspection of Fig. 2.20 shows that there is a noticeable error between the analytic approach and COMSOL for the impact of the cladding index. However, this may not be an issue because EO effects usually don't impact the cladding and TO effects have an extremely weak on the cladding. Typically, the perturbations of cladding index can be ignored in the calculations.

3D Rib Waveguide

The rib (ridge) waveguide is similar to the strip waveguide with the addition of a thin silicon slab below the main rectangular core. The structure of a rib waveguide is shown in Fig. 2.21a. The width of the core is w, the height of the core is h_{wg} and the height of the slab is h_{slab}. The modes of the rib waveguide are similarly classified into quasi-TE and quasi-TM modes, with the fundamental mode as quasi-TE$_{00}$ mode. The field profile of this mode is shown in Fig. 2.21b. As can be seen, the optical power is carried by E_x and H_y components.
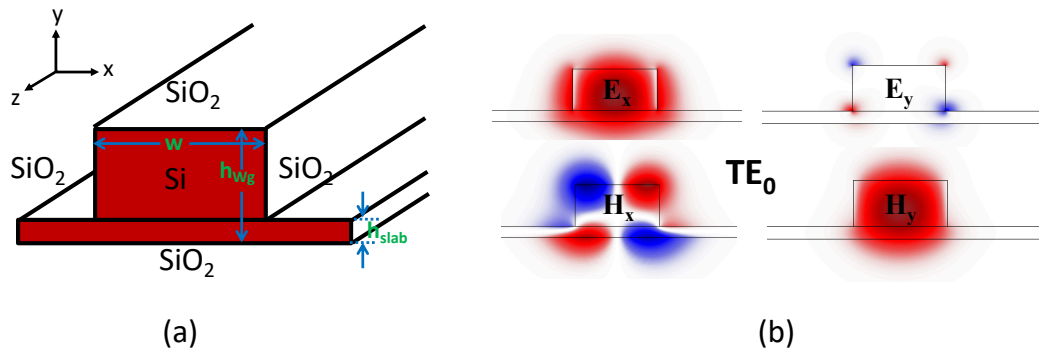
Figure 2.21: (a) Schematic of the 3D Rib waveguide. The rectangular core has a height of h_{wg} and the slab height is denoted by h_{slab}. (b) Field profile of the transverse components of quasi-TE$_{00}$ mode of the rib waveguide.

- **quasi-TE^z Modes of the 3D Rib Waveguide**

Similar to the strip waveguide, since the boundary conditions at all four boundaries cannot be satisfied with either TE^z or TM^z waves, in general the modes of a rib waveguide are hybrid. However, for a class of modes the E_z component is much weaker than the E_x and E_y components, while the H_z component is comparable to H_x or H_y. This class of modes is called quasi-TE^z since these modes resemble TE modes. As shown in Fig. 2.22, the process of estimating the effective index of the optical mode for the rib waveguide consists of breaking the rib waveguide into two slab waveguides in the x and y directions. For the case of TE^z modes, satisfying the boundary conditions of E-field in the x direction approximately results in an eigen mode equation for the TM modes of a slab of width w with a core of Si and substrate and cladding layers with an effective index of the TE mode of the side slab of width h_{slab}. In the same way, satisfying the boundary conditions in the y direction results in an eigen mode equation for the TE modes of a slab of width h_{wg}. By finding the $n_{\text{eff, x}}$ and $n_{\text{eff, y}}$ in both directions, the effective index of the mode can be estimated.

$$\text{SlabX} \Rightarrow \text{TM mode} \rightarrow n_{\text{eff, x}} \rightarrow k_x \tag{2.146}$$

$$\text{SlabY} \Rightarrow \text{TE mode} \rightarrow n_{\text{eff, y}} \rightarrow k_y \tag{2.147}$$

We use this analytic approach to estimate the effective index of the fundamental quasi-TE_{00} mode of the rib waveguide with the dimensions of 450 nm $\times$ 220 nm $\times$ 50 nm as a function of wavelength. The results and comparison with COMSOL are shown in Fig. 2.23a. The maximum error is less than 3%, which indicates a reasonable agreement. The almost constant error over the entire wavelength range is indicative of the fact that the error is most likely not due to the waveguide or material dispersion.

Next, the width of the rib waveguide is swept from 300nm to 1000nm while other parameters and the wavelength are fixed. The comparison with COMSOL results are shown in Fig. 2.23b. One possible explanation for the increase of error as the width decreases is
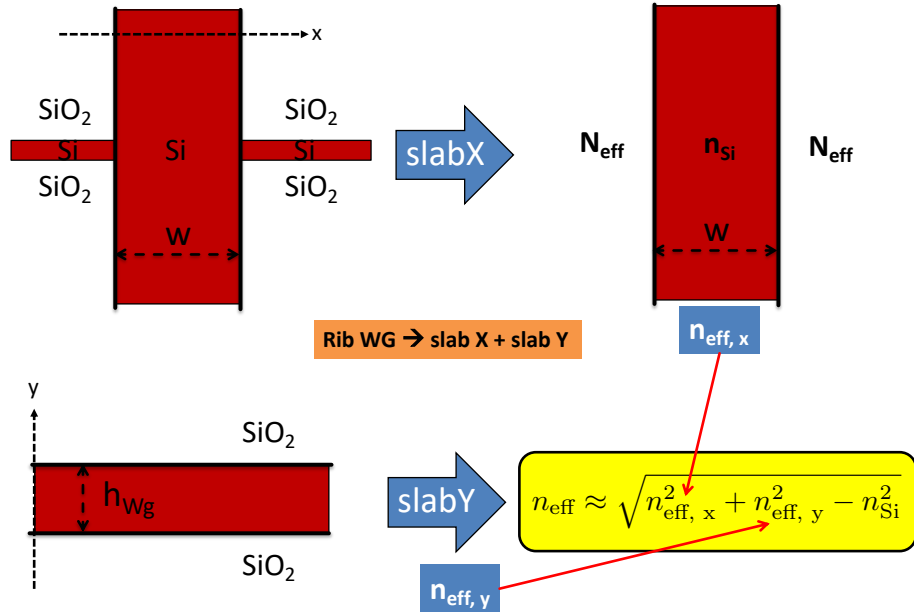
Figure 2.22: Process of the effective index method for the 3D Rib waveguide. The 2D index profile of the cross-section is approximately decomposed into two 1D index profiles. SlabX has its boundaries in the x direction, while SlabY has its boundaries in the y direction. Note the subtle difference in the SlabX in this case with the case for the strip waveguide.

that this analytic approach underestimates the impact of cladding index, which gets worse for smaller widths since it results in less confinement of light inside the silicon core.

- **quasi-TM^z Modes of the 3D Rib Waveguide**

Similar to the quasi-TM modes of the Rib waveguide, we can estimate the effective index of the quasi-TM modes of the Rib waveguide by finding the TE modes of the SlabX and TM modes of the SlabY:

$$\text{SlabX} \Rightarrow \text{TE mode} \rightarrow n_{\text{eff, x}} \rightarrow k_x \tag{2.148}$$

$$\text{SlabY} \Rightarrow \text{TM mode} \rightarrow n_{\text{eff, y}} \rightarrow k_y \tag{2.149}$$

Figure 2.24a shows the results of the mode analysis for the quasi-TE modes as a function of width at 1550 nm wavelength. The cut-off point of each mode is where the effective
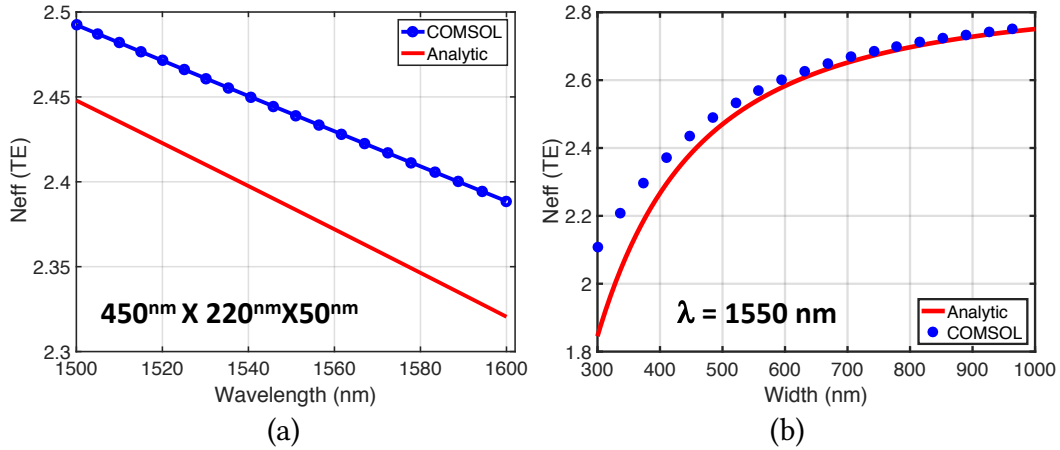
Figure 2.23: Analysis of the fundamental quasi-TE$_{00}$ mode of the Ridge waveguide. (a) Effective index as a function of wavelength. (b) Effective index as a function of width.
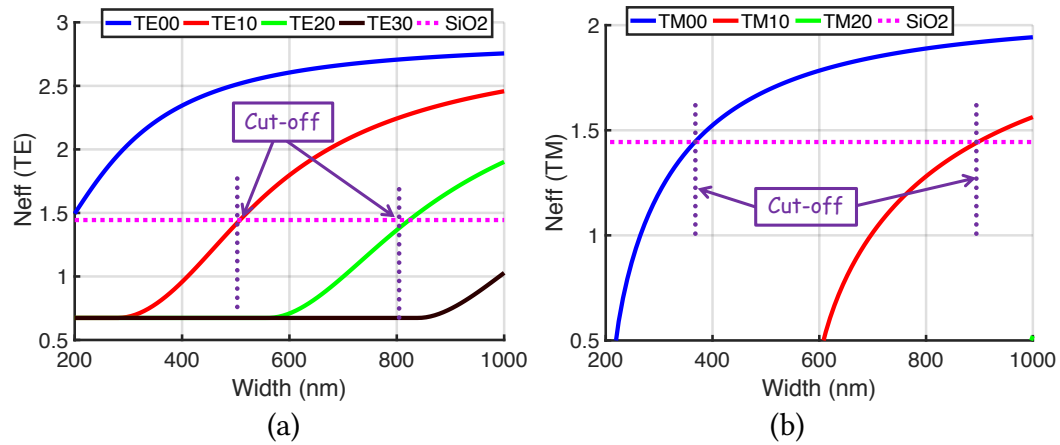
Figure 2.24: (a) Analysis of the effective index of the quasi-TE modes of the Rib waveguide. Up to 500nm of width is single mode operation. (b) Effective index of the quasi-TM modes of the rib waveguide.

index of the mode is lass than the index of SiO$_2$. It can be seen that the fundamental TE$_{00}$ mode is well above cut-off even if the width of the waveguide is 200 nm. At 450 nm width, the TE$_{10}$ is weakly guiding, therefore the TE$_{00}$ mode is the main guiding mode. Figure 2.24b shows the cut-off of the quasi-TM modes of the rib waveguide. It can be seen that up to 500 nm of width, only the TM$_{00}$ mode is above cut-off. Therefore, for a standard 450 nm width rib waveguide, the control of launched polarization is necessary to ensure single mode and polarization operation.

Sensitivity Parameters of Rib Waveguide

Similar to the case for the rectangular strip waveguide, we consider the fundamental quasi-TE mode of the rib waveguide with the dimensions of 450 nm × 220 nm × 50 nm and add some variations to the parameters of the waveguide and see how the effective index changes. Note that we set the variation of width and height to about 5 nm which is the typical variation observed in the current foundries [90]. The sensitivity parameters are calculated as:

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial n_{si}} \approx 1.1261 \text{ (COMSOL: 1.0237)} \tag{2.150}$$

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial n_{sio_2}} \approx 0.2056 \text{ (COMSOL: 0.2514)} \tag{2.151}$$

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial w} \approx 0.002 \ nm^{-1} \text{ (COMSOL: } 0.0016 \ nm^{-1}) \tag{2.152}$$

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial h_{wg}} \approx 0.0039 \ nm^{-1} \text{ (COMSOL: } 0.0031 \ nm^{-1}) \tag{2.153}$$

$$\text{TE}_{00}: \quad \frac{\partial n_{\text{eff}}}{\partial h_{slab}} \approx 0.0012 \ nm^{-1} \text{ (COMSOL: } 0.002 \ nm^{-1}) \tag{2.154}$$

which shows that the sensitivity is strongest for the core (silicon) index. The sensitivity to the height of the rectangular core is higher than the sensitivity to the width, while the mode has the least sensitivity to the height of the slab.

2.5   Perturbation Theory for Waveguides

Although we directly calculated the sensitivity of the optical mode using a mode solver approach, it is also possible to apply the perturbation theory to find the change in the wave vector of the optical mode [63]. This change can be obtained from the complex electromagnetic variational theorem and the result is given by

$$\Delta\beta = \frac{1}{c}\Delta(\omega n_{\text{eff}}) = \frac{\iint \left[\Delta(\omega\epsilon_0 n^2)\mathbf{E}\cdot\mathbf{E}^* + \Delta(\omega\mu)\mathbf{H}\cdot\mathbf{H}^*\right] dx\, dy}{2\iint \text{Re}[\mathbf{E_t} \times \mathbf{H_t^*}].\hat{\mathbf{z}}\, dx\, dy} \tag{2.155}$$

for a waveguide whose materials exhibit dispersion (n is the refractive index of material).
If the refractive index of the core or cladding changes ($n \rightarrow n + \Delta n$) at a given wavelength,
then

$$\Delta\beta = \frac{\omega}{c}\Delta n_{\text{eff}} \approx \omega\epsilon_0 \frac{\iint [n\,\Delta n \mathbf{E} \cdot \mathbf{E}^*]\,dx\,dy}{\iint \text{Re}[\mathbf{E_t} \times \mathbf{H_t^*}].\hat{\mathbf{z}}\,dx\,dy} . \tag{2.156}$$

Using this approximate equation, we can find the sensitivity of the optical mode to the
change in the index of the silicon core:

$$\Delta\beta = \frac{\omega}{c}\Delta n_{\text{eff}} \approx \omega \frac{\iint_{\text{core}} [\epsilon_0 n^{si}\,\Delta n^{si}\,\mathbf{E} \cdot \mathbf{E}^*]\,dx\,dy}{\iint \text{Re}[\mathbf{E_t} \times \mathbf{H_t^*}].\hat{\mathbf{z}}\,dx\,dy} . \tag{2.157}$$

By some simple manipulation and using the definition of the group index of the mode as
the ratio between stored energy and propagating power, the following relation is reached:

$$\Delta\beta = \frac{\omega}{c}\Delta n_{\text{eff}} \approx \frac{\omega}{c} \Gamma_{\text{core}} \left(\frac{n_g}{n_g^{si}}\right) \Delta n^{si} \tag{2.158}$$

or equivalently

$$\Delta n_{\text{eff}} \approx \Gamma_{\text{core}} \left(\frac{n_g}{n_g^{si}}\right) \Delta n^{si} . \tag{2.159}$$

Note that n_g is the group index of the optical mode whereas n_g^{si} is the group index of
silicon material. Γ_{core} is related to the confinement of the optical modal energy density
inside the core and is defined by

$$\Gamma_{\text{core}} = \frac{\iint_{\text{core}} \left[\epsilon_0 n^{si}\,n_g^{si}\,\mathbf{E} \cdot \mathbf{E}^*\right]\,dx\,dy}{\iint \left[\epsilon_0 n\,n_g^M\,\mathbf{E} \cdot \mathbf{E}^*\right]\,dx\,dy} \tag{2.160}$$

where n_g^M refers to the group index of material in either core or cladding region. Clearly,
$0 < \Gamma_{\text{core}} < 1$. If we ignore the dispersion of the materials (e.g. $n^{si} = 3.477$ and $n_g^{si} = 3.605$, hence $n^{si} \approx n_g^{si}$. The error is even less for SiO_2), then Γ_{core} is exactly the fraction
of the modal energy density inside the core. If we further define the confinement of the

intensity of electric field inside the core as

$$\Gamma_E = \frac{\iint_{\text{core}} \mathbf{E} \cdot \mathbf{E}^* dx\, dy}{\iint \mathbf{E} \cdot \mathbf{E}^* dx\, dy} = \frac{\iint_{\text{core}} \mathbf{E} \cdot \mathbf{E}^* dx\, dy}{\iint_{\text{core}} \mathbf{E} \cdot \mathbf{E}^* dx\, dy + \iint_{\text{clad}} \mathbf{E} \cdot \mathbf{E}^* dx\, dy} \tag{2.161}$$

then we see that

$$\Gamma_{\text{core}} = \frac{n^{si}\, n_g^{si}\, \Gamma_E}{n^{si}\, n_g^{si}\, \Gamma_E + n^{sio2}\, n_g^{sio2}\, (1 - \Gamma_E)} . \tag{2.162}$$

Similarly, the modal energy confinement factor of the cladding can be written as

$$\Gamma_{\text{clad}} = \frac{n^{sio2}\, n_g^{sio2}\, (1 - \Gamma_E)}{n^{si}\, n_g^{si}\, \Gamma_E + n^{sio2}\, n_g^{sio2}\, (1 - \Gamma_E)} . \tag{2.163}$$

and as expected we have $\Gamma_{\text{core}} + \Gamma_{\text{clad}} = 1$. Note that the definition of these confinement factors is different from the definition of the power confinement factor of a waveguide.

If both index of the core and index of the cladding experience some small perturbation, then

$$\boxed{\Delta n_{\text{eff}} \approx \Gamma_{\text{core}} \left(\frac{n_g}{n_g^{si}} \right) \Delta n^{si} + \Gamma_{\text{clad}} \left(\frac{n_g}{n_g^{sio_2}} \right) \Delta n^{sio_2}} \tag{2.164}$$

For example, for the quasi-TE$_{00}$ mode of a 450×220 nm Strip waveguide, the E-field intensity confinement factor of the core at $\lambda = 1550$ nm is about 0.57 (i.e. $\Gamma_E \approx 0.57$—this number is obtained from Lumerical Mode package), the group index of silicon is about 3.605, the group index of silica is 1.462, and the group index of the optical mode is about 4.393. Therefore, $\Gamma_{\text{core}} \approx 0.8873$, and the sensitivity of the optical mode to the index of core is

$$\frac{\partial n_{\text{eff}}}{\partial n_{si}} \approx \frac{\Delta n_{\text{eff}}}{\Delta n^{si}} \approx \Gamma_{\text{core}} \left(\frac{n_g}{n_g^{si}} \right) = 0.8873 \times \frac{4.393}{3.605} = 1.0813 \tag{2.165}$$

which is very close to the results of Fig. 2.20c. Note that for this waveguide, the confinement factor of optical power is 0.766 which is different from the sensitivity factor of the optical mode. The change in the effective loss factor of the optical mode $\Delta \alpha_{\text{eff}}$ can be estimated the same way in the presence of material loss ($\Delta \alpha^M$) using the perturbation

theory. To do so, assume that the loss of the silicon core is changed by $\Delta\alpha^{si}$. This means that the index of silicon now gets an imaginary change

$$\Delta n^{si} \to -j\frac{\lambda}{4\pi}\Delta\alpha^{si} \tag{2.166}$$

and therefore $\Delta n_{\text{eff}} \to -j\Delta\alpha_{\text{eff}} \times \lambda/4\pi$ and

$$\Delta\alpha_{\text{eff}} \approx \Gamma_{\text{core}}\left(\frac{n_g}{n_g^{si}}\right)\Delta\alpha^{si}. \tag{2.167}$$

In general, we can see that the sensitivity factors of the optical mode also apply to the loss factor. Therefore

$$\boxed{\Delta\alpha_{\text{eff}} \approx \Gamma_{\text{core}}\left(\frac{n_g}{n_g^{si}}\right)\Delta\alpha^{si} + \Gamma_{\text{clad}}\left(\frac{n_g}{n_g^{sio_2}}\right)\Delta\alpha^{sio_2}} \tag{2.168}$$

or equivalently

$$\Delta\alpha_{\text{eff}} \approx \left(\frac{\partial n_{\text{eff}}}{\partial n_{si}}\right)\Delta\alpha^{si} + \left(\frac{\partial n_{\text{eff}}}{\partial n_{sio_2}}\right)\Delta\alpha^{sio_2}. \tag{2.169}$$

This result can also be obtained simply by applying the Kramers-Kronig (KK) relations which is a direct result of the *causality* of the material and the waveguide. Noting that the Kramers-Kronig relations are linear transformations, we can write:

$$\Delta\alpha_{\text{eff}} = \text{KK}[\Delta n_{\text{eff}}] \approx \text{KK}\left[\left(\frac{\partial n_{\text{eff}}}{\partial n_{si}}\right)\Delta n^{si} + \left(\frac{\partial n_{\text{eff}}}{\partial n_{sio_2}}\right)\Delta n^{sio_2}\right] \tag{2.170}$$

and therefore

$$\Delta\alpha_{\text{eff}} = \left(\frac{\partial n_{\text{eff}}}{\partial n_{si}}\right)\text{KK}\left[\Delta n^{si}\right] + \left(\frac{\partial n_{\text{eff}}}{\partial n_{sio_2}}\right)\text{KK}\left[\Delta n^{sio_2}\right] \tag{2.171}$$

and since $\text{KK}\left[\Delta n^{si}\right] = \Delta\alpha^{si}$ and $\text{KK}\left[\Delta n^{sio_2}\right] = \Delta\alpha^{sio_2}$, the same result is reached.

2.6 Summary of the Chapter

In this chapter, the main properties of the silicon photonic waveguides, namely the effective index and group index of the optical modes, were described and models were presented to estimate their values. The sensitivity of the optical mode to the dimensions of the waveguide and the change in the material index (silicon core and silica cladding) were derived for strip and ridge waveguides. These sensitivity parameters will play an important role in analyzing the thermo-optic and electro-optic response of the optical mode inside the waveguide. This is the topic of the next chapters. The transfer matrix of the interface between two dielectric materials was derived for both TE and TM polarization which will be used later to calculate the coupled modes (supermodes) between two dielectric slab waveguides.

Chapter 3

Compact Modeling of Silicon Photonic Circuits

Abstract– *In this chapter, we present compact models for the basic building blocks for the silicon photonic circuits made with the SOI waveguides. The effective index of the optical mode can be obtained based on a prior mode analysis and the port-to-port transfer functions are characterized based on the evolution of amplitude and phase of light. Two circuit solver methods based on iteration and signal flow graph (SFG) are presented for calculating the transfer function of complex structures. A variational optimization approach is presented for designing the optimal waveguide bends (e.g. 90-degree and 180-degree bends) in silicon photonic platform. It is mathematically proved that the ratio between the loss of an optimal bend and the loss of a circular bend has a lower bound of 0.865.*

Since the propagation of light in photonic elements is governed by Maxwell's equations, the most accurate way of obtaining the transfer function of a photonic circuit is to perform a 3D numerical solution of Maxwell's equations. Methods such as finite difference time domain (FDTD) allow direct solution of Maxwell's equations; however, such solutions are typically inefficient and computationally demanding when applied to large-scale photonic circuits. To reduce the computation demands of analyzing a photonic circuit compact models [77] can be used for each individual photonic element and a circuit solver can be used to simulate the propagation of light in the circuit [173].

Such compact modeling approach requires two main ingredients:

• **A library of compact models for basic photonic building blocks** [178].

As shown in Fig. 3.1, passive silicon photonic devices can be realized with elementary
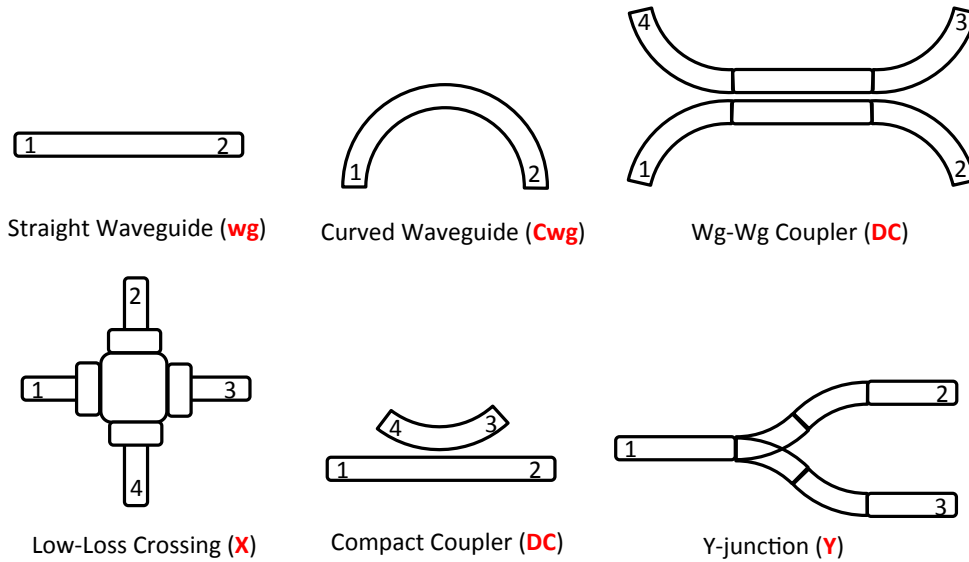
Figure 3.1: Schematic of basic building blocks of passive silicon photonic devices, including waveguides, couplers and power splitters/combiners.

components such as straight or bent waveguides, directional couplers, and power splitters. Active components can be realized by embedding thermo-optic or electro-optic phase shifters in waveguides. For example, in the case of a Mach-Zehnder modulator (MZM), a phase shifter is embedded in a straight waveguide that comprises an arm of the MZM. In the case of a microring modulator, a phase shifter is embedded in a circularly curved waveguide that comprises part of the ring.

- **A circuit solver that models the propagation of light**.

A photonic circuit can be built by connecting the ports of elementary building blocks. A photonic circuit solver is then required to model the flow of light (amplitude and phase) from a given port to another port. The basic idea is to use the scattering matrix of each elementary block and combine them to find the scattering parameter of the given port-to-port connection. In this chapter, two methods are presented to model the flow of light in a photonic interconnect:

The first method is based on *iteration* in which the amplitude and phase of light at each port is updated repeatedly until a steady solution (hence convergence) is achieved.

As we will see, this method is very efficient in cases a unidirectional propagation of light is assumed. For example, in a Mach-Zehnder switch fabric, the flow of light is typically from one input port to an output port only in one direction.

The second method is based on the complex *signal flow graph (SFG)* approach in which a closed-form solution can be obtained for the transfer matrix using Mason's graph formula. This method is particularly efficient in circuits that include feedback loops. first-order and higher-order microring resonators are examples of this kind.

3.1 Elementary building blocks

Here, we list the passive elementary building blocks that are typically used in silicon photonic circuits. These elements are shown in Fig. 3.1.

Straight Waveguide

Straight waveguides are characterized by an ideal rectangular cross-section. The width and height of the silicon core surrounded by SiO_2 cladding in conjunction with the wavelength of operation determine the effective index of the optical mode. The fundamental mode of such rectangular waveguides is the quasi-TE_{00} mode. For a waveguide of length l, the scattering matrix is given by

$$S_{wg} = \begin{bmatrix} 0 & \sqrt{L}\exp(-j\frac{2\pi}{\lambda}n_{\text{eff}}l) \\ \sqrt{L}\exp(-j\frac{2\pi}{\lambda}n_{\text{eff}}l) & 0 \end{bmatrix} \tag{3.1}$$

where n_{eff} is the effective index of the optical mode. L is the linear optical loss of the waveguide defined by the ratio of the output optical power to the input optical power, i.e.

$$L = \frac{P_{out}}{P_{in}} = \exp(-\alpha l) \tag{3.2}$$

where α is the absorption coefficient (in units of 1/m). Commonly, loss (or the absorption coefficient) of optical waveguides is expressed in units of dB/cm and is mainly tied to the roughness of the sidewalls of the waveguides. The conversion between "dB/cm" and "1/m" and "1/cm" is given by:

$$\alpha_{(1/m)} = 100\,\alpha_{(1/cm)}$$

$$\alpha_{(1/m)} = 23\,\alpha_{(dB/cm)}$$

$$\alpha_{(1/cm)} = 0.23\,\alpha_{(dB/cm)}. \tag{3.3}$$

Similarly, we can write:

$$\alpha_{(dB/m)} = 100\,\alpha_{(dB/cm)}$$

$$\alpha_{(dB/cm)} = 0.04343\,\alpha_{(1/m)}$$

$$\alpha_{(dB/m)} = 4.343\,\alpha_{(1/m)}. \tag{3.4}$$

In order to prove the relation between "dB/cm" and "1/m" units, we can directly use the definition:

$$L_{dB} = -10\log_{10}(\exp(-\alpha_{(1/m)}l)) = \alpha_{(1/m)}\,l \times (10\log_{10}(e)) \tag{3.5}$$

and since $10\log_{10}(e)$ = 4.393, the loss in dB per unit length is given by

$$L_{dB}/l = \alpha_{(dB/m)} = 4.393\,\alpha_{(1/m)} \tag{3.6}$$

and finally by noting that 1/4.393 = 0.2303, the other equations are derived as well.

In the scattering matrix modeling, the output waves of the ports are linearly related to the input waves of the ports through the scattering matrix. In the case of a straight

Straight Waveguide (**wg**)
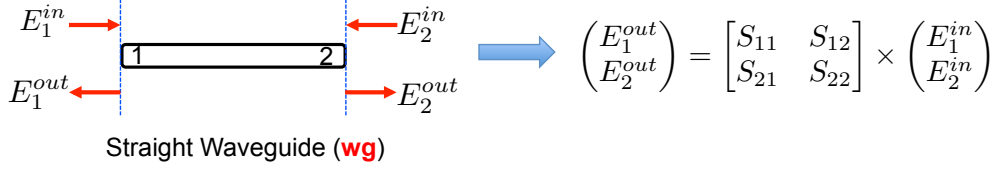
Figure 3.2: Scattering matrix of a straight waveguide. The output waves are related to the input waves through the scattering matrix.

waveguide as shown in Fig. 3.2, this relation is given by

$$
\begin{pmatrix} E_1^{out} \\ E_2^{out} \end{pmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \times \begin{pmatrix} E_1^{in} \\ E_2^{in} \end{pmatrix} . \tag{3.7}
$$

Note that the reflection at each port of the waveguide is assumed negligible in Eq. (3.1) (i.e. $S_{11} = 0$ and $S_{22} = 0$). Although studies have shown that reflections exist in silicon waveguides due to the roughness-induced backscattering [124], for simplicity we assume that there is no reflection at each port. Since a photonic circuit solver takes into account the entire scattering matrix, the assumption of no reflection will not in general impact the circuit solver.

Curved (bent) Waveguide

Similar to the straight waveguide, a curved waveguide (typically circularly curved) can be characterized by the scattering matrix

$$
S_{cwg} = \begin{bmatrix} 0 & \sqrt{L}\exp(-j\frac{2\pi}{\lambda}n_{\text{eff}}l) \\ \sqrt{L}\exp(-j\frac{2\pi}{\lambda}n_{\text{eff}}l) & 0 \end{bmatrix} \tag{3.8}
$$

where $l = R\theta$. Here, R is the radius of curvature and θ is the angle of curvature. Figure 3.3 shows the schematic of a circular waveguide and the scattering matrix description of the waves.

The loss of a curved waveguide generally consists of sidewall roughness scattering
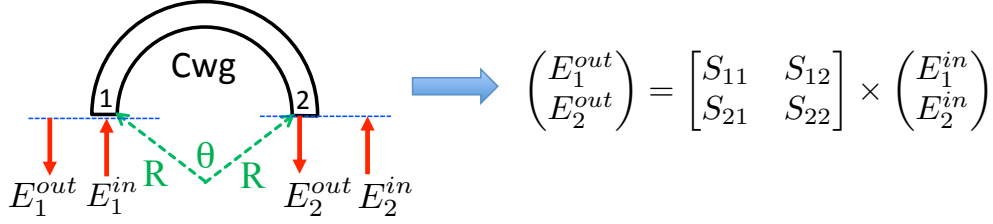
Figure 3.3: Schematic of a circularly curved waveguide. The output waves are related to the input waves through the scattering matrix.

(similar to a straight waveguide) and an additional part reflecting the radiation loss due to the curvature of the waveguide. For a differential length of the curved waveguide, the loss factor (in dB) can be written as

$$\Delta L_{(dB)} = \alpha[R(s)]_{(dB/cm)} \, \Delta s = \alpha[R(\theta)]_{(dB/cm)} \, R[\theta] \, d\theta \qquad (3.9)$$

where $R[s]$ is the local radius of curvature and $\Delta s = R[\theta] \, d\theta$ is the length of the differential part of the waveguide. The average loss factor in "dB/cm" units is then given by

$$\alpha_{0(dB/cm)} = \frac{\int \alpha[R(s)]_{(dB/cm)} \, ds}{\int ds} = \frac{\int \alpha[R(\theta)]_{(dB/cm)} \, R[\theta] \, d\theta}{\int R[\theta] \, d\theta}. \qquad (3.10)$$

For example, in the case of a circularly shaped waveguide, the radius of curvature is constant and independent of the location, therefore $R[\theta] = R_0$ and $\alpha[R(\theta)] = \alpha[R]$ and hence the average loss in dB/cm is equal to the local loss in dB/cm, i.e. $\alpha_0 = \alpha[R]$. Typically, the radiation loss of a waveguide decreases as its radius of curvature increases. For circularly bent waveguide, the loss factor can be modeled either with an exponential function of radius or an inverse power law relation. A power-law relation for the loss can take the following form:

$$\alpha[R]_{(dB/cm)} = a \, R^{-b} + c \qquad (3.11)$$

where a, b, and c are positive constants. With this model, the average loss factor in "dB/cm" for a curved waveguide of any shape can be derived from Eq. (3.10).
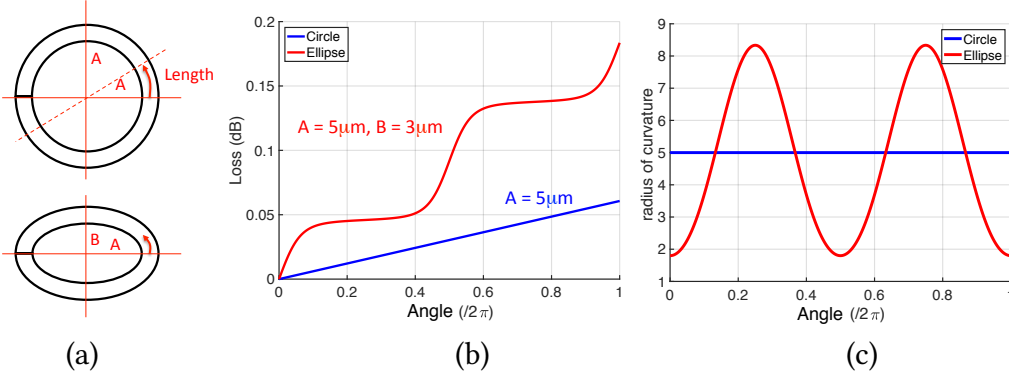
Figure 3.4: (a) Structure of a circular ring resonator and an elliptical ring resonator. (b) Comparison of estimated loss (in dB) for a ring with 5μm radius and an ellipse. (c) Comparison of the radius of curvature between a circular ring and an elliptical ring.

If a curved waveguide is parametrized such that $x = x(t)$ and $y = y(t)$, then the radius of curvature is given by

$$R(t) = \frac{(x'^2 + y'^2)^{3/2}}{|x'\,y'' - y'\,x''|},\tag{3.12}$$

and if the curve is given by $y = f(x)$, then the radius of curvature is given by

$$R(x) = \frac{[1 + (dy/dx)^2]^{3/2}}{|d^2y/dx^2|}.\tag{3.13}$$

Finally, the differential length of the curve is given by

$$\Delta s = \sqrt{x'^2 + y'^2}\,dt = \sqrt{1 + (dy/dx)^2}\,dx.\tag{3.14}$$

As an example, we examine the loss of a circular ring and an elliptical ring as shown in Fig. 3.4a. The bend loss model as a function of radius of curvature is constructed based on the measurements of a circular ring resonator. The parameters used are $a = 2096.3$, $b = 2.9123$, $c \approx 0$ for the strip waveguides with dimensions 400nm$\times$220nm [16], and R in Eq. (3.11) is given in μm. Figure 3.4b shows the loss of the curved waveguides (in dB) as a function of its curvature angle. It is seen that the ellipse whose major axis is 5μm and

72
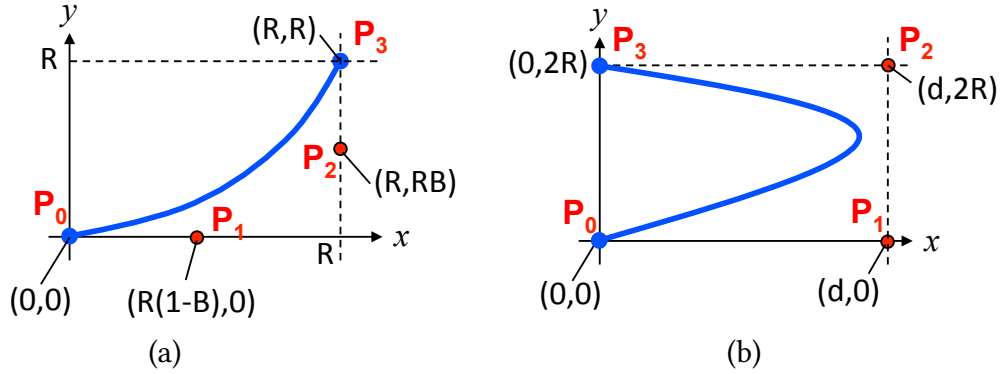
Figure 3.5: (a) 90 degree Bezier bend ($0 < B < 1$). (b) 180 degree Bezier bend.

minor axis is 3μm exhibits higher loss than a circularly curved waveguide. The reason is the variation of the radius of curvature of the ellipse at different points. As shown in Fig. 3.4, the radius of curvature varies in the range of 2μm–8μm for the ellipse, whereas it has a constant value of 5μm for the circular ring.

- **Bezier curves vs. Circular curves**

Bezier curves have been shown to provide less bending loss for silicon photonic waveguides, hence being a better choice for high quality factor race-track resonators [26, 30, 114, 115]. A Bezier curve is generally described by

$$\mathbf{P}(t) = (1 - t)^3 \, \mathbf{P_0} + 3(1 - t)^2 t \, \mathbf{P_1} + 3(1 - t)t^2 \, \mathbf{P_2} + t^3 \, \mathbf{P_3} \tag{3.15}$$

where t is the curve parameter ($0 \leq t \leq 1$). $\mathbf{P}(t) = (x(t), y(t))$ is the trajectory of the Bezier curve which clearly passes through the points $\mathbf{P_0}$ and $\mathbf{P_3}$. Typically, the Bezier curve does not pass through $\mathbf{P_1}$ and $\mathbf{P_2}$. These two points are used to determine the directionality of the curve. The first derivative of the Bezier curve is given by

$$\mathbf{P}'(t) = 3(1 - t)^2 \, (\mathbf{P_1} - \mathbf{P_0}) + 6(1 - t)t \, (\mathbf{P_2} - \mathbf{P_1}) + 3t^2 \, (\mathbf{P_3} - \mathbf{P_2}) \tag{3.16}$$

(a)

(b)

(c)

(d)
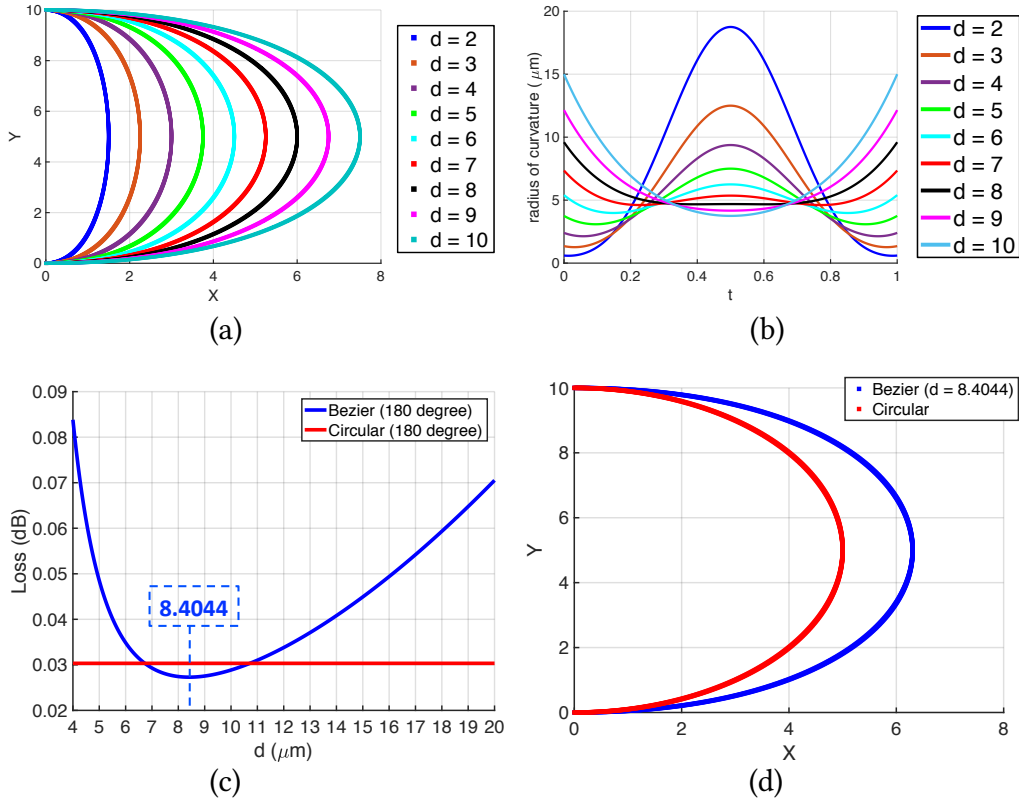
Figure 3.6: (a) Plot of 180-degree Bezier curves for $R = 5\mu$m and various d values (in μm). (b) Radius of curvature for the Bezier curves in part a. (c) Minimization of propagation loss (in dB) of the Bezier curves by changing the d parameter when R is set to 5μm. It is seen that at $d = 8.4044$, the loss is minimum and less than a fully circular 180-degree bend. (d) Plot of the optimal Bezier curve compared to the fully circular bend. X and Y axes are in μm units.

and the second derivative is given by

$$\mathbf{P}''(t) = 6(1 - t)\left(\mathbf{P_2} - 2\mathbf{P_1} + \mathbf{P_0}\right) + 6t\left(\mathbf{P_3} - 2\mathbf{P_2} + \mathbf{P_1}\right). \tag{3.17}$$

Figure 3.5a shows an example of a 90 degree Bezier curve that passes through (0,0) and (R, R) points. This curve has a dimensionless parameter B that varies between 0 and 1. For this curve: $\mathbf{P_0} = (0, 0)$, $\mathbf{P_1} = (R(1 - B), 0)$, $\mathbf{P_2} = (R, RB)$, $\mathbf{P_3} = (R, R)$.

Figure 3.5b shows an example of a 180 degree Bezier curve that passes through (0,0) and $(0, 2R)$ points. This curve has a parameter d that has a length dimension. For this curve: $\mathbf{P_0} = (0, 0)$, $\mathbf{P_1} = (d, 0)$, $\mathbf{P_2} = (d, 2R)$, $\mathbf{P_3} = (0, 2R)$. Since the 180-degree curve

is an important part of designing a race-track resonator, we examine this type of bends in more details. As an example, we assume that $R = 5\mu$m and try to find the value of d that minimizes the loss of the Bezier bend. Figure 3.6a plots Bezier bends for $R = 5\mu$m and various d values. In Fig. 3.6b we calculate and plot the radius of curvature (μm) at any point along the Bezier curves. It is seen that when $d \approx 8$, the minimum value of the radius of curvature is close to 5μm. This means that we expect to see a minimum point around 8μm. As shown in Fig. 3.6c, there is indeed a minimum for the loss at $d = 8.4044$ which is lower than the circular bend. The value for the loss of this Bezier curve is about 0.0273 dB, whereas for the circular bend the loss value is about 0.03034 dB. To compare the shape of the two bends, we plot the XY diagrams in Fig. 3.6d. Clearly, the optimum Bezier curve has a wider spread than the circular bend because $d_{\text{opt}} = 8.4044 \ \mu$m.

- **A variational approach for minimizing the loss of 180-degree bends**

Observing that a Bezier bend can lead to a lower propagation loss than a circular bend, we try to find a *symmetric* curve $y = y(x)$ such that it minimizes the propagation loss of a 180-degree bend. This curve is shown in Fig. 3.7a. The curve must pass through two given points $P_0 = (-R, d)$ and $P_1 = (R, d)$ where R is given but d is to be determined. By combining Eq. (3.13) and Eq. (3.14) and assuming that the constant part of the loss model in Eq. (3.11) is zero, we reach the following equation for the propagation loss (in dB units):

$$\text{Loss(dB)} = a \int_{-R}^{R} |y''|^b \left(1 + y'^2\right)^{(1-3b)/2} dx \qquad (3.18)$$

where a and b are parameters of the loss model. A change of variable $y' = f(x)$ and the assumption that $y'' \geq 0$ over the range of integration leads to

$$J[f] = a \int_{-R}^{R} (f')^b (1 + f^2)^{(1-3b)/2} \, dx. \qquad (3.19)$$

Here, $J[f]$ is the functional associated with the loss of the bend (in dB units). The objective
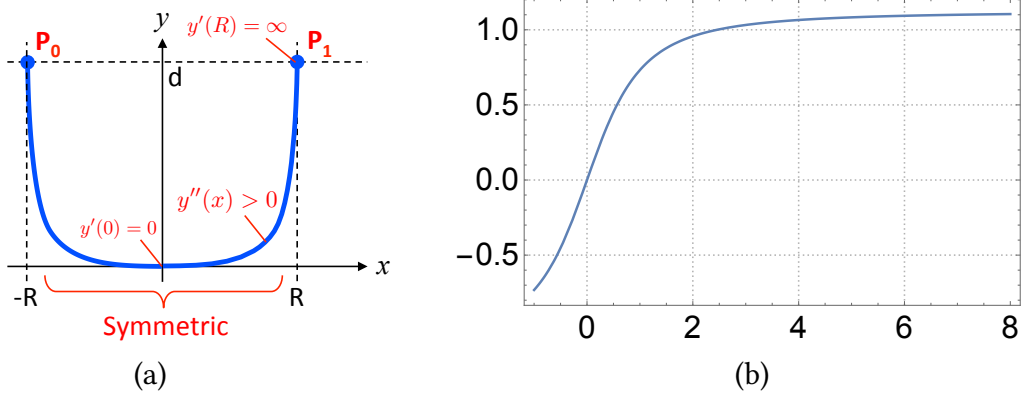
Figure 3.7: (a) Characteristics of the sought-after optimal curve. (b) Plot of the function $h(x;\xi) = x \times \text{HyperGeometric2F1}(0.5, \xi, 1.5, -x^2)$.

is find a function $f(x)$ such that it minimizes the loss function $J[f]$. The sought-after $f(x)$ function is a function at which the variation of the loss functional defined by

$$\delta J[f] = \frac{d}{d\epsilon} J[f + \epsilon \eta]_{\text{at } \epsilon=0} \tag{3.20}$$

is equal to zero $\delta J[f] = 0$. According to Euler-Lagrange equation, if a functional is defined by

$$J[f] = \int_{x_1}^{x_2} L(x, f, f') \, dx \tag{3.21}$$

the conditoin $\delta J[f] = 0$, leads to the following differential equation:

$$\frac{\partial L}{\partial f} - \frac{d}{dx}\left(\frac{\partial L}{\partial f'}\right) = 0. \tag{3.22}$$

Therefore, the function $f(x)$ that minimizes the loss functional of the bend must satisfy this equation when $L = (f')^b(1 + f^2)^{(1-3b)/2}$. The partial derivatives are calculated as:

$$\frac{\partial L}{\partial f} = (1 - 3b) f (1 + f^2)^{-(1+3b)/2} (f')^b$$
$$\frac{\partial L}{\partial f'} = b (f')^{b-1} (1 + f^2)^{(1-3b)/2} . \tag{3.23}$$

76

The end result is the following differential equation

$$f'' \left(1 + f^2\right) + \left(\frac{1 - 3b}{b}\right) f'^2 \, f = 0 \tag{3.24}$$

or equivalently

$$y''' \left(1 + y'^{\,2}\right) + \left(\frac{1 - 3b}{b}\right) y''^{\,2} \, y' = 0 \tag{3.25}$$

As expected, the parameter b in the loss model is the only parameter that defines the shape of the curve. Note that $f(x) = y'$. This equation is a nonlinear ODE that can be solved numerically.

Application of Beltrami identity for the variation of loss functional: This identity is a simplified version of Euler-Lagrange identity in Eq. (3.22) which applied in cases that the integrand of function in Eq. (3.21) does not explicitly depend on x, i.e. $\partial L / \partial x = 0$. Starting from Eq. (3.22) and multiplying both sides by f' and considering the chain rule

$$\frac{dL}{dx} = \frac{\partial L}{\partial f} f' + \frac{\partial L}{\partial f'} f'' + \frac{\partial L}{\partial x} \tag{3.26}$$

we end up with the following equation:

$$\frac{d}{dx} \left(L - f' \frac{\partial L}{\partial f'} \right) = \frac{\partial L}{\partial x} \tag{3.27}$$

which leads to the Beltrami identity

$$L - f' \frac{\partial L}{\partial f'} = C_0 \tag{3.28}$$

where C_0 is a constant. By substituting Eq. (3.23) into this equation and some simplifica-

tion we reach the following differential equation

$$\frac{df}{dx} = \left(\frac{C_0}{1-b}\right)^{1/b} (1+f^2)^{(3b-1)/2b}.$$ (3.29)

Considering that $b > 1$, the factor C_0 must be a negative factor. By defining $A = (C_0/(1-b))^{1-b}$ and refactoring, we have

$$\frac{df}{(1+f^2)^{(3b-1)/2b}} = A\,dx.$$ (3.30)

In order to determine the factor A, we take the integral of both sides of this equation. For the right side, the integral goes from 0 to some point x between 0 and R. Subsequently, the integral on the left side goes from 0 to some value $f(x)$ because $f(x = 0) = y'(0) = 0$. Therefore

$$\int_0^f \frac{df}{(1+f^2)^{(3b-1)/2b}} = \int_0^x A\,dx.$$ (3.31)

The left hand side has a closed-form expression in terms of the "HyperGeometric2F1" function:

$$h(f;\xi) = f \times \text{HyperGeometric2F1}(0.5, \xi, 1.5, -f^2) = A\,x$$ (3.32)

where $\xi = (3b-1)/2b$ and the "HyperGeometric2F1" function is defined as

$$\text{HyperGeometric2F1}(a, b; c; x) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 \frac{t^{b-1}(1-t)^{c-b-1}}{(1-t\,x)^a}dt$$ (3.33)

and $\Gamma(x)$ is the gamma function. A plot of the function $h(f;\xi)$ as a function of f is given in Fig. 3.7b. As can be seen, this function has a horizontal asymptote for $f \to \infty$ which occurs when $x \to R$ since $f(x) = y'(x)$. This fact allows us to find the coefficient A as:

$$A = \frac{1}{R}\left[\lim_{f\to\infty} h\left(f;\frac{3b-1}{2b}\right)\right].$$ (3.34)

For the loss model in Eq. (3.11) with $b \approx 2.9123$, the numerical value for the limit of the $h(f; \xi)$ is about 1.12454. Therefore

$$A \approx \frac{1.12454}{R}. \tag{3.35}$$

Quite interestingly, using the value of A and Eq. (3.29) in the loss functional $J[f]$ in Eq. (3.19), the total propagation loss (in dB) of the optimal 180-degree bend is simply given by

$$\min(J[f]) = a\,A^b\,(2R) \tag{3.36}$$

which can be easily compared to the loss of a 180-degree circular bend:

$$\frac{\text{Optimal Loss (dB)}}{\text{Circular Loss (dB)}} = \frac{2}{\pi}\left[\lim_{f \to \infty} h\left(f; \frac{3b-1}{2b}\right)\right]^b \approx 0.896\,. \tag{3.37}$$

This shows that the loss of the optimal 180-degree bend is 90% of the loss of a circular 180-degree bend that passes through P_0 and P_1 points in Fig. 3.7a. In order to find the optimal bend $y = y(x)$, we take the following steps:

- **Step1**: solve the following differential equation:

$$\frac{df}{dx} = A\,(1 + f^2)^{(3b-1)/2b} \tag{3.38}$$

  using the initial condition $f(0) = 0$ and the value of A from Eq. (3.34).

- **Step2**: calculate the following integral:

$$y(x) = \int_0^x f(t)\,dt \tag{3.39}$$

  to find the $y(x)$ in the $0 < x < R$ interval. Note that $y(-x) = y(x)$ due to the

(a)

(b)

(c)
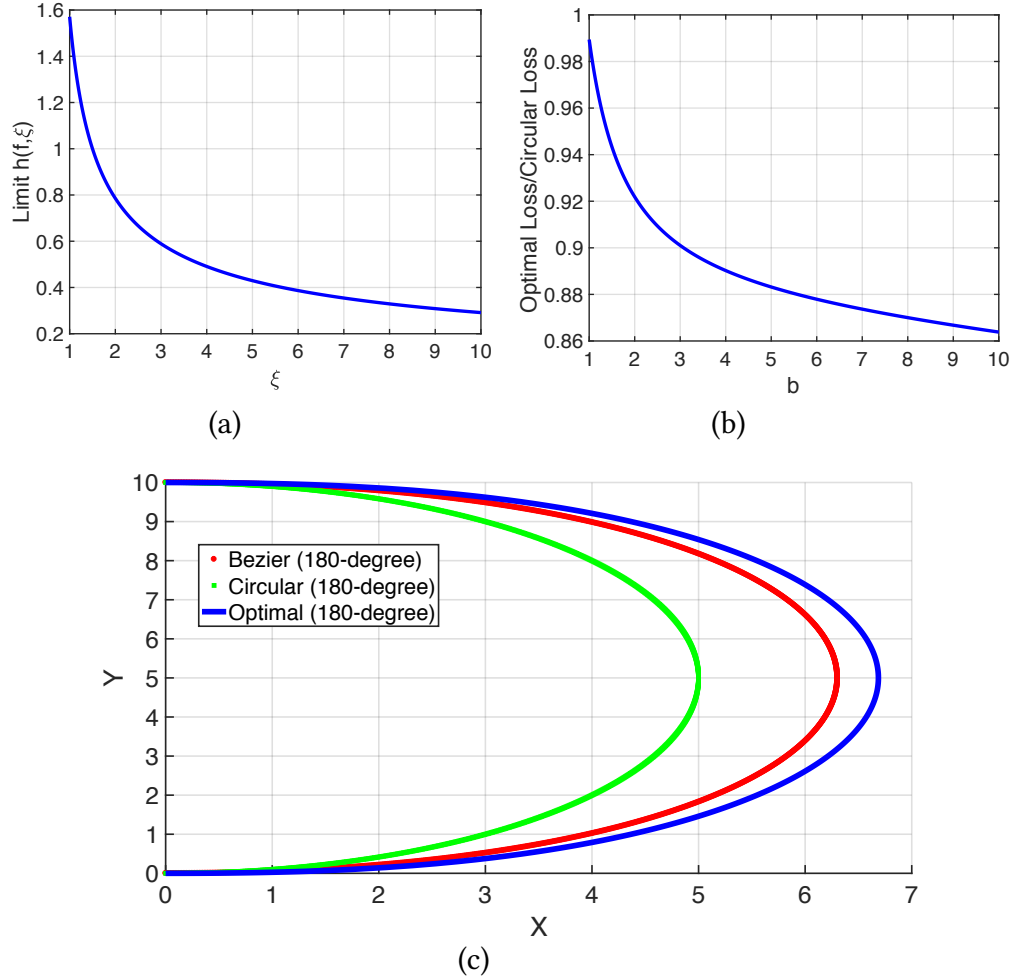
Figure 3.8: (a) Plot of $\lim_{f\to\infty} h\left(f;\xi\right)$ as a function of ξ. (b) Plot of the ratio of the loss of the optimal 180-degree bend to the circular bend as a function of b. (c) Comparison of the optimal 180-degree curve with the circular and Bezier curves.

symmetry of the curve. The parameter d is also calculated by

$$d = \int_0^R f(x)\,dx. \tag{3.40}$$

As an example, we calculate and plot the optimal 180-degree bend for $R = 5\mu$m and compare it against the circular bend and Bezier bend. Figure 3.8a plots the values of $\lim_{f\to\infty} h\left(f;\xi\right)$. Figure 3.8b plots the ratio of the optimal loss to the circular loss a function of loss model parameter b. It is seen that as the parameter b increases, the loss ratio in Eq. (3.37) becomes smaller. However, there is a lower bound for the ratio as $b \to \infty$.

To prove this, we start by noting that the function $p(\xi) = \lim_{f \to \infty} h(f; \xi)$ has an exact value of 1 at $\xi = 1.5$, i.e. $p(1.5) = 1$, therefore the limit in Eq. (3.37) is in the form of $(\to 1^+)^{(\to \infty)}$ which has an ambiguity that needs to be removed. In order to remove this ambiguity, we first perform a first-order expansion of $p(\xi)$ at $\xi = 1.5$:

$$p(\xi) = \lim_{f \to \infty} h(f; \xi) \approx p(1.5) + \frac{\partial P}{\partial \xi}(\xi - 1.5) \tag{3.41}$$

and by defining $x = 1/2b$ and noting that $\partial P / \partial \xi < 0$, we can see that

$$\lim_{b \to \infty} \left[p \left(\frac{3b-1}{2b} \right) \right]^b = \lim_{x \to 0^+} \left[1 + \left| \frac{\partial P}{\partial \xi} \right| x \right]^{1/2x} = \exp \left(\frac{1}{2} \left| \frac{\partial P}{\partial \xi} \right| \right) \tag{3.42}$$

and since

$$\left| \frac{\partial P}{\partial \xi} \right|_{\text{at } \xi = 1.5} = 0.613706 \tag{3.43}$$

the lower bound is given by

$$\boxed{\frac{\text{Optimal Loss (dB)}}{\text{Circular Loss (dB)}} = \frac{2}{\pi} \left[\lim_{f \to \infty} h \left(f; \frac{3b-1}{2b} \right) \right]^b \geq 0.8653} \tag{3.44}$$

Finally, fig. 3.8 compares the optimal 180-degree bend against the circular and Bezier bends. It is seen that Bezier bend provides a very close-to-optimal solution. In this example, the loss of the optimal 180-degree bend is about 0.02718 dB and the loss of the Bezier bend is about 0.0273 dB, both of which are smaller than the loss of the circular bend (0.03034 dB).

Directional coupler and power splitter

Optical power bounces back and forth between two silicon waveguides when put at close proximity to each other (typically at a distance less than 500 nm). This is due to the existence of the optical supermodes in the coupling region [161]. Assuming that the two

waveguides have identical cross-sections, the *transfer matrix* of the coupling is given by

$$
\begin{pmatrix} E_3^{out} \\ E_2^{out} \end{pmatrix} = \begin{bmatrix} t & -j\kappa \\ -j\kappa & t \end{bmatrix} \times \begin{pmatrix} E_4^{in} \\ E_1^{in} \end{pmatrix}
\tag{3.45}
$$

where t and κ are in general complex numbers that satisfy a lossless splitting of power, i.e. $|t|^2 + |\kappa|^2 \approx 1$. Figure 3.9 shows the structure of a directional coupler (DC) and the transfer matrix of coupling from the input ports 4 and 1 to the output ports 3 and 2. The scattering matrix of the directional coupler is then given by

$$
S_{DC} = \begin{bmatrix} r & t & -j\kappa & jx \\ t & r & jx & -j\kappa \\ -j\kappa & jx & r & t \\ jx & -j\kappa & t & r \end{bmatrix}.
\tag{3.46}
$$

Note that typically we assume that there is no reflection at each port (i.e. $r = S_{ii} = 0$) which is in general a good assumption. However, there can be some *back-coupling* from each port to its adjacent port (e.g. from port 1 to port 4) which can be of high importance in high quality resonators [94]. The parameter x in the scattering matrix accounts for such back-coupling. If the coupler structure is considered lossless, then $|t|^2 + |\kappa|^2 + |x|^2 + |r|^2 = 1$ and the scattering matrix should be a unitary matrix (each two rows or columns are orthogonal to each other). With this fact, we can find r and x in terms of t and κ as:

$$
\begin{aligned}
x &= t\sqrt{\frac{1 - (t^2 + \kappa^2)}{t^2 + \kappa^2}} \\
r &= \kappa\sqrt{\frac{1 - (t^2 + \kappa^2)}{t^2 + \kappa^2}}
\end{aligned}
\tag{3.47}
$$

Therefore, in an ideal directional coupler with $t^2 = 0.5$ and $\kappa^2 = 0.5$, we should have $x = 0$ and $r = 0$. If some amount of optical loss is associated with the coupler, then we
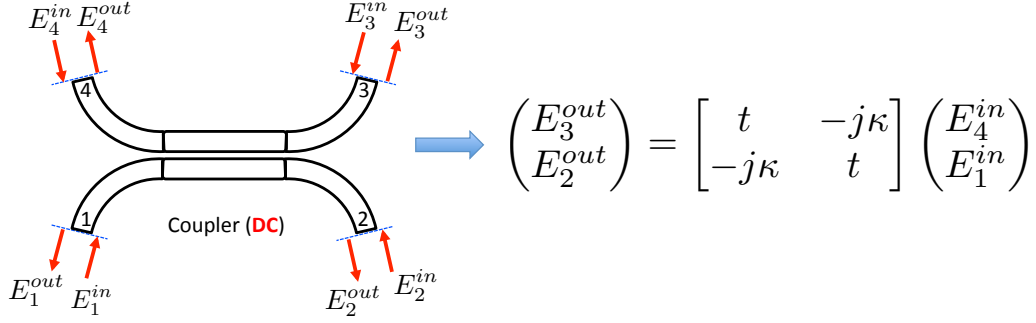
Figure 3.9: Structure of a directional coupler (DC) and input and output waves. The transfer matrix of coupling from ports 4 and 1 to ports 3 and 2 is shown.

can write

$$|t|^2 + |\kappa|^2 + |x|^2 + |r|^2 = \frac{P_{out}}{P_{in}} = \text{loss factor.} \tag{3.48}$$

For example, in a directional coupler with 0.5 dB loss we have $t^2 + \kappa^2 + x^2 + r^2 = 0.89$. By default, we assume that the directional couplers are ideal and there is no back-coupling ($|x| \approx 0$) or loss in the coupler. Finally, we point out that for a *3dB coupler* with no loss and no back-coupling the transfer matrix of coupling is

$$\begin{pmatrix} E_3^{out} \\ E_2^{out} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -j \\ -j & 1 \end{bmatrix} \times \begin{pmatrix} E_4^{in} \\ E_1^{in} \end{pmatrix} \tag{3.49}$$

which means there is a 90° phase shift between the two output waves. The dependence of t and κ on the physical parameters (length), optical modes and wavelength is established later in this thesis.

Waveguide crossing

When two silicon waveguides intersect with each other, a waveguide crossing is formed. Such crossing is considered as a discontinuity in the index profile along the path of the optical mode, hence outward radiation of optical power (loss) and leakage to the other paths (crosstalk) is inevitable. Low loss and low crosstalk crossings can be achieved by care-
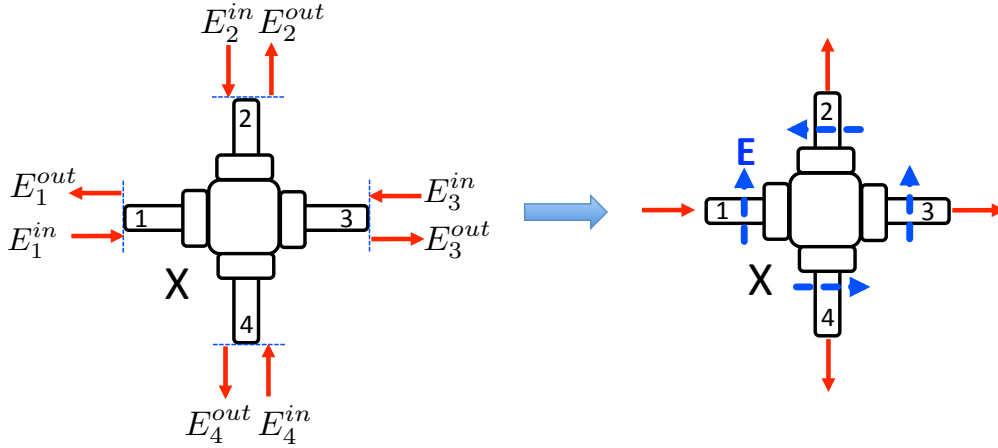
Figure 3.10: Schematic of a waveguide crossing. It is seen that the electric fields of crosstalk leakage have opposite direction for the TE_0 mode of the waveguides.

fully engineering the shape of the crossing region such that the optical mode experiences a smooth transition (Fig. 3.10). This is usually accomplished by tapering the waveguides near the crossing region [32, 110, 37, 111]. Ma *et al.* [111] have shown an optimization approach leading to 0.03 dB of optical loss and -37 dB of crosstalk. The scattering matrix of the crossing is then given by

$$
S_X = \begin{bmatrix} r & jx & t & -jx \\ jx & r & -jx & t \\ t & -jx & r & jx \\ -jx & t & jx & r \end{bmatrix} \tag{3.50}
$$

for the quasi-TE mode of the waveguides. The reason that the crosstalk terms in each row have opposite signs is shown in Fig. 3.10. In this matrix, t^2 and x^2 represent the fraction of optical power going through (e.g. from port 1 to port 3 or from port 2 to port 4 in Fig. 3.10), and optical crosstalk (e.g. from port 1 to port 2 or from port 1 to port 4 in Fig. 3.10). If we assume that the *radiation loss* is negligible and consider a lossless crossing,

the relation between the parameters are given by

$$r = 1 - t$$

$$x = \sqrt{t(1-t)}. \tag{3.51}$$

This equation shows that in a lossless crossing, the reflection at each port cannot be zero. In practice, however, a smooth expansion of the optical mode at each port allows for almost no reflection. In order to resolve this issue, we can introduce a *virtual port* as a fifth port in the compact model of the crossing. For simplicity, here we assume that the reflection at each port is negligible, $r \approx 0$, hence $t^2 + 2x^2 < 1$ due to the radiation loss. For a given through loss, the worst-case of crosstalk is found from $x^2_{\mathrm{max}} = 0.5(1 - t^2)$. For example, in a crossing with 0.03 dB of through loss, the worst-case of crosstalk is about -24.6 dB.

Y-Junction

Similar to the waveguide crossings, a Y-junction exhibits radiation loss due to the sharp discontinuity in the index profile [12, 200]. Zhang *et al.* have developed an optimization method based on particle swarm optimization for silicon-based Y-junctions that minimizes the radiation loss [212]. The radiation loss can be accounted for in the compact model of the junction using a virtual port as shown in Fig. 3.11 [153]. If we assume that there is no reflection in port 1 and the there is no radiation loss when port 1 is excited, the scattering matrix of the Y-junction is given by

$$S_Y = \begin{bmatrix} 0 & \sqrt{\frac{1+\delta}{2}} & \sqrt{\frac{1-\delta}{2}} & 0 \\ \sqrt{\frac{1+\delta}{2}} & 0 & 0 & \sqrt{\frac{1-\delta}{2}} \\ \sqrt{\frac{1-\delta}{2}} & 0 & 0 & -\sqrt{\frac{1+\delta}{2}} \\ 0 & \sqrt{\frac{1-\delta}{2}} & -\sqrt{\frac{1+\delta}{2}} & 0 \end{bmatrix} \tag{3.52}$$
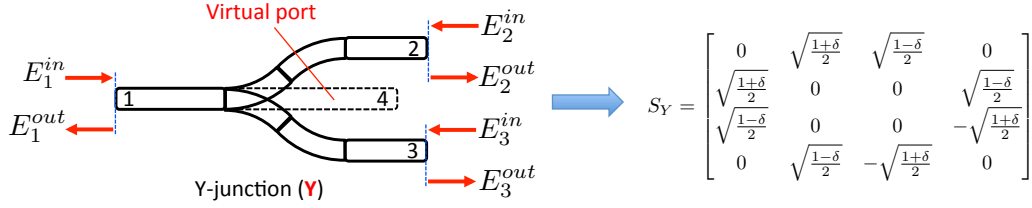
Figure 3.11: Schematic of a Y-junction structure with a virtual port to account for radiation loss.

which means there the power splitting ratio cannot be exactly the same from port 1 to port 2 and port 3. The imbalance in the power splitting ratio is given by the parameter δ:

$$\text{power splitting ratio} = \frac{1-\delta}{1+\delta}. \tag{3.53}$$

For an ideal Y-junction the power splitting ratio is equal to 1, hence $\delta = 0$.

3.2 Photonic circuit solver

A photonic circuit solver is required to find the transfer function of light in complex circuits. The circuit solver takes the scattering matrices of all the elements in the circuit as well as the connection of ports (i.e. which port is connected to which port) and finds the sought after transfer function. In this thesis, two general circuit solvers are considered:

Iteration method

For any photonic circuit made of linear photonic elements, it is possible to find the overall transfer function from one input port to one output port through iteration. This is due to the principle of superposition in linear circuits. To implement this method, a given port is excited in the first step of iteration cycle and then *is set to zero in the next steps* (this is very important). This allows the light to propagate through the structure. The total wave at each port in each iteration step is the sum of all the previous waves at that port.
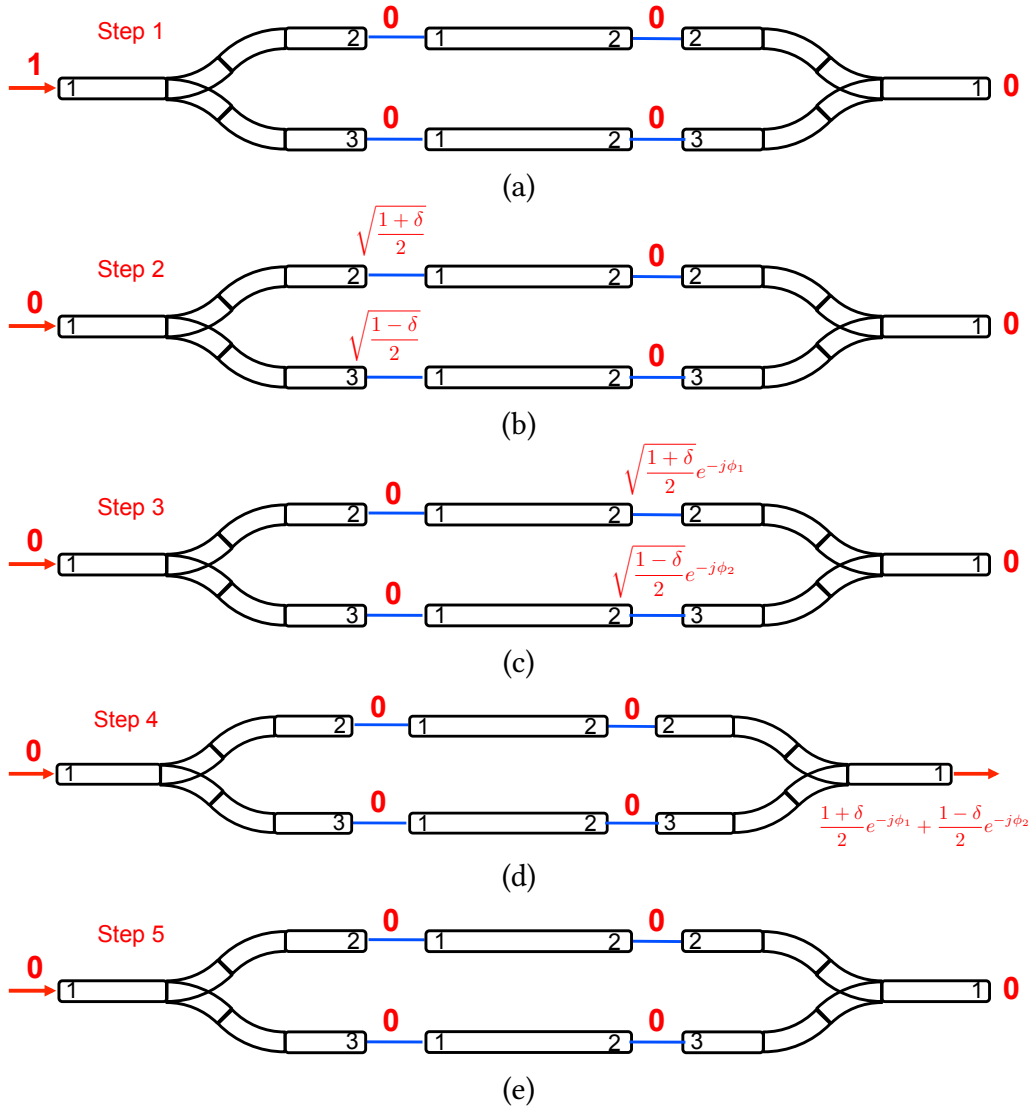
Figure 3.12: Illustration of finding the transfer function using the iteration method. (a)-(e) show step 1 through step 5 of the iterations with a complete convergence of the result.

Typically, the convergence is achieved within a few iteration cycles if a uni-directional propagation of light is considered.

To illustrate this approach, consider the photonic circuit shown in Fig. 3.12a made of two identical Y-junctions and two straight waveguides with different lengths. In the beginning, all the input and output waves in all the ports are initialized to zero. The goal is to find the transfer function from the left input to the right output port. In the first step, an excitation equal to 1 is put in port 1 of the left Y-junction (Fig. 3.12a). In the step 2,
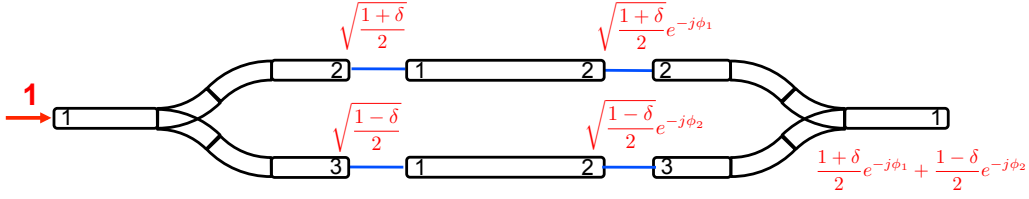
Figure 3.13: Calculated transfer functions for the propagation of light from left input to the right output based on the iteration method.

the excitation is propagated through the first Y-junction and the port 2 and port 3 of the Y-junction are updated to new values. The excitation is set to zero at the end of this step (Fig. 3.12b). In the step 3, the output ports of the two straight waveguides are updated to the new values and the zero input has now propagated to the port 2 and port 3 of the Y-junction (Fig. 3.12c). In the step 4, the output of the second Y-junction is updated with the new value and other ports are updated with zero (Fig. 3.12d). This value is saved and will be added to the values from next steps. In the step 5, the output is updated with a zero value (Fig. 3.12d). This means that the value of the output is not changing and the propagation is complete and a full convergence of the solution is reached.

The value of the wave at each port is basically the sum of the values of all the steps. In other words, the superposition of Fig. 3.12a-e will result in Fig. 3.13 which gives the transfer function fro the left input to any of the ports in the circuit. In this example, the input-output transfer function is given by

$$\frac{E_{out}}{E_{in}} = \frac{1+\delta}{2} \exp(-j\phi_1) + \frac{1-\delta}{2} \exp(-j\phi_2) \tag{3.54}$$

and the power transfer function is given by

$$\left| \frac{E_{out}}{E_{in}} \right|^2 = \cos^2\left(\frac{\Delta\phi}{2}\right) + \delta^2 \sin^2\left(\frac{\Delta\phi}{2}\right) \tag{3.55}$$

where $\Delta\phi = \phi_2 - \phi_1$. This is the general equation of a Mach-Zehnder interferometer under imbalanced power splitting [203].

Signal flow graph method

Signal flow graph is a mathematical method for calculating the transfer function from one node to another in a directional graph. In this approach, each photonic element is translated into a directional graph based on its scattering matrix. Then, the transfer function, T, of the graph is calculated based on Mason's gain formula:

$$T = \frac{O}{I} = \frac{\sum_{k=1}^{N} G_k \, \Delta_k}{\Delta} \tag{3.56}$$

where

$$\Delta = 1 - \sum L_i + \sum L_i L_j - \sum L_i L_j L_k + \sum L_i L_j L_k L_r - \cdots \tag{3.57}$$

and

- Δ = the determinant of the graph
- I = input-node variable
- O = output-node variable
- T = complete gain between I and O
- N = total number of forward paths between I and O
- G_k = path gain of the kth forward path between I and O
- L_i = loop gain of each closed loop in the system
- $L_i L_j$ = product of the loop gains of any two *non-touching* loops (no common nodes)
- $L_i L_j L_k$ = product of the loop gains of any three pairwise *non-touching* loops. Same definition applies for higher order terms as well.
- Δ_k = the cofactor value of Δ for the kth forward path, with the loops touching the kth forward path removed.

Basically, in this method first we find all the loops and their associated gain in the graph to calculate the determinant of the graph Δ. Then we find all the forward paths and their gain and their associated cofactors Δ_k and finally we find the transfer function.
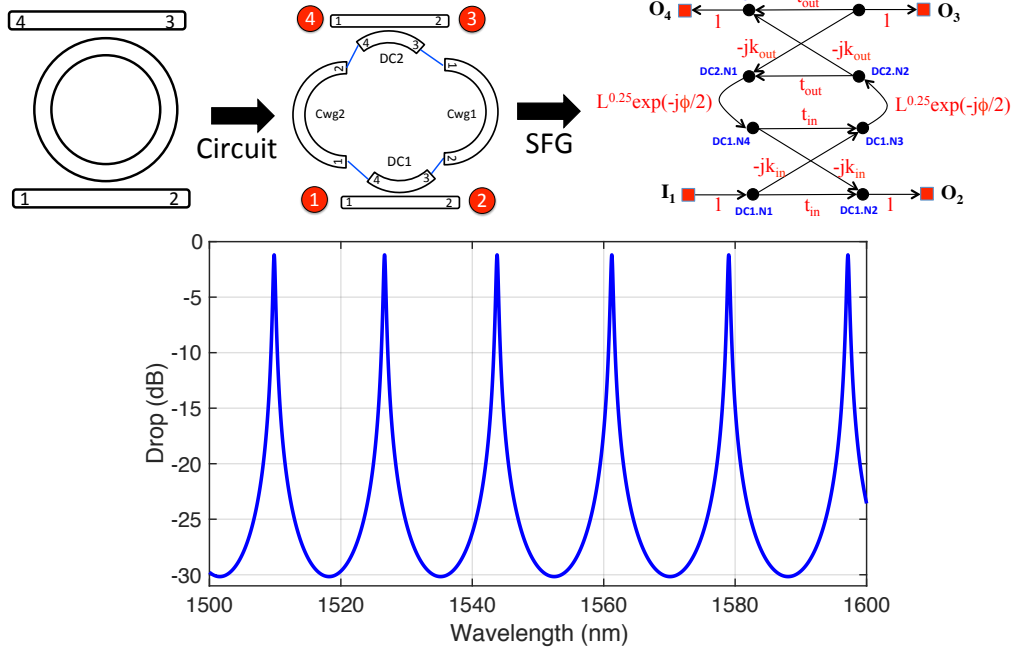
Figure 3.14: Decomposition of the add-drop ring structure into an interconnection of elementary building blocks and formation of its equivalent signal flow graph (SFG). The calculated spectrum of a 5μm ring with 10 dB/cm propagation loss made of 400nm$\times$220nm strip waveguides.

To illustrate the implemented SFG solver, we examine the transfer function of an add-drop structure as shown in Fig. 3.14. The add-drop filter can be constructed as a photonic circuit using the elementary building blocks (couplers and curved waveguides). The goal is to find the transfer function from port 1 to port 4, therefore a signal flow graph is constructed as shown in Fig. 3.14. There SFG solver first finds all the loops in the graph which is only one loop in this case constructed from nodes "DC1.N3", "DC2.N2", "DC2.N1", "DC1.N4" whose gain is

$$L_1 = t_{in}\, t_{out}\, \sqrt{L}\, \exp(-j\phi) \tag{3.58}$$

where L is the round trip power attenuation of the ring. Therefore, the determinant of the graph is calculated as $\Delta = 1 - L_1$. Then the SFG solver finds all the forward paths from port 1 to port 4 which is only one path in this case: "DC1.N1", "DC1.N3", "DC2.N2",

"DC2.N4" whose path gain is

$$G_1 = (-j\kappa_{in})(L^{0.25}\exp(-j\phi/2))(-j\kappa_{out}) \qquad (3.59)$$

and since this path is touching the only loop in the graph, its cofactor is calculated by removing all the terms that include L_1 from the determinant of the graph. Hence, $\Delta_1 = 1$. Finally, according to Mason's formula, the transfer function from port 1 to port 4 is given by

$$\frac{O_4}{I_1} = \frac{-\kappa_{in}\kappa_{out}\,L^{0.25}\exp(-j\phi/2)}{1 - t_{in}t_{out}\sqrt{L}\exp(-j\phi)} \qquad (3.60)$$

which is exactly the transfer function of the drop path of an add-drop ring. Using the developed mode solver, the power transfer function (in dB units) is calculated and plotted in Fig. 3.14 for a 5μm ring with 10 dB/cm of propagation loss made of 400nm$\times$220nm strip waveguides.

Note that in general the signal flow graph approach can provide good computation efficiency (the closest efficiency compared to having a closed-form equation) because it only needs to solve the graph once (to find the paths and determinant of the graph). Then for each wavelength, the main computation is performing multiplications and divisions in Eq. (3.60). Unfortunately, this is not the case for the iteration method in which the complete iteration procedure must be performed for each wavelength.

Modeling of backscattering in ring resonators

As an application of the SFG method, we examine the contradirectional coupling of two optical modes inside a ring resonator. This can occur due to the backscattering inside the resonator as shown in Fig. 3.15. Based on the analysis presented by Little *et al.* [104], we assume that the randomness of the scattering from the sidewalls of the ring resonator has no spatial preference. Therefore, the cumulative impact of the back-reflections inside the ring can be lumped at one location along the ring. We then built a signal flow graph
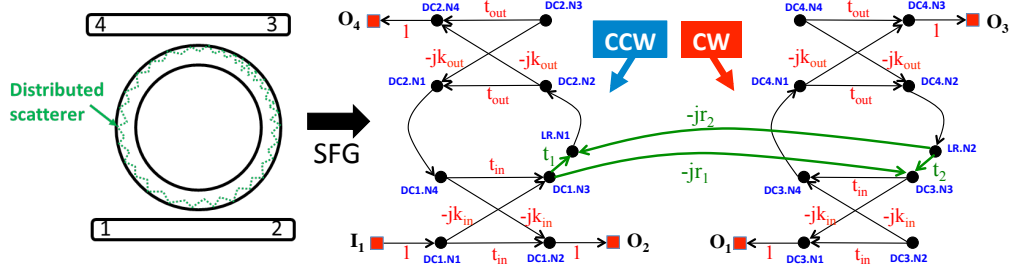
Figure 3.15: Signal flow graph model of contradirectional coupling inside a ring resonator.

that mutually couples the original counter clockwise (CCW) wave inside the ring to the back scattered clockwise (CW) wave. The signal flow graph essentially has two counter directional SFGs that are coupled to each other through the reflectance coefficients $,jr_1$ and $,jr_2$. If the scattering loss is given as α_s (considered in units of 1/m) and assumed to be the same for coupling in both directions (CW and CCW) inside the ring, then the following relations should hold:

$$|r_1| = |r_2| = r$$

$$|t_1| = |t_2| = t$$

$$t^2 + r^2 = \exp(-2\pi R \alpha_s) = L_s \approx 1 \,. \tag{3.61}$$

If we assume that the shift of the original resonance of the ring (in the absence of backscattering) is negligible, we can consider t_1 and t_2 to be real numbers. The splitting of resonance happens due to the fact that $r \neq 0$ and whether the right peak or the left peak is the stronger is determined by the cumulative phase shift, ϕ_s, due to the reflectance (i.e. $\angle r_1 + \angle r_2 = \phi_s$). The goal is to find the transfer functions O_4/I_1 (drop) using the SFG model.

- **Forward paths**

92

There are two forward paths in the graph from input I_1 to the drop O_4. By defining

$$p_0 = (-j\kappa_{in})(-j\kappa_{out})L_i^{0.25} \exp(-j\phi/2) \tag{3.62}$$

the two forward paths gain are given by

$$p_1 = t_1\, p_0$$

$$p_2 = (-jr_1)(-jr_2)\, p_0\, G = -r_1 r_2 G p_0 \tag{3.63}$$

where G is the loop gain of the ring: $G = t_{in}t_{out}\sqrt{L_i}\exp(-j\phi)$.

- **Loops and graph determinant**

There are three loops in the signal flow graph. Ring loop for the CCW wave (left part of SFG):

$$L_1 = t_1\, G \tag{3.64}$$

and the ring loop for the CW wave (right part of SFG):

$$L_2 = t_2\, G \tag{3.65}$$

and the loop between the left and right parts of SFG:

$$L_3 = (-jr_1)(-jr_2)\, G\, G = -r_1 r_2 G^2. \tag{3.66}$$

Graph determinant is given by

$$\Delta = 1 - (L_1 + L_2 + L_3) + L_1 L_2 = 1 - (t_1 + t_2)G + (r_1 r_2 + t_1 t_2)\, G^2 \tag{3.67}$$

since only L_1 and L_2 loops are not touching each other.
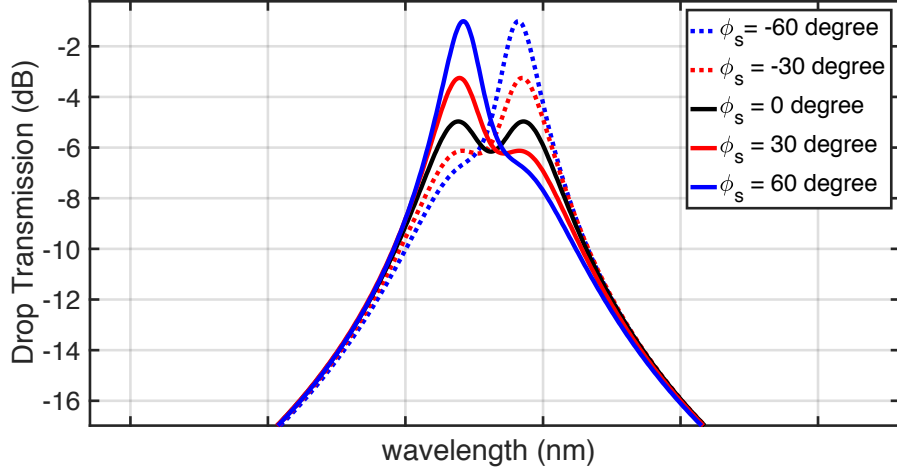
Figure 3.16: Example of calculating the drop spectrum of a ring resonator with backscattering for $r = 0.01$ and various accumulated scattering phase.

- **Cofactors of the forward paths**

The cofactors of the two forward paths are given by

$$\Delta_1 = 1 - L_2 = 1 - t_2\, G$$

$$\Delta_2 = 1 \tag{3.68}$$

- **Transfer function**

Finally, the transfer function is given by

$$\frac{O_4}{I_1} = \frac{p_1 \Delta_1 + p_2 \Delta_2}{\Delta} = p_0 \left(\frac{t_1 - (r_1 r_2 + t_1 t_2)G}{1 - (t_1 + t_2)G + (r_1 r_2 + t_1 t_2)G^2} \right) \tag{3.69}$$

It can be seen that if there is no backscattering in the ring, then $t_1 = t_2 = 1$ and $r_1 = r_2 = 0$ and the transfer function is reduced to

$$\frac{O_4}{I_1} = \frac{p_0}{1 - G} \tag{3.70}$$

which is exactly the result of Eq. (3.60). Examples of this transfer function are plotted in Fig. 3.16 for the backscattering strength of 0.01 (i.e. $r_1 = r_2 = 0.01$) and various accumulated scattering phase. The sign of the scattering phase determines which pick is higher and $\phi_s = 0$ corresponds to the case of the symmetric splitting.

3.3 Temporal Coupled Mode Theory (tCMT)

The equations for the coupling of microring resonators can be simplified for cases that the coupling of the rings to the waveguides is weak, i.e. $\kappa^2 \ll 1$, and the intrinsic loss of the ring is also very low, i.e. $L = \exp(-2\pi R\alpha) \approx 1$. With these assumptions the following first order approximations hold:

$$t = 1 - \delta t = \sqrt{1 - \kappa^2} \approx 1 - \frac{1}{2}\kappa^2 \Rightarrow \boxed{\delta t \approx \frac{1}{2}\kappa^2}$$
$$L = 1 - \delta L \approx 1 - 2\pi R\alpha \Rightarrow \boxed{\delta L \approx 2\pi R\alpha} \tag{3.71}$$

All-pass configuration

For an all-pass configuration, the transfer function from the input port to the output port can be easily written using the equivalent SFG model. The result is

$$T = \frac{t - \sqrt{L}\,e^{-j\phi}}{1 - t\sqrt{L}\,e^{-j\phi}} \tag{3.72}$$

where ϕ is the round-trip phase of the ring, which can be expanded to the first order at frequencies close to the resonance:

$$\phi(\omega) \approx 2\pi R\left(\beta(\omega_0) + \frac{d\beta}{d\omega}\delta\omega\right) = 2m\pi + \frac{2\pi R}{v_g}\delta\omega. \tag{3.73}$$

Here, $\delta\omega = \omega - \omega_0$ is the frequency deviation from the resonance, and v_g is the group velocity of the optical mode inside the waveguide. Clearly, the coefficient of $\delta\omega$ is the

group delay of one round-trip of the optical mode which is equal to the inverse of the free spectral range (in Hz units):

$$\tau_g = \frac{2\pi R}{v_g} = \frac{1}{\text{FSR}_{hz}} \, . \tag{3.74}$$

By substituting the linear expansion of the round-trip phase into the transfer function, we have

$$T \approx \frac{j\delta\omega + \frac{\text{FSR}}{\sqrt{L}}(t - \sqrt{L})}{j\delta\omega + \frac{\text{FSR}}{t\sqrt{L}}(1 - t\sqrt{L})} \tag{3.75}$$

which can be simplified using the first order approximations of t and L:

$$\frac{t - \sqrt{L}}{\sqrt{L}} \approx (1 - \delta t - 1 + \frac{\delta L}{2})(1 + \frac{\delta L}{2}) \approx \frac{\delta L}{2} - \delta t$$

$$\frac{1 - t\sqrt{L}}{t\sqrt{L}} \approx (1 - (1 - \delta t)(1 - \frac{\delta L}{2}))(1 + \delta t)(1 + \frac{\delta L}{2}) \approx \frac{\delta L}{2} + \delta t \, . \tag{3.76}$$

Finally, the transfer function is given by

$$T \approx \frac{j\,\delta\omega + \text{FSR}\left(\frac{\delta L}{2} - \delta t\right)}{j\,\delta\omega + \text{FSR}\left(\frac{\delta L}{2} + \delta t\right)} \tag{3.77}$$

which is usually expressed in terms of the intrinsic and coupling decay time constants:

$$T \approx \frac{j\delta\omega + \left(\frac{1}{\tau_i} - \frac{1}{\tau_c}\right)}{j\delta\omega + \left(\frac{1}{\tau_i} + \frac{1}{\tau_c}\right)} \tag{3.78}$$

where

$$\frac{1}{\tau_i} = \text{FSR}\frac{\delta L}{2} \Rightarrow \tau_i = \frac{\tau_g}{\pi R \alpha} = \frac{2n_g}{c\alpha}$$

$$\frac{1}{\tau_c} = \text{FSR}\,\delta t \Rightarrow \tau_c = \frac{2}{\kappa^2}\tau_g = \frac{2}{\kappa^2}\frac{2\pi R n_g}{c} \, . \tag{3.79}$$

The *photon lifetime* of the cavity, τ_p is defined as the time constant with which the optical power is decaying inside the cavity. Since $1/\tau = 1/\tau_i + 1/\tau_c$ describes the decay time

constant of the field, we have

$$\tau_p = \frac{\tau}{2} \Rightarrow \Delta f_{\text{3dB}} = \frac{1}{2\pi\tau_p} = \frac{\text{FSR}}{2\pi}(\delta L + 2\delta t) \tag{3.80}$$

which provides an approximation for the *finesse* of the resonator:

$$\text{Finesse} = \frac{\text{FSR}}{\Delta f_{\text{3dB}}} \approx \frac{2\pi}{\kappa^2 + 2\pi R \alpha}. \tag{3.81}$$

Add-drop configuration

Similar approximations can be made for the transfer function of the add-drop configuration. Starting from the SFG model, we have

$$T = \frac{-\kappa_1\kappa_2 L^{0.25} e^{-j\phi/2}}{1 - t_1 t_2 \sqrt{L} e^{-j\phi}} \approx \frac{-\kappa_1\kappa_2}{j\tau_g\,\delta\omega + \delta t_1 + \delta t_2 + \delta L/2} \tag{3.82}$$

which is usually written as

$$T = \frac{-\mu_1\mu_2}{j\delta\omega + \left(\frac{1}{\tau_i} + \frac{1}{\tau_{c,1}} + \frac{1}{\tau_{c,2}}\right)} \tag{3.83}$$

and

$$\frac{1}{\tau_i} = \frac{\delta L/2}{\tau_g}, \frac{1}{\tau_{c,1}} = \frac{\delta t_1}{\tau_g}, \frac{1}{\tau_{c,2}} = \frac{\delta t_2}{\tau_g}, \mu_1 = \frac{\kappa_1}{\sqrt{\tau_g}}, \mu_2 = \frac{\kappa_2}{\sqrt{\tau_g}}. \tag{3.84}$$

Finally, we note that

$$\mu_{1,2} = \sqrt{\frac{\kappa_{1,2}^2}{\tau_g}} = \sqrt{\frac{2}{\tau_{c_{1,2}}}}. \tag{3.85}$$

These equations can also be derived independently from the conservation of energy in the ring-waveguide system.

3.4 Summary of the Chapter

In this chapter, compact models of elementary building blocks of silicon photonic circuits were presented based on the scattering matrix. We presented a variational optimization approach for bent waveguides and proved that optimal bends can provide lower than their circular counterparts. It was also shown that a lower bound exists for the ratio of the optimal bend loss (in dB) to that of a circular bend and the exact value of this lower bound was calculated to be 0.865256. Finally, two methods for analyzing complex photonic circuits. The first method was based on iteration due to the superposition principle in linear systems. This method was shown to be very efficient when a uni-directional propagation of light is considered. The second method was based on the theory of directional graphs which proved to give the closed-form solution with minimal computational burden. As an example, the signal flow graph method was used to model the contradirectinal coupling of light inside a ring resonator.

Chapter 4

Optimal Routing Tables for Silicon Photonic Switch Fabrics

Abstract– *In this chapter, the analysis of scalability of silicon photonic switch fabrics based on 2×2 Mach-Zehnder switch (MZS) elements is presented. In particular, Benes topology is studied. It is shown that such switch topologies usually offer multiple switch configurations for a particular input-output port mapping. The idea is to choose the best possible switch configuration for each routing mapping based on the physical layer performance (loss, crosstalk) of photonic elements.*

Demand for high communication bandwidth at low cost, e.g. in ultra-large scale computing systems [160], has led to the deployment of photonic links side-by-side of electronics. On one hand, the distance-independence and wavelength-division-multiplexing features of photonics together deliver high bandwidth for transmission; on the other hand, electronic, potentially high-radix packet switches provide for switching needs. The two realms of photonics and electronics are bridged via electrical-to-optical (E/O) and O/E direct conversions, necessary at every hop. Such conversions can be avoided if electronic switches are replaced by reconfigurable electro-optical switches. One option to realize such optical switches, of arbitrary radix, is to cascade and interconnect (potentially numerous) elementary 2×2 switches [59, 108, 149]. Silicon photonics is an excellent platform for fabrication of such architectures due to its compatibility with CMOS technology and capability to scale to large component counts. 2×2 switch elements (SE) based on Mach-Zehnder interferometers (MZI) [59, 60] or microring resonators (MRR) [129, 27, 41] have been proposed and extensively studied in this platform. Mach-Zehnder switches (MZS),

although requiring a large footprint to reduce the required switching voltage, have the significant advantage of permitting broadband operation, contrary to MRR.

The cascading of elementary switches makes large scale designs extremely sensitive to characteristics of single 2×2 elements. In this chapter, we take advantage of our physical layer modeling approach to analyze the performance of an electro-optic 2×2 MZS. Based on this initial result, we characterize the optical losses and crosstalk performances (worst-case) of 4×4 and 8×8 Benes switching topologies. We show that waveguide crossings play an important role in the propagation of crosstalk through the switch fabric, especially for the largest port counts. The worst-case power penalty of each topology is summarized into a graph representation of the switch as the result of this study [19].

Furthermore, we use our modeling platform to automatically generate generic path-mapping schemes for silicon integrated switch fabrics [42]. The bounds of path-mapping variations are for the first time defined for Benes topology and the particular routing with *equalized path-dependent power penalties* is selected for each permutation. Quantitative analysis on an 8×8 electro-optic Benes switch is presented by a representative case of equal permutation. Results indicate that this advanced path-mapping approach improves the worst-path power penalty by 6 dB and achieves a root mean square error (RMSE) of less than 0.3 dB for all 8 paths routing through the switch [40].

4.1 Characteristics of Benes Topology

The $N \times N$ ($N = 2^m$ and $m \geq 1$) Benes topology corresponds to the Clos architecture with two banks of 2×2 elements forming the input and output stages, while the inner layer is implemented using two smaller Clos networks recursively, as shown by Fig. 4.1. An $N \times N$ Benes switch requires $N/2$ ($= 2^{m-1}$) rows of 2×2 elements and $2log_2N - 1$ stages ($= 2m-1$), summing up to a total of $N \log_2 N - N/2$ number ($= 2^{m-1}(2m-1)$) of 2×2 elements. Therefore the total number of ways that a Benes switch can be configured
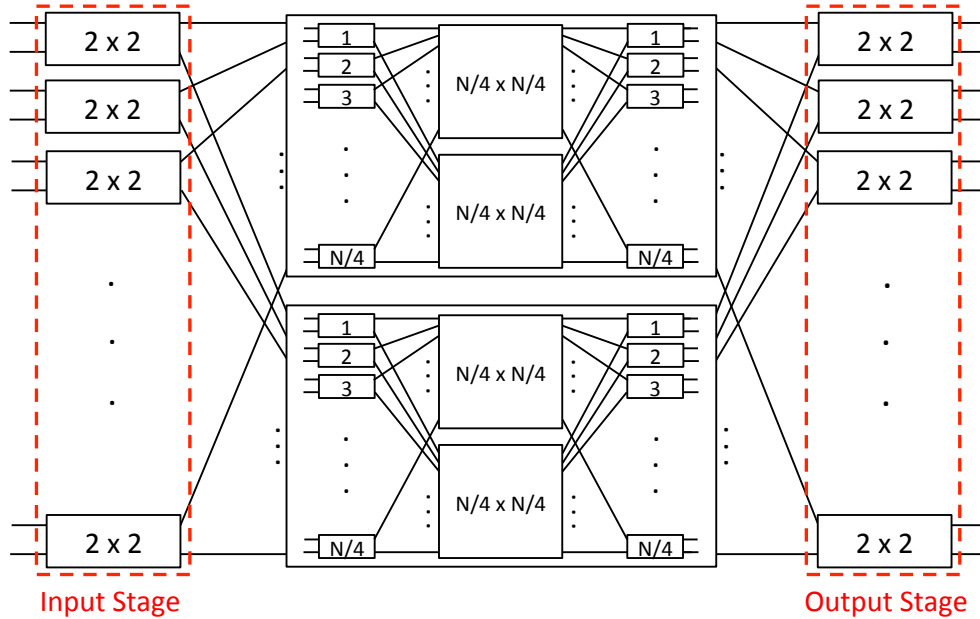
Figure 4.1: Schematic of an $N \times N$ Benes topology. This shows how the $N \times N$ switch is recursively constructed using two $N/2 \times N/2$ switches.

is $2^{(N \log_2 N - N/2)}$ since each 2×2 element can be configured in either "Bar" or "Cross" state. For example, an $8x \times 8$ switch requires 20 switch elements which leads to 2^{20} (=1048576) ways that the switch can be configured. On the other hand, an $N \times N$ switch fabric can map the N input ports to the N output ports in only $N!$ (= $N(N-1)(N-2) \cdots 1$) ways which is typically much lower than the number of switch configurations. This means that multiple switch configurations can provide the same mapping of input ports to the output ports.

For any given input-to-output mapping permutation, the upper bound of path-mapping variations is reached when the outputs of each layer share the same 2×2 switching element as their corresponding inputs. This is because the state of 2×2 switching elements in the outer left column (input stage in Fig. 4.1), with a total number of $N/2$ (=2^{m-1}) switching elements, can all be set independently while the configuration of the output stage is predetermined by the chosen mapping and the configuration of the previous stages of the switch. Considering the recursive construction of the Benes topology and assuming the number of configurations for $N \times N$ Benes is C_m, a recursive equation

thus exists as

$$C_m \left(\text{upper} \right) = 2^{2^{m-1}} \times C_{m-1}^2 \tag{4.1}$$

for the upper bound of the configuration repetitions of a given input-output permutation. In order to solve this recursive equation, we define $C_m = 2^{a_m}$ and take the $log_2()$ of both sides of the equation:

$$a_m(\text{upper}) = 2\, a_{m-1} + 2^{m-1} \tag{4.2}$$

with the initial conditions $a_1 = 0$ and $a_2 = 2$. The initial condition $a_1 = 0$ is due to the fact that a 2×2 Benes switch is basically a simple switch element which only has one configuration for a given input-output mapping (either in Bar or Cross state). Therefore, $C_1 = 1$ and $log_2 C_1 = 0$. With the same reasoning, it is easy to see that the maximum number of switch configurations for a given mapping in a 4×4 Benes is 4, hence $C_2(\text{upper}) = 4$ and $a_2(\text{upper}) = 2$. The solution of Eq. (4.2) for a_m has the format $a_m = A_m \times 2^m$ where $A_1 = 0$ and $A_2 = 1$, hence:

$$A_m - A_{m-1} = \frac{1}{2} \Rightarrow \sum_{k=2}^{m} (A_k - A_{k-1}) = \sum_{k=2}^{m} \frac{1}{2} \Rightarrow A_m = \frac{m-1}{2} \tag{4.3}$$

and

$$\boxed{a_m(\text{upper}) = (m-1)\, 2^{m-1} \quad , \quad C_m(\text{upper}) = 2^{(m-1)2^{m-1}}} \tag{4.4}$$

This means that in a 4×4 Benes switch ($m = 2$), an input-output port mapping can correspond to a maximum number of 4 different switch configurations, and in an 8×8 Benes switch ($m = 3$), this number is increased to 256 different configurations.

It is also possible to find a lower bound on the number of switch configurations for a given input-output mapping. Referring back to Fig. 4.1, we see that there are cases in which the input stage and the output stage must have a particular configuration or its complement (i.e. all the Bar switches go to Cross and all the Cross switches go to Bar). This happens if all the input pairs from different switch elements in the input stage are
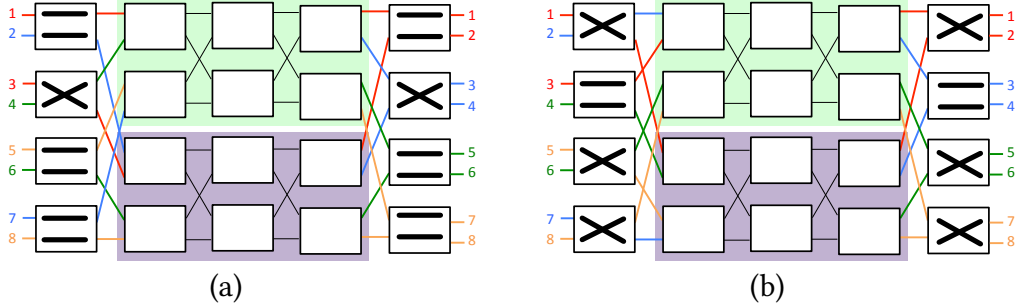
Figure 4.2: Two possible configuration of the input stage and output stage of an 8×8 Benes topology for the input-output mapping (1, 2, 3, 4, 5, 6, 7, 8)→(1, 3, 2, 5, 7, 6, 4, 8).

routed to the same switch element in the output stage. For example, consider the input-output mapping (1, 2, 3, 4, 5, 6, 7, 8)→(1, 3, 2, 5, 7, 6, 4, 8) as shown in Fig. 4.2. Since the inputs of each switch element in the input stage are routed to different switch elements in the output stage, there are two possible configurations (given in Fig. 4.2a and 4.2b) for the input and output stages. Therefore, the recursive equation for the minimal configuration states is given by

$$C_m(\text{lower}) = 2 \times C_{m-1}^2 \tag{4.5}$$

with the initial conditions $C_1 = 1$ and $C_2 = 2$. The change of variable $C_m = 2^{a_m}$ in this case leads to

$$a_m(\text{lower}) = 2\,a_{m-1} + 1\,. \tag{4.6}$$

The general format of the solution to this recursive equation is $a_m = A\,2^m + B$. It is easy to see that $B = -1$ to satisfy the equation and $A = 1/2$ to satisfy the initial condition $a_1 = 0$. Therefore

$$\boxed{a_m(\text{lower}) = 2^{m-1} - 1 \quad , \quad C_m(\text{lower}) = 2^{(2^{m-1}-1)}} \tag{4.7}$$

This means that in a 4×4 Benes switch ($m = 2$), an input-output port mapping can correspond to a minimum number of 2 different switch configurations, and in an 8×8 Benes switch ($m = 3$), this number is increased to 8 different configurations.
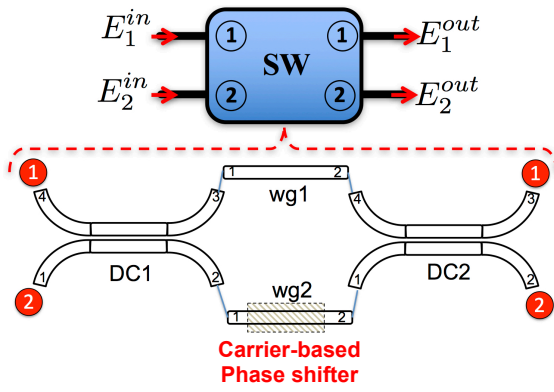
Parameter	Value
Waveguide	Strip 450x220
	Ridge 450x220x50
Wavelength	1550 nm
dn_{eff}/dn_{si}	1.1
Wg loss	2 dB/cm
Phase shifter	PIN (injection)
Coupler loss	0.1 dB
Ideal 50/50	YES

Figure 4.3: A photonic circuit representing the 2×2 Mach-Zehnder switch element.

4.2 Modeling of 2×2 MZS Elements

The structure of a 2×2 Mach-Zehnder switch element is shown in Fig. 4.3. We assume that the input and output power splitters are made of directional couplers and the length of the two arms is the same. For simplicity, we assume that only one arm is equipped with a high-speed carrier-based electro-optic (EO) phase shifter and the switch element is used in the single-arm drive mode. We also assume that the device is made of 450×220 nm strip waveguides and the phase shifter is made of 450×220nm ridge waveguide with a 50 nm slab. It is further assumed that the propagation loss of the waveguides is about 2 dB/cm and the directional couplers are ideal 3dB splitters at $\lambda = 1550$ nm with 0.1 dB of additional loss.

Transfer function

The transfer matrix of the switch from input ports 1 and 2 to the output ports 1 and 2 is defined as

$$\begin{pmatrix} E_1^{out} \\ E_2^{out} \end{pmatrix} = \begin{bmatrix} T_{11} & T_{21} \\ T_{21} & T_{22} \end{bmatrix} \begin{pmatrix} E_1^{in} \\ E_2^{in} \end{pmatrix} \tag{4.8}$$

which can be calculated by cascading the transfer matrices of the three sections:

$$
\begin{bmatrix} T_{11} & T_{21} \\ T_{21} & T_{22} \end{bmatrix} = \begin{bmatrix} t_2 & -j\kappa_2 \\ -j\kappa_2 & t_2 \end{bmatrix} \begin{bmatrix} \sqrt{L_1}e^{-j\phi_1} & 0 \\ 0 & \sqrt{L_2}e^{-j\phi_2} \end{bmatrix} \begin{bmatrix} t_1 & -j\kappa_1 \\ -j\kappa_1 & t_1 \end{bmatrix} \tag{4.9}
$$

where ϕ_1 and ϕ_2 are the phase shifts of each arm and L_1 and L_2 are the power attenuation factor of each arm (wg1 and wg2). The result of the multiplication is given by

$$
T_{11} = t_1 t_2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 - \delta_1 \delta_2 \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right)
$$

$$
T_{12} = -j\kappa_1 t_2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 + \frac{\delta_2}{\delta_1} \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right)
$$

$$
T_{21} = -j\kappa_2 t_1 \sqrt{L_1}\, e^{-j\phi_1} \left(1 + \frac{\delta_1}{\delta_2} \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right)
$$

$$
T_{22} = -\kappa_1 \kappa_2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 - \frac{1}{\delta_1 \delta_2} \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right) \tag{4.10}
$$

where $\delta_1 = \kappa_1/t_1$ and $\delta_2 = \kappa_2/t_2$ denote the imbalance in the power dividers (couplers), $L_\Delta = L_2/L_1$ is the imbalance in the loss of the two arms, and $\Delta\phi = \phi_2 - \phi_1$ is the imbalance in the propagation phase of the two arms. Note that in writing these equations, "wg1" is taken as the *reference path*, hence the terms outside the parenthesis are the forward propagation through "wg1" from the input port to the output port. Assuming that the two power splitters (DC1 and DC2) in Fig. 4.3 are *identical*, we have $t_1 = t_2 = t$ and $\kappa_1 = \kappa_2 = \kappa$ and $\delta_1 = \delta_2 = \delta$. Therefore:

$$
T_{11} = t^2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 - \delta^2 \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right)
$$

$$
T_{12} = -j\kappa t \sqrt{L_1}\, e^{-j\phi_1} \left(1 + \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right)
$$

$$
T_{21} = -j\kappa t \sqrt{L_1}\, e^{-j\phi_1} \left(1 + \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right)
$$

$$
T_{22} = -\kappa^2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 - \frac{1}{\delta^2} \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right). \tag{4.11}
$$

Finally, we also assume that the power splitters have a perfect 50/50 splitting ratio, hence $\kappa = t$ and $\delta = 1$ and

$$
\begin{aligned}
T_{11} &= \kappa^2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 - \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right) \\
T_{12} &= -j\kappa^2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 + \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right) \\
T_{21} &= -j\kappa^2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 + \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right) \\
T_{22} &= -\kappa^2 \sqrt{L_1}\, e^{-j\phi_1} \left(1 - \sqrt{L_\Delta}\, e^{-j\Delta\phi} \right).
\end{aligned}
\tag{4.12}
$$

With these assumptions, it can be seen that $T_{11} = -T_{22}$ (thru) and $T_{12} = T_{21}$ (cross). The thru and cross power transfer functions are given by

$$
\begin{aligned}
\text{Thru:} \quad |T_{11}|^2 &= \kappa^4 L_1 \left[\left(1 - \sqrt{L_\Delta} \right)^2 + 4\sqrt{L_\Delta} \sin^2(\Delta\phi/2) \right] \\
\text{Cross:} \quad |T_{21}|^2 &= \kappa^4 L_1 \left[\left(1 - \sqrt{L_\Delta} \right)^2 + 4\sqrt{L_\Delta} \cos^2(\Delta\phi/2) \right]
\end{aligned}
\tag{4.13}
$$

Clearly, the thru transmission has a minimum and the cross transmission has a maximum at $\Delta\phi = 0$ ("Cross" state of the switch). Similarly, the thru transmission has a maximum and the cross transmission has a minimum at $\Delta\phi = \pi$ ("Bar" state of the switch). Therefore, the *optimum switching performance* is achieved when the phase shifter in wg2 adds an extra π phase shift.

Electro-Optic phase shifter

A symmetrical carrier-injection PIN diode is considered in the lower arm of the 2×2 MZS element. Based on the applied voltage and change of carrier density ($\Delta N = \Delta P$ in units of cm^{-3}), the index of refraction and loss (in units of cm^{-1}) of the silicon core change through the empirical equations of plasma dispersion effect [174, 155] in silicon at 1550

nm wavelength:

$$\Delta n_{si} = - \left[8.8 \times 10^{-22} \Delta N + 8.5 \times 10^{-18} (\Delta P)^{0.8}\right]$$

$$\Delta \alpha_{si} = 8.5 \times 10^{-18} \Delta N + 6 \times 10^{-18} \Delta P \tag{4.14}$$

where $\Delta N = \Delta P$ is assumed. The effective index of the optical mode and the effective loss factor of the optical mode are then calculated based on the sensitivity factor of the optical mode to the core index:

$$\Delta n_{\text{eff}} \approx \frac{\partial n_{\text{eff}}}{\partial n^{si}} \Delta n_{si} = \Gamma_{\text{core}} \left(\frac{n_g}{n_g^{si}}\right) \Delta n_{si} \approx 1.1 \, \Delta n_{si}$$

$$\Delta \alpha_{\text{eff}} \approx \frac{\partial n_{\text{eff}}}{\partial n^{si}} \Delta \alpha_{si} = \Gamma_{\text{core}} \left(\frac{n_g}{n_g^{si}}\right) \Delta \alpha_{si} \approx 1.1 \, \Delta \alpha_{si} \tag{4.15}$$

where we assumed the approximate value of $\partial n_{\text{neff}} / \partial n_{si} \approx 1.1$ for a 450×220 nm strip or ridge waveguide. From these equations we see that the relation between Δn_{si} and $\Delta \alpha_{si}$ is given by

$$\Delta n_{si} = -10^{-4} \left[0.6069 \, \Delta \alpha_{si} + 2.5138 \, (\Delta \alpha_{si})^{0.8}\right] . \tag{4.16}$$

and the phase shift generated by the phase shifter is given by

$$\Delta \phi = \frac{2\pi}{\lambda} \Delta n_{\text{eff}} \, l \tag{4.17}$$

where l is the length of the arm. Since the phase shifter is driven to provide 0 or π phase shift, the transfer functions of the switch are given by

- **CROSS state:** $\Delta \phi = 0$ and $L_\Delta = 1$. Therefore the insertion loss and crosstalk values are given by

$$\text{XTalk:} \qquad |T_{11}|^2 = 0$$

$$\text{IL:} \qquad |T_{21}|^2 = 4 \, L_1 \kappa^4 \tag{4.18}$$

which shows there is no crosstalk leakage.

- **BAR state**: $\Delta\phi = \pi$ and $L_\Delta = \exp(-\Delta\alpha_{\text{eff}}l)$. Therefore the insertion loss and crosstalk values are given by

$$
\begin{aligned}
\text{IL:} \qquad & |T_{11}|^2 = \kappa^4 \, L_1 \left(1 + \sqrt{L_\Delta}\right)^2 \\
\text{XTalk:} \qquad & |T_{21}|^2 = \kappa^4 \, L_1 \left(1 - \sqrt{L_\Delta}\right)^2
\end{aligned}
\qquad (4.19)
$$

which shows that there is some crosstalk leakage.

The switching extinction defined as the ratio between $|T_{11}|^2$ and $|T_{21}|^2$ in this case only depends on $\sqrt{L_\Delta}$. For any given length of the Mach-Zehnder's arms, the loss and crosstalk levels (in dB) of the Bar state are determined by the $\Delta\alpha_{si}$ that corresponds to a π phase shift. We set the length of the arms to 300 μm to keep it short enough and then build the photonic circuit model of the MZS as shown in Fig. 4.3. The transfer functions are found using the iteration-based circuit solver which only requires 4 iteration steps for complete convergence (because of unidirectional propagation of light). The length and gap of the 3dB power splitters are set to 12.3 μm and 200 nm and an additional loss of 0.1 dB is associated with each splitter, resulting in $\kappa^2 \approx 0.4891$ and $t^2 \approx 0.48776$. Figure 4.4 plots the thru and cross transmission of the designed MZS element as a function of the excess loss of the phase shifter. The switch is in the cross state for $\Delta\alpha = 0$ and is in the bar state for $\Delta\alpha = 10.81 \text{ cm}^{-1}$.

The main point is the significant difference in the insertion loss and crosstalk level of the switch in Bar and Cross state, which can greatly impact the performance of the switch. Since, the MZS shows better performance in the cross state, it is better to have more 2×2 MZS elements in the cross state when configuring a large scale switch fabric. This is the idea behind optimal routing tables for silicon photonic switch fabrics.
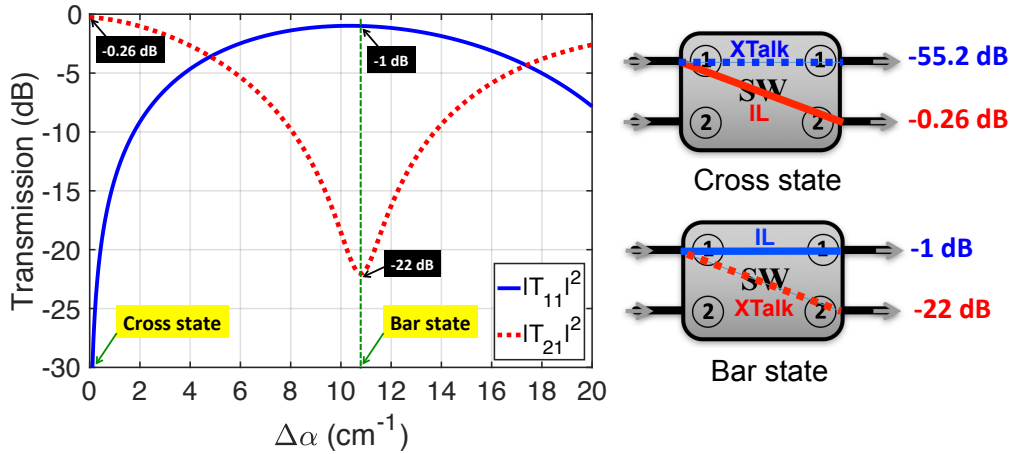
Figure 4.4: Characterization of loss and crosstalk for Bar and Cross states of 2x2 MZS element. The length of each arm of Mach-Zehnder is set to 300 μm.

Wavelength dependence

Since the performance of 3dB power splitters based on directional couplers depends on the wavelength of operation, the imbalance of splitting happens when wavelength is changed from 1550 nm. As shown in Fig. 4.5a, a directional coupler based on 450×220 nm strip waveguides has a 50/50 splitting at 1550 nm for $l = 12.3\mu$m when the coupling gap is set to 200 nm. However, Fig. 4.5b shows that the splitting of power strongly diverges from 50/50 as the wavelength is changed. Therefore, it is very important to see how the insertion loss and the crosstalk of the 2×2 MZS element change as a function of wavelength. To do so, the wavelength is swept and the photonic circuit solver is run for each wavelength. The result of the calculations for the cross state are shown in Fig. 4.5c and for the bar state are shown in Fig. 4.5d.

The interesting observation is that the performance of the switch in the Bar state is almost independent of the wavelength and the insertion loss in the cross state is also nearly independent of the wavelength. However, XTalk leakage in the cross state heavily depends on the wavelength. Typically, -20 dB is defined as the reference for the XTalk. Therefore, the MZS element presented here has a -20 dB crosstalk bandwidth of about 29.63 nm.
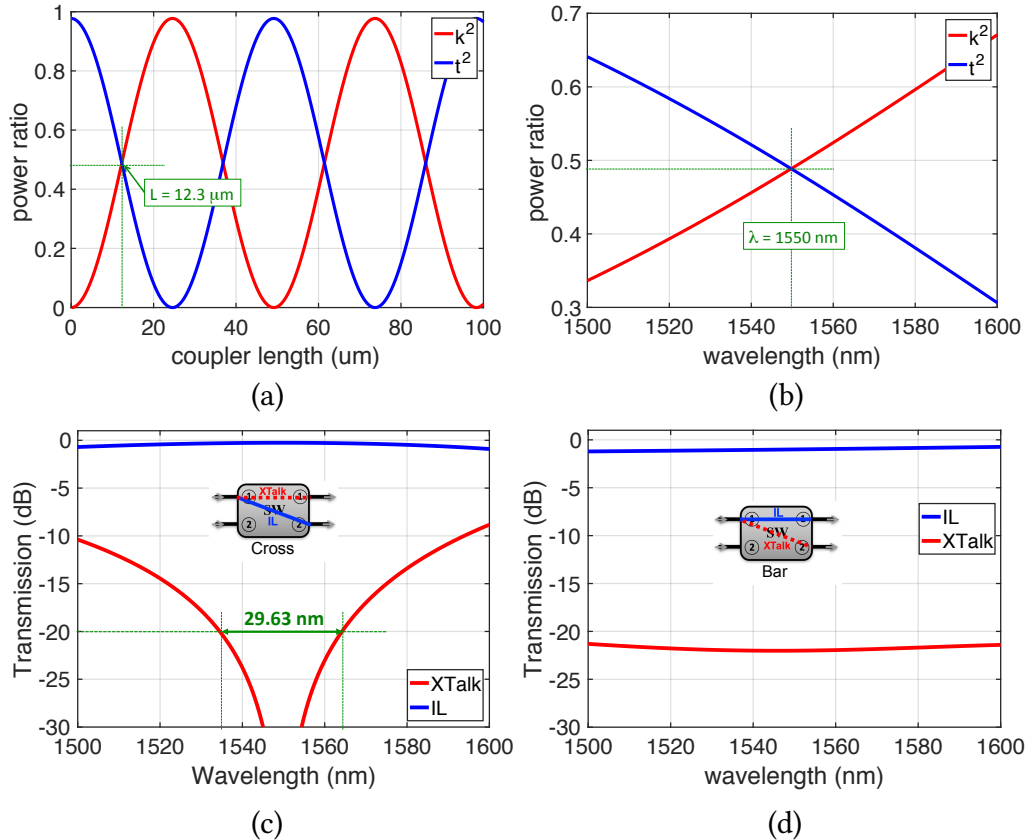
Figure 4.5: Wavelength dependence of the MZS element. (a) Power splitting ratio of the 3dB coupler as a function of coupler length at λ = 1550 nm. (b) Dependence of the power splitting coefficients of the 3dB coupler on wavelength. (c) Dependence of the IL and XTalk of the switch on wavelength in the cross state. (d) Dependence of the IL and XTalk of the switch on wavelength in the bar state.

4.3   4×4 Benes Switch Fabric

A 4×4 Benes topology is constructed using six 2×2 MZS elements as shown in Fig. 4.6. There are three stages each consisting of two MZS elements. A passive waveguide shuffle is between successive stages. We assume that the switch stages are about 100 μm apart. For each waveguide crossing, we assume a thru loss of 0.03 dB and a crosstalk of 37 dB [111].

In order to analyze the port-to-port performance of this switch fabric, we first construct the photonic circuit of each passive shuffle and then connect then to the photonic circuits of all the 2×2 MZS elements. Finally, since the unidirectional propagation of light
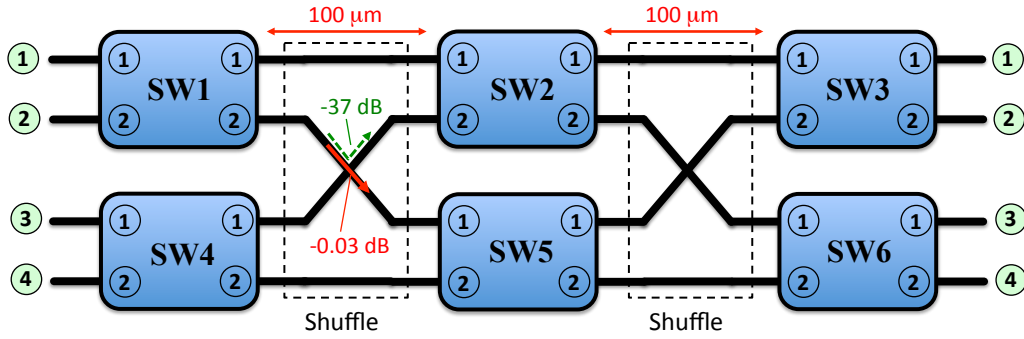
Figure 4.6: Structure of a $4{\times}4$ Benes switch.

is considered, the *iteration-based* photonic circuit solver is used.

Worst-case Performance Graph

We perform an exhaustive enumeration of all the 64 different switch configurations and find the input-output value of the transmission. 0 dBm of optical power is put in each input port and all the 64 configurations are swept and for each case we find all the output powers. Then we sort the values of optical power at each output port and plot them with their corresponding switch configurations. For example, Fig. 4.7 shows the sorted output powers at output ports 1 and 2 when input port 1 is excited with 0 dBm (note that the x-axis is not the same in the two plots). The high values of these plots are IL, corresponding to the switch configurations that allow the input port 1 go to that output port. A big jump is observed indicating the amount of optical crosstalk at each output port. Note that our presentation of switch configuration on the x-axis of these two plots is in the format of $X_1X_2|X_3X_4|X_5X_6$ which represents the state of the switches in each stage or column (i.e. X_1 corresponds to SW1, X_2 corresponds to SW4, X_3 corresponds to SW2, and so on).

The edge of the first big jump in each plot of Fig. 4.7 indicates the switch configurations that lead to the worst-case of insertion loss and crosstalk. By exciting all the four input ports (one at a time) and identifying the worst cases, the $4{\times}4$ Benes switch can be represented by a worst-case performance graph as shown in Fig. 4.8a. Such worst-case
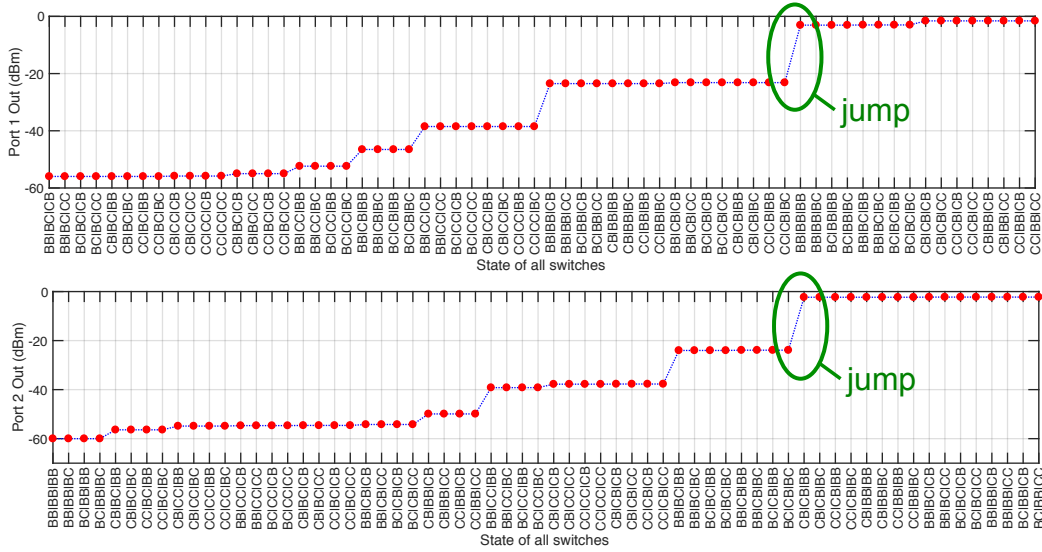
Figure 4.7: Example of exhaustive enumeration of all the 64 switch configurations for 4×4 Benes topology when input port 1 is excited with 0 dBm of optical power.

representation is particularly useful when trying to design an optical link with a photonic switch in it. For example, by inspecting the worst-case graph we can conclude that the maximum loss for routing input port 1 to the output port 1 is about 3.06 dB and the maximum possible crosstalk leakage is the sum of the leakages from input ports 2, 3 and 4. Therefore, the worst crosstalk at output port 1 is about -18.4 dB. Since the performance of each 2×2 MZS depends on the wavelength, the worst-case graph changes with the wavelength as well. As an example, Fig. 4.8b shows the spectra of input 1 to output 1 transmission for all the 64 possible switch configurations. It is seen that crosstalk leakage can significantly increase for wavelengths away from 1550 nm.

If switch topologies other than Benes are considered, the main change occurs in the structure of the passive shuffle stages, and the parameters of the 2×2 MZS elements and the waveguide crossings generally remain unchanged. This means that the light propagation inside the switch fabric based on iteration method should be evaluated with the corresponding models for the passive shuffle stages.
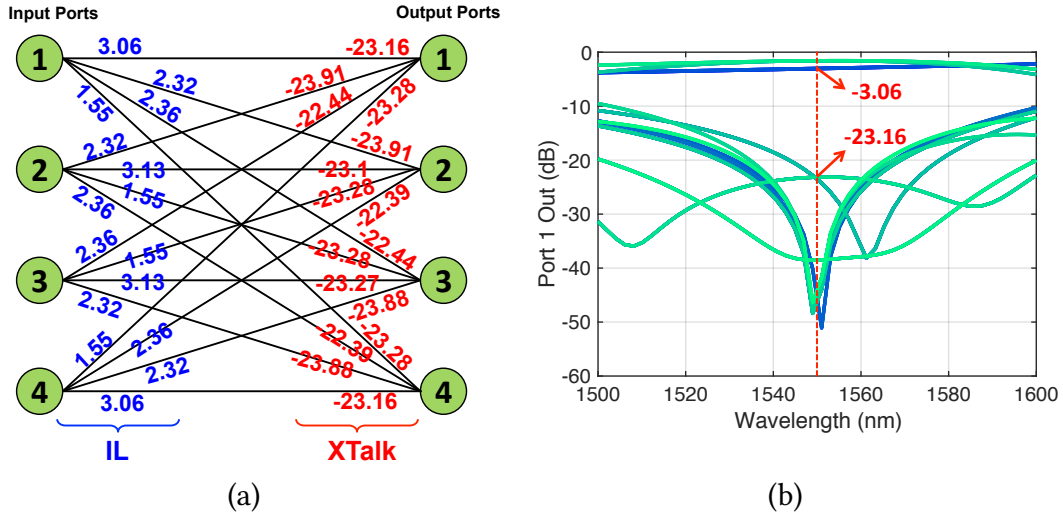
Figure 4.8: (a) Construction of worst-case performance graph for a 4×4 Benes topology. (b) Transmission spectra of input port 1 to output port 1 for all the 64 switch configurations. Worst-case of loss and crosstalk are also marked.

Optimal Routing Table

Due to the repetition of switch configurations for a given input-output port mapping of the switch, we can choose a switch configuration that has the best performance. The first step is find the input-output mapping that corresponds to a given switch configuration. To do so, as shown in Fig. 4.9 we define the boolean transition matrix of each switch stage (column) as

$$C_{4\times4} = \begin{bmatrix} SWi_{2\times2} & 0_{2\times2} \\ 0_{2\times2} & SWj_{2\times2} \end{bmatrix} \tag{4.20}$$

where $SWi_{2\times2}$ and $SWj_{2\times2}$ are the transition matrices of the two 2×2 switch elements in the switch stage. The transition matrix of a 2×2 MZS can be written as

$$\text{BAR:} \quad SW_{2\times2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad , \quad \text{CROSS:} \quad SW_{2\times2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{4.21}$$

for both bar and cross states. The transition matrix of the passive shuffle for the 4×4
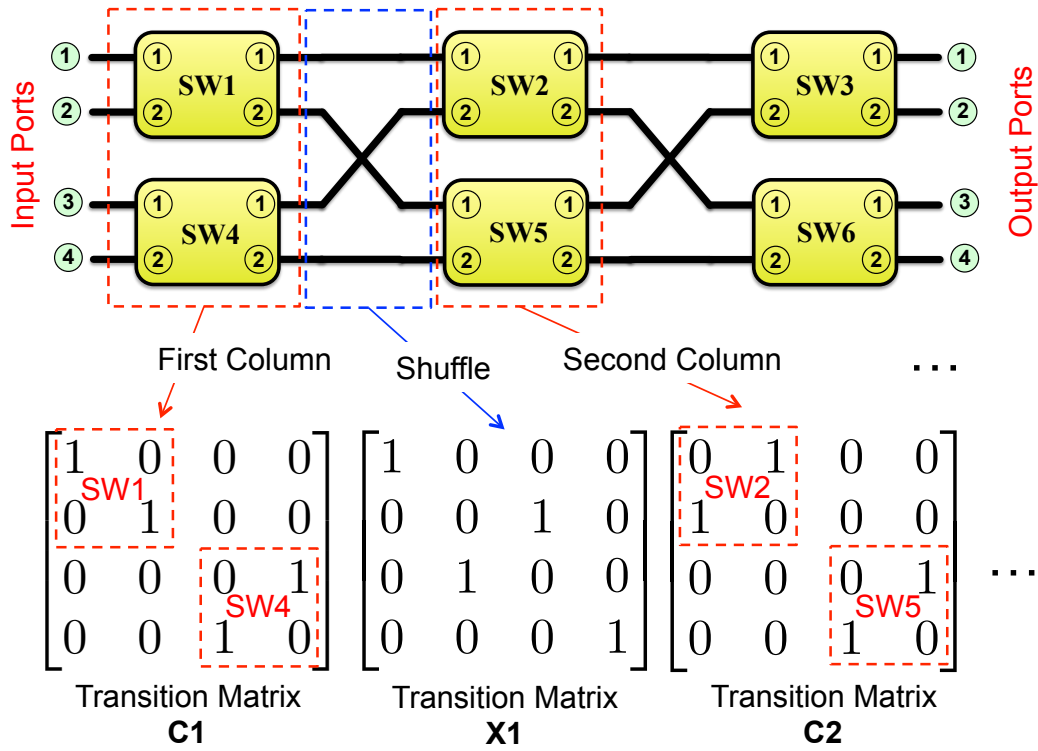
Figure 4.9: Definition of transition matrices for each switch stage and passive shuffle.

Benes is given by

$$\mathbf{X}_{4\times4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.22}$$

and finally, the input-output transition matrix of the switch is given by

$$\mathbf{T}_{4\times4} = \mathbf{C3} \times \mathbf{X2} \times \mathbf{C2} \times \mathbf{X1} \times \mathbf{C1}\,. \tag{4.23}$$

In order to find out which output port corresponds for example to the input port 1, we
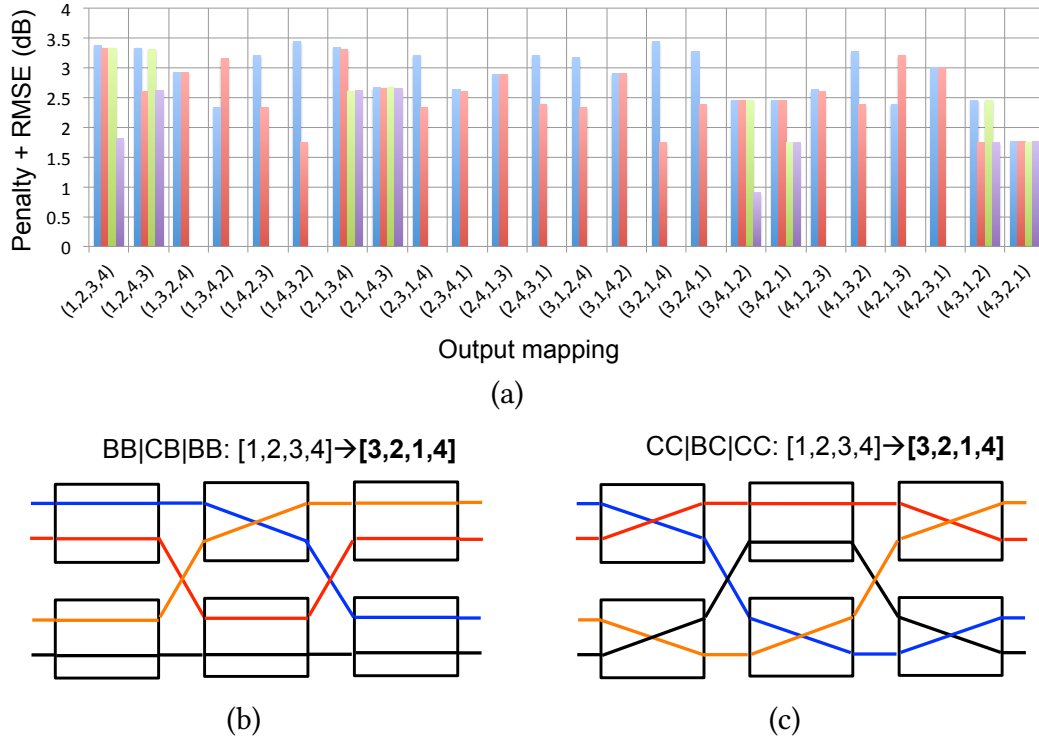
Figure 4.10: (a) Average power penalty plus the RMSE value for all the 24 input-output mappings and 64 switch configurations. It is seen that for each mapping, there is a particular configuration that provides the lowest value which is taken as the optimal routing. (b) Worst-case of routing for (1,2,3,4)→(3,2,1,4) mapping. (c) Best-case of routing for (1,2,3,4)→(3,2,1,4) mapping.

write

$$
\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \mathbf{T} \times \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\tag{4.24}
$$

where only one of the y_i's is nonzero and the rest are zero, indicating the corresponding output port.

In order to find the optimal routing strategy for each input-output mapping (1,2,3,4)→(x_1, x_2, x_3, x_4), we first perform an exhaustive enumeration of all the 64 switch configurations and for each switch configuration we find its corresponding mapping. For each configuration, we find the individual input-output path power penalties by combin-
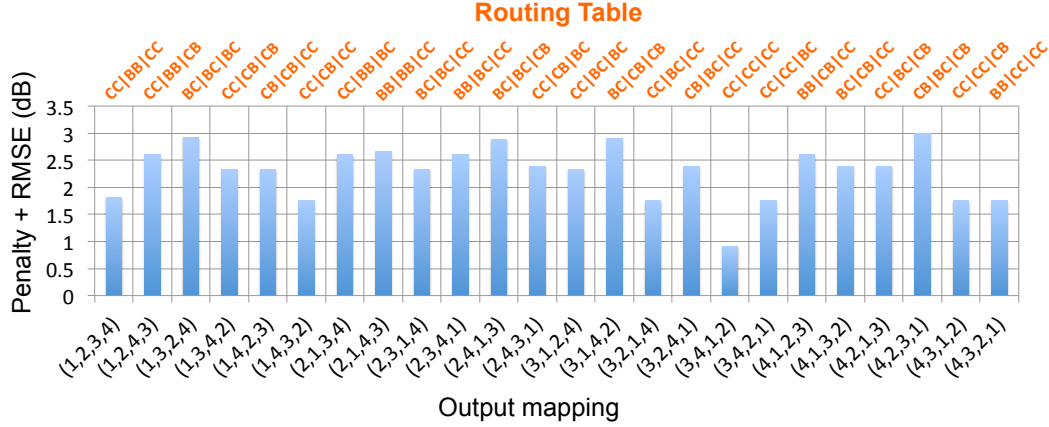
Figure 4.11: Optimal routing table for 4×4 electro-optic Benes topology.

ing the loss of that path and the crosstalk leakage from other paths. We then define the average power penalty of the switch configuration and and its root-mean-square-error (RMSE) as

$$PP_{\mathrm{avg}} = \frac{1}{N} \sum_{k=1}^{N} PP_k$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (PP_i - PP_{\mathrm{avg}})^2} \tag{4.25}$$

for $N = 4$. We then define a figure of merit (FOM) as the sum of average power penalty of the switch configuration and its RMSE. Since it is desirable for all the paths to have more or less the same power penalty, for each input-output mapping we choose the switch configuration that corresponds to the minimum value for FOM. Figure 4.10a shows the calculated FOM for all the 24 possible mappings and 64 possible switch configurations. It is seen that some mappings have 4 repetitions while other have 2 repetitions. For each mapping, there is one switch configuration that provides the minimum FOM. For example, consider the mapping (1,2,3,4)→(3,2,1,4). In this mapping the input 1 is routed to output 3, input 2 is routed to output 2, input 3 is routed to output 1, and input 4 is routed to output 4. There are two possible switch configurations that allow this mapping as shown in Fig.

(a)

(b)

(c)
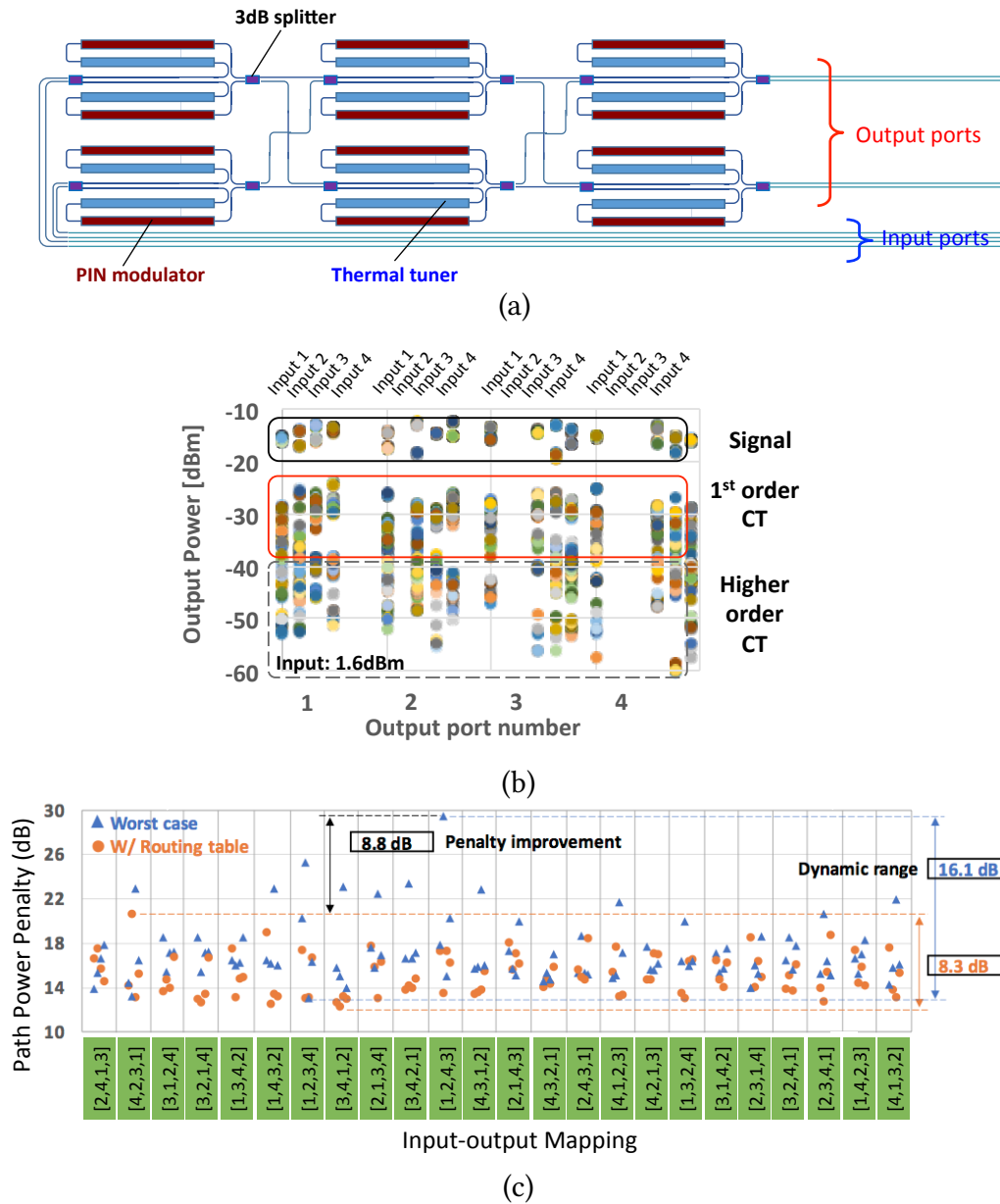
Figure 4.12: Experimental demonstration of optimal routing table for 4×4 electro-optic Benes topology. (a) Schematic drawing of the switch layout. (b) Measured output power of all the 16 input-output combinations for all 64 switch configurations. (c) Comparison of the port-to-port penalty for optimal routing and worst-case routing.

4.10b and Fig. 4.10c. The difference between FOM for these two cases is more than 1.5 dB. This shows that an optimal choice of the switch configuration can save a significant amount of optical power penalty. Figure 4.11 shows the optimal routing table for all the 24 input-output permutations.

Experimental Demonstration of Routing Table

In order to experimentally demonstrate the impact of the physical layer performance of the switch fabric on its optimal routing choice, we used a 4×4 Benes switch. A schematic drawing of the actual layout of this switch is shown in Fig. 4.12a. Each Mach-Zehnder switch element in this silicon photonic chip is equipped with a thermo-optic phase tuner and electro-optic phase shifter in both arms. We chose use each MZS in the single-arm operation.

An input laser source was put at each input port and the optical power was measured at each output port for all the four input ports and 64 switch configurations. Figure 4.12b shows all the 4×4×64 (= 1024) collected data points, indicating both path loss and crosstalk leakage. The total accumulated loss of the input and output coupling of light to/from the silicon chip is estimated to be about 11 dB. Based on these measurements, the power penalty of each input-output path was calculated and then plotted for the worst-case and the best case of the path mapping according to Fig. 4.10a. As expected, the optimal routing table provides the best uniformity of the four path penalties in each mapping. A difference of 8.8 dB is observed in our measurements for the worst path penalty with and without the routing table. This simply means that just by using optimal routing table the dynamic range requirements of optical power at the output ports of the switch is reduced by 8.8 dB.

4.4   8×8 Benes Switch Fabric

Similar to the 4×4 Benes, the goal is to investigate the impact of the physical layer performance on the path mappings. An 8×8 Benes topology and the transition matrix of each passive shuffling stage is shown in Fig. 4.13. This Benes topology has 20 MZS elements which results in 2^{20} (= 1048576) switch configurations. A given input-output mapping such as (1,2,3,4,5,6,7,8)→(1,2,3,4,5,6,7,8) in this topology has 256 switch configurations

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\qquad
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\qquad
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

Transition Matrix   Transition Matrix   Transition Matrix
X1   **X2, X3**   **X4 = (X1)^t**
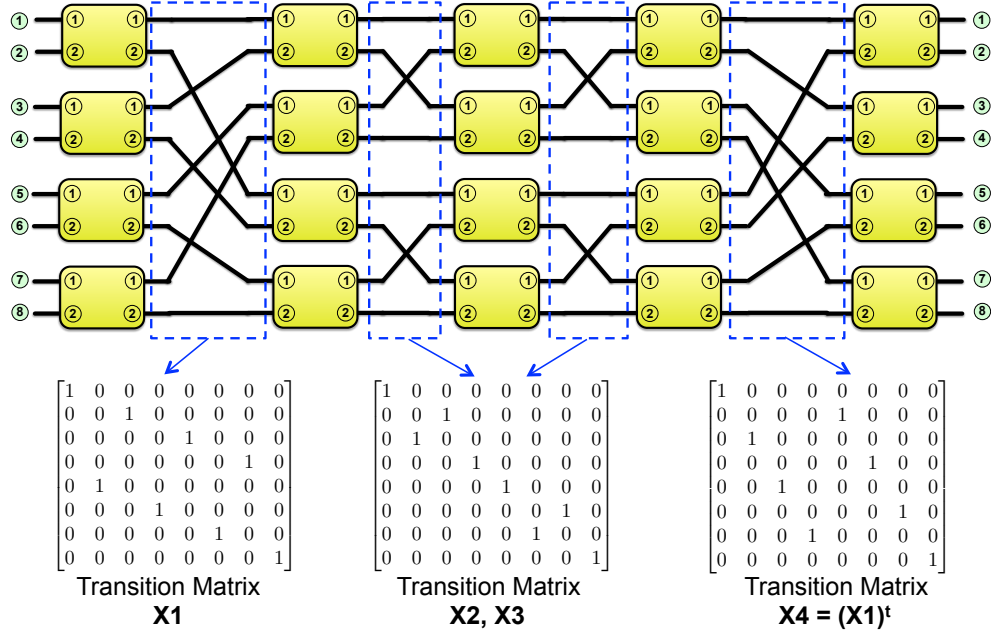
Figure 4.13: Schematic of an 8×8 Benes topology. Transition matrix of each passive shuffling is also shown.

associated with it. Therefore, it is important to determine which one of these 256 configurations provide the best performance.

As shown in Fig. 4.14a, the RMSE of the path power penalties for each configuration corresponding to the (1,2,3,4,5,6,7,8)→(1,2,3,4,5,6,7,8) mapping is calculated [40]. The minimum RMSE is chosen as the optimal routing for this mapping since it *equalizes* the power penalty of all the input-output paths. The minimum RMSE of this mapping corresponds to the CCCC|CCCC|BBBB|CCCC|BBBB configuration as shown in Fig. 4.14b. On the other hand, the maximum RMSE for this mapping corresponds to the CCCB|CCCB|BBBB|CCCB|CCCB configuration as shown in Fig. 4.14c. Clearly, the minimum RMSE provides a good uniformity for all the eight input-output paths whereas the maximum RMSE leads to one path having a much higher penalty that the others. This example again signifies the fact that we can take advantage of the element of repetition in large scale silicon photonic switch fabrics to build up an optimal routing table based on the physical layer performance of the switch.

(a)

(b)   (c)
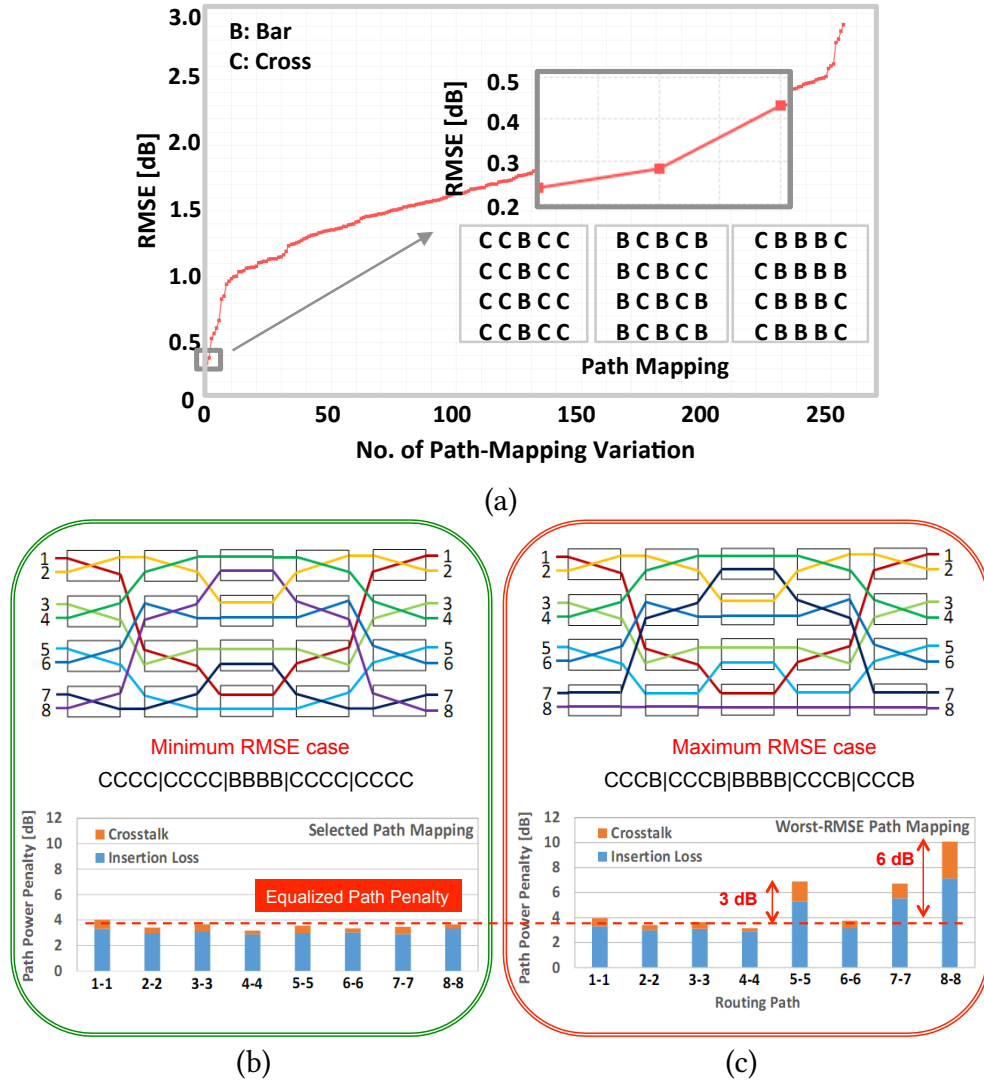
Figure 4.14:   (a) Plot of RMSE of the input-output path power penalties for the (1,2,3,4,5,6,7,8)→(1,2,3,4,5,6,7,8) mapping. (b) Minimum RMSE configuration. (c) Maximum RMSE configuration [40].

4.5   Summary of the Chapter

In this chapter, the physical layer performance of silicon photonic switch fabrics based on Benes topology and Mach-Zehnder switch elements was discussed. The *loss* and *first-order (as well as higher-order and multi-path) crosstalk* were evaluated in a switch fabric by propagating the light inside the structure using the iteration method. We showed that large scale switches provide a repetition in terms of the number of switch configurations

that correspond to a particular input-output mapping. It was also demonstrated that an optimal routing table with fairly equalized path power penalties can be constructed based on the physical layer performance of the switch fabric.

Chapter 5

Design Space Exploration of Microring Resonators

Abstract–*In this chapter, two models are introduced to characterize the design space of microring resonators in silicon photonics platform. The first model aims at predicting how the optical mode coupling present at the interface between a ring and a waveguide is affected by geometrical parameters, namely the ring-waveguide distance and ring radius. The concept of the curvature function of coupling is proposed to evaluate the impact of the nonuniform coupling region on the coupling coefficients. The second model relates the optical loss experimentally observable in the ring to its diameter. We then exemplify how our method can ease the selection of ring design parameters under certain constraints for WDM applications.*

5.1 Ring Resonator Modeling

The most accurate and complete way to explore the design space of ring resonators (i.e. estimation of the coupling strength and radiation loss) is by means of brute-force full-wave methods, such as three-dimensional finite-difference-time-domain (3D FDTD) [96, 116, 26] or finite-element method (FEM) [84, 138]. As shown in Fig. 5.1, such numerical simulations do not include any intermediate steps and once setup, they can directly calculate the desired parameters. Although attractive for simulating structures of limited complexity, these methods are based on time-consuming calculation procedures, which render them rather ineffective when applied to complex structures such as higher-order MRR filters [113, 146]. Even for a simple MRR structure, the coupling strength between the ring and the bus waveguides depends on multiple parameters, namely the width and

height of waveguides, radius of the ring, the wavelength of operation, and the gap between the ring and the waveguide. *This complexity makes the design space exploration of MRRs by means of FDTD or FEM simulations particularly cumbersome.*

Alternatively, the design space of MRRs can be explored by means of abstract compact models. A ring of the radius R can be defined by a loss coefficient α in units of cm^{-1} or dB/cm (hence a round-trip loss of L = $\exp(-2\pi R\,\alpha_{1/\mathrm{cm}})$ or $\mathrm{L_{dB}} = 2\pi R\,\alpha_{\mathrm{dB/cm}}$, and one (all-pass configuration), two (first-order add-drop configuration), or even more (higher-order add-drop configuration [146]) through and cross coupling coefficient(s) (t and κ – unitless parameters) describing the coupling strength of the electric and magnetic fields inside the ring with the adjacent waveguide(s). As long as the coupling between the ring and waveguides is assumed lossless, the relation $\kappa^2 + t^2$ = 1 holds, hence knowing only one of the two parameters is sufficient. These models are key to understand the dynamics of the resonant behavior [105], and to quickly seize the principles at play for a large-scale system. However, they do not capture the relationships between the physical dimensions of the structure, the fabrication-induced effects, and the coefficients used. As further discussed by Nikdast *et al.* [128], Le Maitre *et al.* [90], and Lu *et al.* [109] fabrication-induced effects are of major importance when a structure as compact as a MRR is considered. Compact models agnostic to the physical dimensions of the structure are thus insufficient to clearly assess the capabilities of MRRs for silicon photonic interconnects in practice. *This motivates us to establish accurate yet efficient models to bridge the gap between abstract compact models, time-consuming full-wave simulations, and the fabrication aspects of MRRs for silicon-on-insulator (SOI) platform.* The resulting models allow us to sweep a design space corresponding to different wavelengths, ring radii, and coupling gap sizes. For each considered ring, we calculate various figures of merit and exclude the designs failing to fulfill a set of minimal requirements. Finally, by considering the remaining "realm of feasibility", we derive conclusions on MRR capabilities.

As shown in Fig. 5.1, our proposed approach for estimating coupling coefficients re-
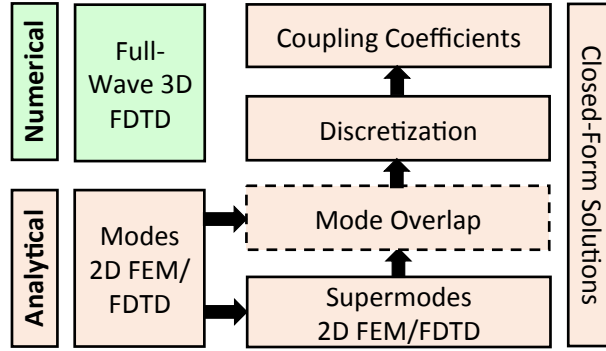
Figure 5.1: Comparison between the numerical method and our proposed analytical method for estimating the coupling between a ring and a waveguide. A fully numerical approach requires only one step while our analytical approach requires intermediate steps for calculating optical modes. In return, the analytical approach provides closed-form solutions for the coupling coefficients.

quires some intermediate steps for calculating the optical modes of the waveguides [see Fig. 5.1]. This can be realized by building a waveguide cross-section specific database of effective indices inside an isolated waveguide (for different wavelengths), as well as inside a coupled pair of waveguides (for different wavelengths and gaps).

5.2 Coupling of 2D Dielectric Slab Waveguides

In order to find the coupled modes of two dielectric slab waveguides, we can take use of the transfer matrix method for interfaces. We start by considering the case where the two waveguides are identical, i.e. they have the same width and are made of the same dielectric materials. Consider the two waveguides of width w placed at a distance g from each other as depicted in Fig. 5.2. In order for the optical modes to establish a strong coupling, the condition of *Phase Matching* must be satisfied. This means that the coupled supermodes are formed from the identical modes of the two waveguides. For example, the TE$_0$ mode of the left waveguide can only coupled to the TE$_0$ mode of the right waveguide. This interaction is either a "positive" coupling in which an even mode is formed or is a "negative" coupling in which an odd symmetry is formed in the field profile.
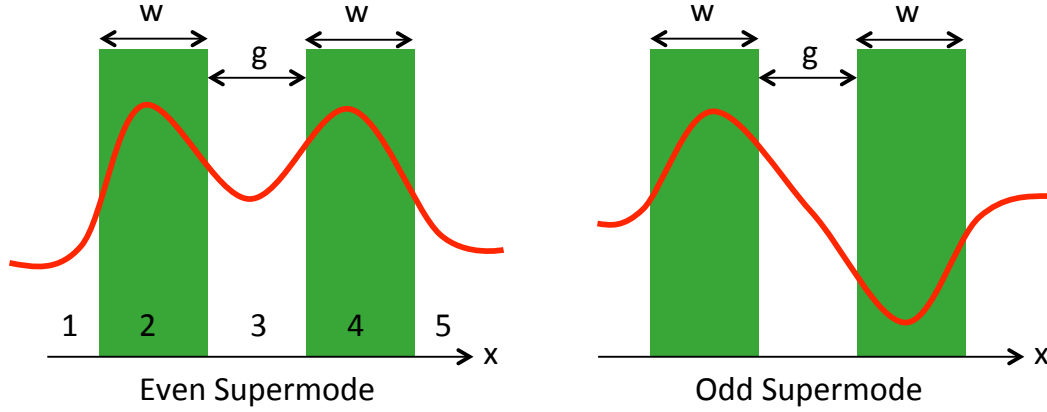
Figure 5.2: (left) Even supermode of two identical dielectric slab waveguide. A perfect symmetry is observed in the field profile. (right) Odd supermode of two identical waveguides.

Since there are four interfaces in this structure, the overall transfer matrix is written as

$$\mathbf{Q}_{1\to5} = \mathbf{Q}_{4\to5}(x = g + 2w) \times \mathbf{Q}_{3\to4}(x = g + w) \times \mathbf{Q}_{2\to3}(x = w) \times$$
$$\mathbf{Q}_{1\to2}(x = 0) \tag{5.1}$$

and therefore the optical modes are the solutions to the mode equation:

$$\mathbf{Q}_{1\to5} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \Rightarrow |q_{22}| = 0 \tag{5.2}$$

Note that polarization of the mode is included in the type of the transfer matrix to be used (i.e. TE interface matrix or TM interface matrix). After finding the effective index of the optical coupled mode, the field profiles can be easily obtained from the transfer matrix of each interface. If the modal index of the even mode is used, the field profile is automatically calculated to have an even distribution. If the modal index of the odd mode is used, the field profile will end up with an odd distribution. Figure 5.3a shows the calculated effective modal index for the even and odd TE_0 modes of a 450 nm silicon slab

125
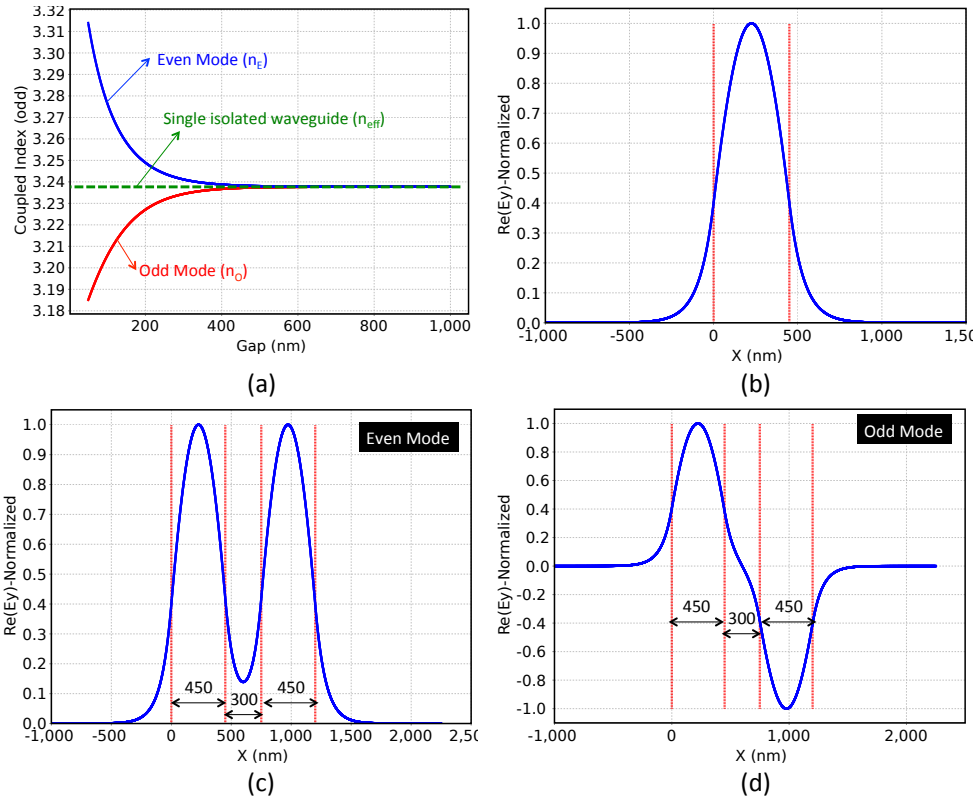
Figure 5.3: (a) Calculation of the even and odd effective indices of two identical silicon slab waveguides with $w = 450$ nm and $g = 300$ nm. The index of a single isolated waveguide is also plotted as a reference. (b) TE$_0$ mode profile of the isolated waveguide. (c) Even supermode of two identical waveguides. (d) Odd supermode of two identical waveguides.

waveguide with a gap size of 300 nm. As can be seen, when the gap is too large (> 600 nm), the effective indices of both even and odd modes approach the index of a single isolated waveguide. This means that there is no coupling between the two waveguides. Figure 5.3b shows the E_y component of the TE$_0$ mode of an isolated slab waveguide. Figure 5.3c shows the calculated mode profile of the coupling by using the effective index of the even supermode. As can be seen, the field profile has a perfect even symmetry. Likewise, Fig. 5.3d shows the calculated field profile of the coupled mode based on the effective index of the odd mode. In this case, a perfect asymmetry is observed.

5.3 Coupling of 3D Silicon Waveguides

Unlike 2D dielectric slab waveguides, the 3D strip waveguides and ridge waveguides do not have closed-form solutions for the effective indices of the supermodes. Although it is possible expand the approximate effective index method to the coupled waveguides, it fails to provide accurate enough solutions for the coupled modes. Therefore, we used COMSOL to create a database of the effective indices of the even and odd supermodes (for 400×220 nm, 450×220 nm, and 500×220 nm strip waveguides) in order to deliver an even and odd index for any gap and wavelength using interpolation. To populate this database, the coupling gap distance separating the coupled straight waveguides is swept from 50 nm to 1000 nm (on a step of 5 nm) while the wavelength is swept from 1500 nm to 1600 nm on the step of 2.5 nm. Figure 5.4a shows the simulated mode profiles for 450×220 nm strip waveguides in COMSOL. Figure 5.4b shows the effective index of the even and odd modes as function of wavelength for 200nm gap and Fig. 5.4c shows the effective index as function of gap for 1550nm wavelength.

5.4 Transfer Matrix of Coupling

Now we consider a directional coupler structure as shown in Fig. 5.5. Region 1 on the left side indicates the optical mode for $z < 0$ given by

$$\Psi_L(x, y, z) = a_1(z)\Psi_1(x, y)\exp(-j\beta_1 z) + a_2(z)\Psi_2(x, y)\exp(-j\beta_2 z) \qquad (5.3)$$

where $\Psi_1(x, y)$ and $\Psi_2(x, y)$ are the individual modes (typically the electric field) of each waveguide and $a_1(z)$ and $a_2(z)$ correspond to the amplitudes. Region 2 in the middle indicates the coupling region in which the optical mode can be expressed as the superposition

(a)

(b)

(c)
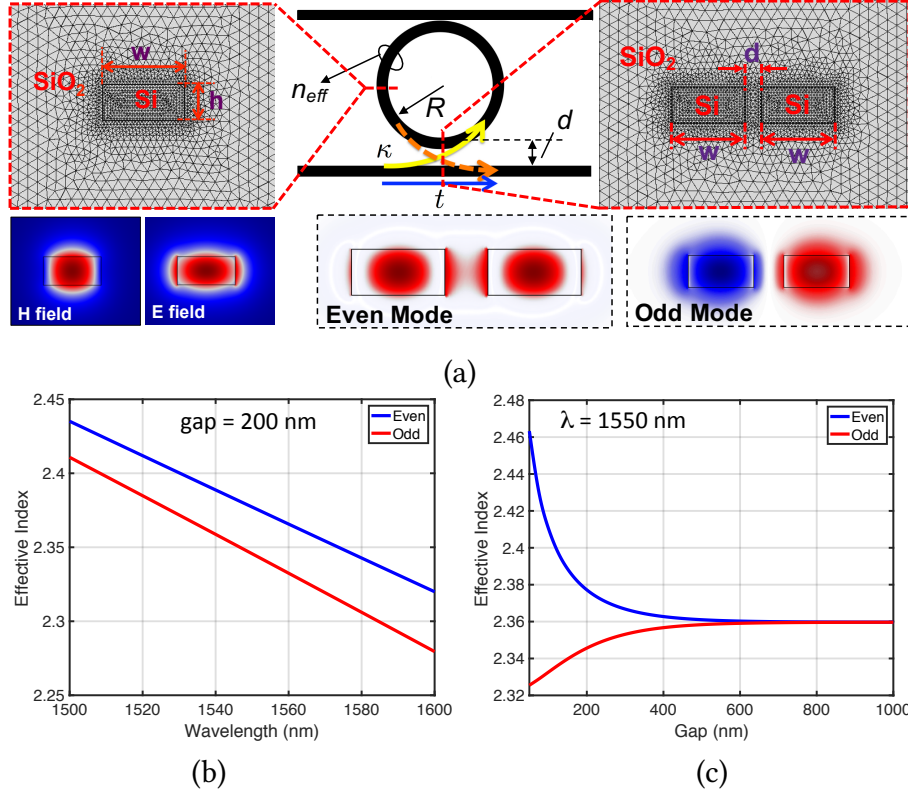
Figure 5.4: (a) Schematic of a ring resonator with two coupling waveguides. The cross-section of single strip and coupled strip waveguides and the mode profiles from COMSOL are also shown. (b) Effective index of even and odd supermodes as function of wavelength for $450{\times}220$ strip waveguide. (c) Effective index as function of gap.

of the even, Ψ_e, and odd, Ψ_o, coupled supermodes:

$$\Psi_C(x, y, z) = A_e \Psi_e(x, y) \exp(-j\beta_e z) + A_o \Psi_o(x, y) \exp(-j\beta_o z). \qquad (5.4)$$

Finally, region 3 on the right side indicates the optical mode for $z > l$ given by

$$\Psi_R(x, y, z) = a_1(z) \Psi_1(x, y) \exp(-j\beta_1(z - l))$$

$$+ a_2(z) \Psi_2(x, y) \exp(-j\beta_2(z - l)). \qquad (5.5)$$

Note that the reference of the phase fronts is set at $z = 0$, i.e. the beginning of the coupling section.
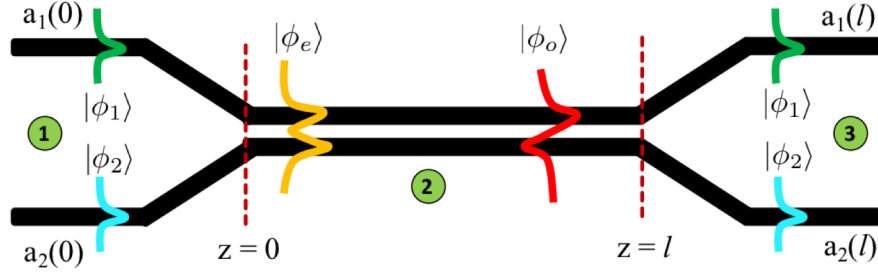
Figure 5.5: Structure of a directional coupler made of two straight waveguide in a very close proximity in the yz-plane. No coupling is assumed in region 1 and region 3.

In order to find the transfer matrix of this directional coupler that relates the amplitudes of the optical modes in regions 1 and 3, boundary conditions at $z = 0$ and $z = l$ must be satisfied. Applying the boundary condition at $z = 0$ results in

$$a_1(0)\Psi_1(x,y) + a_2(0)\Psi_2(x,y) = A_e\Psi_e(x,y) + A_o\Psi_o(x,y). \tag{5.6}$$

Using the orthogonality property of the modes (or supermodes) in the coupling region as given by

$$< \Psi_e|\Psi_o >= \iint_{-\infty}^{\infty} \Psi_e(x,y)\Psi_o^*(x,y)\, dx\, dy = 0 \tag{5.7}$$

results in

$$\begin{pmatrix} A_e \\ A_o \end{pmatrix} = \mathbf{M} \times \begin{pmatrix} a_1(0) \\ a_2(0) \end{pmatrix} \tag{5.8}$$

where

$$\mathbf{M} = \begin{pmatrix} \frac{<\Psi_e|\Psi_1>}{<\Psi_e|\Psi_e>} & \frac{<\Psi_e|\Psi_2>}{<\Psi_e|\Psi_e>} \\ \frac{<\Psi_o|\Psi_1>}{<\Psi_o|\Psi_o>} & \frac{<\Psi_o|\Psi_2>}{<\Psi_o|\Psi_o>} \end{pmatrix}. \tag{5.9}$$

129

Similarly, applying the boundary condition at $z = l$ results in

$$\begin{pmatrix} A_e \\ A_o \end{pmatrix} = \mathbf{E}^{-1} \times \mathbf{M} \times \begin{pmatrix} a_1(l) \\ a_2(l) \end{pmatrix} \tag{5.10}$$

where

$$\mathbf{E} = \begin{pmatrix} \exp(-j\beta_e l) & 0 \\ 0 & \exp(-j\beta_o l) \end{pmatrix}. \tag{5.11}$$

Combining these equations finally leads to the transfer matrix of the directional coupler:

$$\begin{pmatrix} a_1(l) \\ a_2(l) \end{pmatrix} = (\mathbf{M}^{-1}\mathbf{E}\mathbf{M}) \times \begin{pmatrix} a_1(0) \\ a_2(0) \end{pmatrix}. \tag{5.12}$$

The transfer matrix of the coupling region is therefore defined as

$$\mathbf{T} = \mathbf{M}^{-1}\mathbf{E}\mathbf{M}. \tag{5.13}$$

It is easy to verify that the normalization of the even and odd supermodes, i.e. $< \Psi_e | \Psi_e >$ and $< \Psi_o | \Psi_o >$ does not matter for the entries of this transfer matrix. To see this, first we introduce an abbreviation notation for the inner products:

$$I_{xy} \equiv < \Psi_x | \Psi_y > \tag{5.14}$$

where x and y index can be "e" or "o" or "1" or "2". With this notation, the $\mathbf{M}$ and $\mathbf{M}^{-1}$ matrices are written as

$$\mathbf{M} = \begin{pmatrix} \frac{I_{e1}}{I_{ee}} & \frac{I_{e2}}{I_{ee}} \\ \frac{I_{o1}}{I_{oo}} & \frac{I_{o2}}{I_{oo}} \end{pmatrix} \quad , \quad \mathbf{M}^{-1} = \frac{I_{ee}I_{oo}}{I_{e1}I_{o2} - I_{e2}I_{o1}} \begin{pmatrix} \frac{I_{o2}}{I_{oo}} & -\frac{I_{e2}}{I_{ee}} \\ -\frac{I_{o1}}{I_{oo}} & \frac{I_{e1}}{I_{ee}} \end{pmatrix} \tag{5.15}$$

Therefore, the elements of the **T** matrix are given by

$$T_{11} = \frac{1}{I_{e1}I_{o2} - I_{e2}I_{o1}} \left(I_{e1}I_{o2}E_e - I_{e2}I_{o1}E_o \right) \tag{5.16}$$

$$T_{12} = \frac{1}{I_{e1}I_{o2} - I_{e2}I_{o1}} \left(I_{e2}I_{o2}(E_e - E_o) \right) \tag{5.17}$$

$$T_{21} = \frac{1}{I_{e1}I_{o2} - I_{e2}I_{o1}} \left(-I_{e1}I_{o1}(E_e - E_o) \right) \tag{5.18}$$

$$T_{22} = \frac{1}{I_{e1}I_{o2} - I_{e2}I_{o1}} \left(-I_{e2}I_{o1}E_e + I_{e1}I_{o2}E_o \right) \tag{5.19}$$

which are completely independent of I_{ee} and I_{oo} elements. Here, $E_e = \exp(-j\beta_e l)$ and $E_o = \exp(-j\beta_o l)$. Therefore, we set the initial normalization of the even and odd super-modes to unity, i.e.

$$I_{ee} = < \Psi_e | \Psi_e > = 1 \, , \ I_{oo} = < \Psi_o | \Psi_o > = 1 \quad . \tag{5.20}$$

Note that the off-diagonal elements T_{21} and T_{12} can be simplified according to

$$E_e - E_o = \exp(-j\beta_+ l) \times (-2j) \, \sin(\beta_- l) \tag{5.21}$$

where

$$\beta_+ = (\beta_e + \beta_o)/2, \tag{5.22}$$

$$\beta_- = (\beta_e - \beta_o)/2. \tag{5.23}$$

For the case of the coupling of two identical waveguides, by considering the optical mode profiles plotted on Fig. 5.5, the following relations can be concluded:

$$I_{e1} = < \Psi_e | \Psi_1 > \approx < \Psi_e | \Psi_2 > = I_{e2} \equiv I_e \tag{5.24}$$

$$I_{o1} =<\Psi_o|\Psi_1> \equiv I_o \approx -<\Psi_o|\Psi_2> = -I_{o2}. \tag{5.25}$$

This leads to the transfer matrix elements given by

$$T_{11} = \frac{1}{-2I_e I_o}(-I_e I_o)(E_e + E_o) = \frac{1}{2}(E_e + E_o) \tag{5.26}$$

$$T_{12} = \frac{1}{-2I_e I_o}(-I_e I_o)(E_e - E_o) = \frac{1}{2}(E_e - E_o) \tag{5.27}$$

$$T_{21} = \frac{1}{-2I_e I_o}(-I_e I_o)(E_e - E_o) = \frac{1}{2}(E_e - E_o) \tag{5.28}$$

$$T_{22} = \frac{1}{-2I_e I_o}(-I_e I_o)(E_e + E_o) = \frac{1}{2}(E_e + E_o) \tag{5.29}$$

which results in the following transfer matrix:

$$\mathbf{M}^{-1}\,\mathbf{E}\,\mathbf{M} = \exp(-j\beta_+ l) \begin{pmatrix} \cos(\beta_- l) & -j\sin(\beta_- l) \\ -j\sin(\beta_- l) & \cos(\beta_- l) \end{pmatrix}. \tag{5.30}$$

Note that the exact value of I_e and I_o is not important in this case as long as the assumptions $I_{e1} = I_{e2}$ and $I_{o1} = I_{o2}$ hold. This is a very good assumption for coupling of identical waveguides.

5.5   Formulation of Coupled Mode Theory

Considering the general form of the transfer matrix of coupling given by

$$\mathbf{T} = \mathbf{M}^{-1}\,\mathbf{E}\,\mathbf{M} = \exp\left(-j(\mathbf{M}^{-1}\mathbf{B}\mathbf{M})l\right) \tag{5.31}$$

where **B** is the diagonal matrix of the even and odd propagation constants, it is clear that this transfer matrix is the solution of a linear differential system of the form

$$\frac{d}{dz} \begin{pmatrix} a_1(z) \\ a_2(z) \end{pmatrix} = -j(\mathbf{M}^{-1}\mathbf{B}\mathbf{M}) \times \begin{pmatrix} a_1(z) \\ a_2(z) \end{pmatrix}. \tag{5.32}$$

for the region $0 < z < l$. Here, $\mathbf{M}^{-1}\mathbf{B}\mathbf{M}$ is a constant matrix whose entries are given by

$$(\mathbf{M}^{-1}\mathbf{B}\mathbf{M})_{11} = \frac{1}{I_{e1}I_{o2} - I_{e2}I_{o1}} \left(I_{e1}I_{o2}\beta_e - I_{e2}I_{o1}\beta_o \right) \tag{5.33}$$

$$(\mathbf{M}^{-1}\mathbf{B}\mathbf{M})_{12} = \frac{1}{I_{e1}I_{o2} - I_{e2}I_{o1}} \left(I_{e2}I_{o2}(\beta_e - \beta_o) \right) \tag{5.34}$$

$$(\mathbf{M}^{-1}\mathbf{B}\mathbf{M})_{21} = \frac{1}{I_{e1}I_{o2} - I_{e2}I_{o1}} \left(-I_{e1}I_{o1}(\beta_e - \beta_o) \right) \tag{5.35}$$

$$(\mathbf{M}^{-1}\mathbf{B}\mathbf{M})_{22} = \frac{1}{I_{e1}I_{o2} - I_{e2}I_{o1}} \left(-I_{e2}I_{o1}\beta_e + I_{e1}I_{o2}\beta_o \right) \tag{5.36}$$

Therefore, the entries of the $\mathbf{M}^{-1}\mathbf{B}\mathbf{M}$ matrix are resembling the propagation constants. If the two waveguides are identical, then this matrix simplifies to

$$\mathbf{M}^{-1}\mathbf{B}\mathbf{M} = \begin{pmatrix} (\beta_e + \beta_o)/2 & (\beta_e - \beta_o)/2 \\ (\beta_e - \beta_o)/2 & (\beta_e + \beta_o)/2 \end{pmatrix} = \begin{pmatrix} \beta_+ & \beta_- \\ \beta_- & \beta_+ \end{pmatrix} \tag{5.37}$$

in which β_+ and β_- are defined as before. This differential equation shows that the amplitude of the waves inside each waveguide is coupled to the amplitude of the wave in the other waveguide. This is another interpretation of the supermodes that exist in the coupling region. This coupling equation can be also rigorously derived from Maxwell's equations using the perturbation theory.
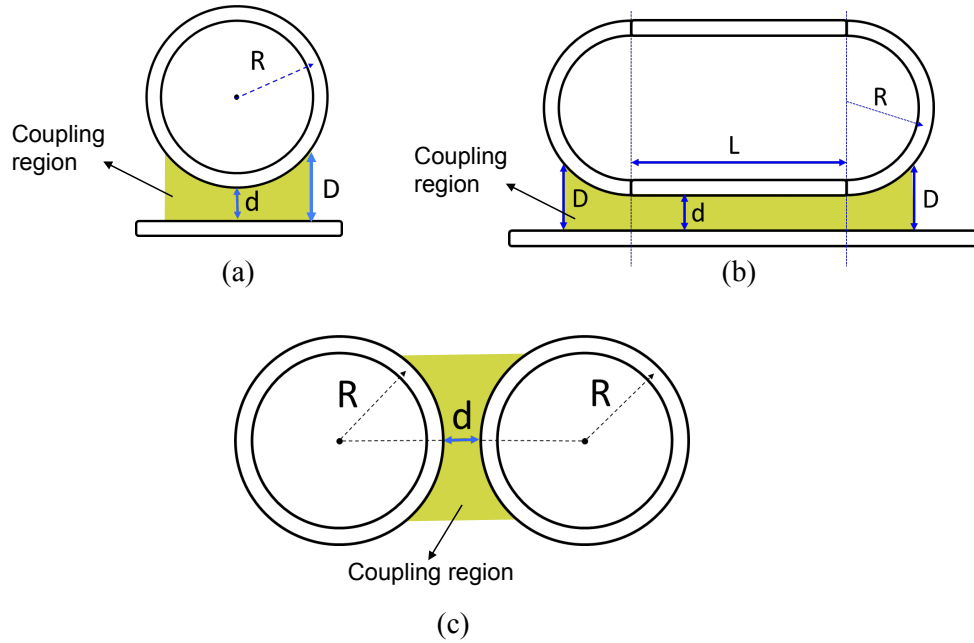
Figure 5.6: Examples of common cases of nonuniform coupling between two waveguides. (a) Coupling of a circular waveguide to a straight waveguide. (b) Mixed coupling of a circular-straight waveguide to a straight waveguide. (c) Coupling of two circular waveguides.

5.6   Nonuniform Coupling: Introduction of the Curvature Function

Now we consider a case where the coupling region between the two waveguides is not uniform along the direction of propagation (i.e. along the z direction). This is a common case in structures where bent waveguides are involved. Typical examples are the coupling of a ring resonator to a straight waveguide (Fig. 5.6a), or a race-track configuration (Fig. 5.6b), or coupling of two ring resonators to each other (Fig. 5.6c), .

In order to calculate the coupling coefficients in these structures, we propose the following steps:

- **Step 1**: Discretize the nonuniform coupling region between the two waveguides into very small intervals. The coupling in each small interval can be considered as a uniform coupling whose transfer matrix is readily available.

- **Step 2**: The total transfer matrix is obtained by cascading the transfer matrices of all the small intervals. The accuracy of the result increases as the number of intervals is increased.

- **Step 3**: The summation over all the small intervals can be turned into a Riemann Sum and in the limit into an integral form.

Considering a rectangular cross-section for the silicon waveguides (e.g. 450×220 nm), we intend to capture the dependence of the cross and through coupling coefficients (κ and t) on the ring radius and the gap size separating a ring structure from the waveguide. For the case of coupling of two identical waveguides, these coefficients are defined based on the transfer matrix of coupling:

$$\begin{pmatrix} \cos(\beta_- l) & -j\sin(\beta_- l) \\ -j\sin(\beta_- l) & \cos(\beta_- l) \end{pmatrix} = \begin{pmatrix} t & -j\kappa \\ -j\kappa & t \end{pmatrix}. \tag{5.38}$$

Therefore $t = \cos(\beta_- l)$ and $\kappa = \sin(\beta_- l)$. Although it is obvious that κ has a strong dependence on the minimum gap separating the ring from the bus (denoted by d in Fig. 5.6a), which is typically described by an exponential fitting, its dependence on the radius is not clear at first glance. Figure 5.7a demonstrates the impact of radius and the coupling gap (minimum distance between the ring and the waveguide) on the ring-waveguide electric field coupling coefficients. The colored area around each ring indicates the decaying tail of the optical mode inside the ring. Although this tail is infinite in practice, for the sake of the example, let us assume that no coupling occurs beyond the colored region. The first and second structures have the same gap distance, but the larger radius of the ring in the first structure results in a longer coupling region (interaction region between ring and waveguide denoted by double-arrow). A larger coupling region then leads to a larger cross coupling coefficient (κ). The second and third structures have the same radius, but the smaller gap in the third one results in a larger coupling region, hence stronger coupling.

(a)

(b)
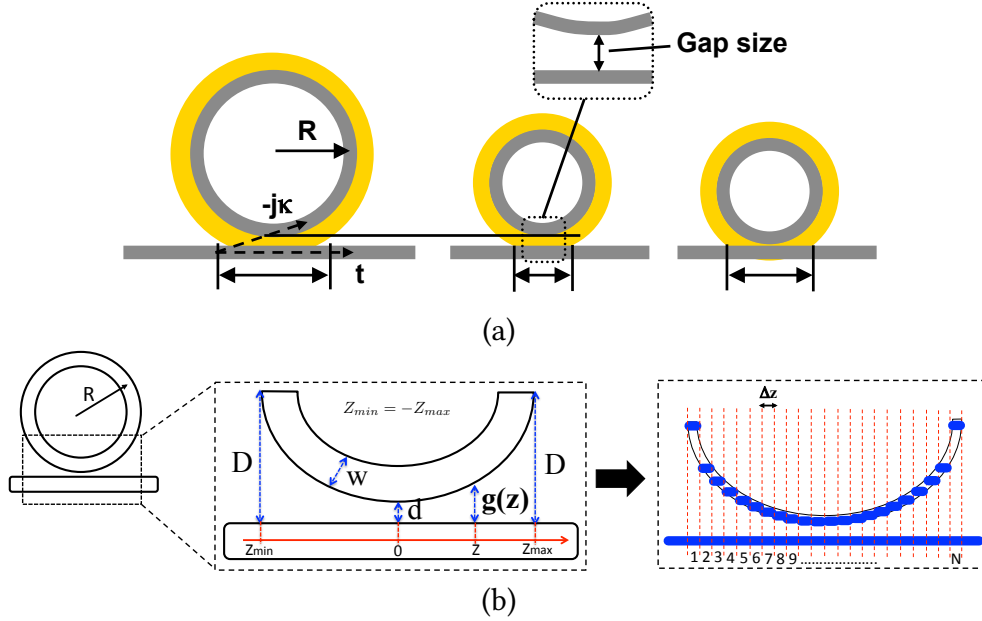
Figure 5.7: (a) Effect of ring radius and minimum coupling gap size on the thru and cross coupling coefficients of the ring-bus structure. (b) Example of discretization of the nonuniform coupling between two waveguides. In this example, the coupling of a circular waveguide to a straight waveguide is shown.

As shown in Fig. 5.7b, the coupling gap along the coupling region, $g(z)$, can be estimated by the following equation:

$$g(z) = d + (R + w/2) - \sqrt{(R + w/2)^2 - z^2} \tag{5.39}$$

where R is the radius of the ring (center to midpoint), w is the width of the ring, d is the minimum gap distance between the ring and the waveguide, and z is the distance relative to the point where the coupling gap is at its minimum (for this particular structure it is in the middle). Note that $g(0) = d$, and $g(R + w/2) = d + (R + w/2)$. Assuming that the coupling is only significant within a distance D ($\approx 1 \mu$m), the length of the coupling region is thus $Z_{\text{max}} - Z_{\text{min}} = 2 Z_{\text{max}}$ (as $Z_{\text{min}} = -Z_{\text{max}}$ due to the symmetry), with

$$Z_{\text{max}} = \sqrt{(D - d)(2(R + w/2) - (D - d))}. \tag{5.40}$$

It is clear that unlike a directional coupler made of two straight waveguides where the coupling region is uniform, due to the curvature of the ring structure, the coupling region formed between the ring and the waveguide is nonuniform. To properly capture the through and cross coupling coefficients of the structure (t and κ), this non-uniformity must be taken into account.

To this aim, consider a uniform discretization of the non-uniform region into N very small intervals of size $\Delta z = (Z_\text{max} - Z_\text{min})/N$ in the z direction, as shown in Fig. 5.7b. The gap distance corresponding to each position on the z-axis (i.e. z_i, $i = 1, 2, \ldots, N$), denoted by g_i, can be obtained from Eq. (5.39). Assuming the ith interval extends from z_{i-1} to z_i, its corresponding gap can be estimated by $g_i = g((z_{i-1} + z_i)/2)$. Small intervals permit to consider a uniform coupling within each interval. We then write the transfer matrix of the ith interval as

$$T_i = \exp(-j\phi_i) \begin{bmatrix} t_i & -j\kappa_i \\ -j\kappa_i & t_i \end{bmatrix} \tag{5.41}$$

where

$$\phi_i = \frac{2\pi}{\lambda} \Delta z \frac{n_\text{even}[g_i] + n_\text{odd}[g_i]}{2}, \tag{5.42a}$$

$$t_i = \cos\left(\frac{2\pi}{\lambda} \Delta z \frac{n_\text{even}[g_i] - n_\text{odd}[g_i]}{2}\right), \tag{5.42b}$$

$$\kappa_i = \sin\left(\frac{2\pi}{\lambda} \Delta z \frac{n_\text{even}[g_i] - n_\text{odd}[g_i]}{2}\right). \tag{5.42c}$$

In these equations, $n_\text{even}[g_i]$ and $n_\text{odd}[g_i]$ are the effective indices of the even and odd supermodes as a function of the gap distance g_i at a given wavelength. For an identical pair of coupled waveguides, these functions can be fitted by exponential curves as a function of gap size, and thus written as

$$n_E \approx n_\text{eff} + a_E \exp(-\gamma_E g) \tag{5.43}$$
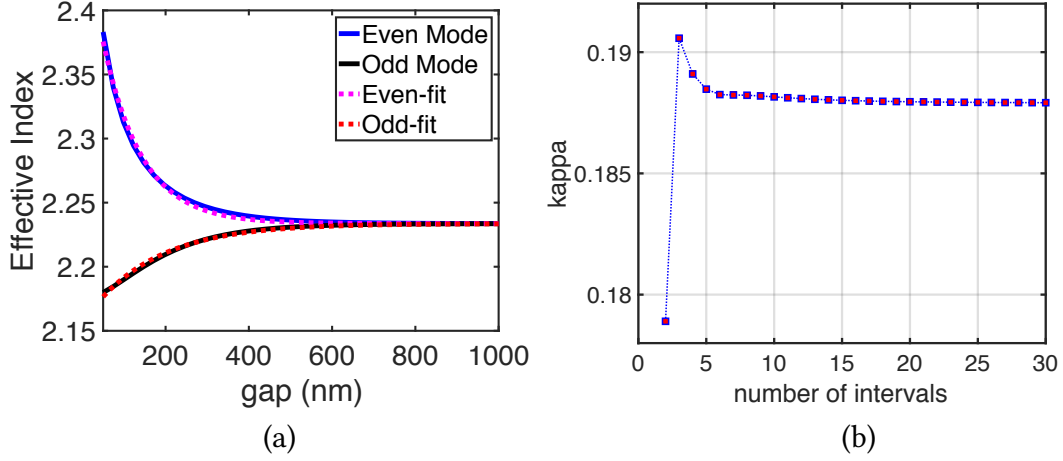
Figure 5.8: (a) Exponential fitting to the effective index of even and odd supermodes of two strip waveguides. (b) Example of convergence of the discretization method for kappa ($450{\times}220$ nm strip waveguide, λ = 1550nm, R = 5 μm, gap = 150 nm).

for the even mode, and

$$n_O \approx n_{\text{eff}} - a_O \exp(-\gamma_O g) \tag{5.44}$$

for the odd mode. An example of exponential fitting is shown in Fig. 5.8a. n_{eff} is the effective index of the optical mode of each waveguide in the absence of coupling.

The full impact of the coupling region can then be evaluated by cascading the intervals, thus by multiplying the transfer matrices of all the intervals. The resulting global transfer matrix $T = T_N \times T_{N-1} \times \ldots T_2 \times T_1$, has the same form as Eq. (5.41) and is given by

$$T = \exp(-j\phi_+) \begin{bmatrix} \cos(\phi_-) & -j\sin(\phi_-) \\ -j\sin(\phi_-) & \cos(\phi_-) \end{bmatrix} \tag{5.45}$$

where

$$\phi_+ = \frac{2\pi}{\lambda} \sum_{i=1}^{N} \frac{n_{\text{even}}[g_i] + n_{\text{odd}}[g_i]}{2} \Delta z, \tag{5.46a}$$

$$\phi_- = \frac{2\pi}{\lambda} \sum_{i=1}^{N} \frac{n_{\text{even}}[g_i] - n_{\text{odd}}[g_i]}{2} \Delta z. \tag{5.46b}$$

Therefore coupling coefficients (t and κ) are obtained simply by inspecting the ele-

ments of T matrix in Eq. (5.45), i.e. $t = \cos(\phi_-)$ and $\kappa = \sin(\phi_-)$. Figure 5.8b shows an example of the evaluation of kappa as function of number of intervals used in the discretization. It is seen that the convergence is fast and a stable solution is found after 20 intervals. In the limit where $N \rightarrow \infty$, the summation in these equations turns into an integral since $\Delta z \rightarrow 0$. Using the two exponential fittings of n_E and n_O we can then turn Eq. (5.46b) into an integral form and write κ as

$$\kappa = \sin\left(\frac{\pi}{\lambda}\int_{z_{min}}^{z_{max}}[n_E(z) - n_O(z)]dz\right) \tag{5.47}$$

where $Z_{\text{min}} = -Z_{\text{max}}$ due to the symmetry. In order to calculate the integral, the gap must be expressed as a function of z. It is easier to use polar coordinates where

$$z = (R + w/2)\sin(\theta), \tag{5.48}$$

$$g(z) = d + (R + w/2)(1 - \cos(\theta)), \tag{5.49}$$

$$z_{max} = (R + w/2)\sin(\theta_0), \tag{5.50}$$

and θ_0 is the angle that corresponds to the maximum coupling distance between the ring and the waveguide. It can be calculated by

$$\theta_0 = \cos^{-1}\left(1 - \frac{D-d}{R + w/2}\right). \tag{5.51}$$

Using the above equations, the integral in Eq. (5.47) is reduced to

$$
\int_{z_{min}}^{z_{max}} [n_E(z) - n_O(z)]dz =
$$

$$
2\frac{a_E x_E}{\gamma_E} e^{-\gamma_E d} \int_0^{\theta_0} e^{-x_E(1-\cos(\theta))} \cos(\theta)d\theta +
$$

$$
2\frac{a_O x_O}{\gamma_O} e^{-\gamma_O d} \int_0^{\theta_0} e^{-x_O(1-\cos(\theta))} \cos(\theta)d\theta \tag{5.52}
$$

where we defined $x_{E,O} = \gamma_{E,O}(R + w/2)$. The two integrals in the above equations are values of the function $F(x, \phi)$ defined by

$$
F(x, \phi) = \int_0^{\phi} \exp(-x(1 - \cos\theta)) \cos\theta \, d\theta. \tag{5.53}
$$

Finally, we note that in the limit the coupling between the ring and the waveguide can go all the way to the side edges of the ring, i.e. $D \to d + R + w/2$ which results in $\theta_0 \to \pi/2$. In that case by defining $B(x) = 2x\, F(x, \pi/2)$, the result of the integration for kappa is given by

$$
\kappa = \sin\left(\frac{\pi}{\lambda}\left[\frac{a_E}{\gamma_E}e^{-\gamma_E d}B(x_E) + \frac{a_O}{\gamma_O}e^{-\gamma_O d}B(x_O)\right]\right) \tag{5.54}
$$

where a_E, a_O, γ_E, γ_O are the parameters of the fitted exponential curves of Eq. (5.43) and (5.44), which are *independent of the shape of the coupling region.* We define $B(x)$ as the curvature function of the coupling region and is given by

$$
B(x) = 2x \int_0^{\pi/2} \exp\left(-x(1 - \cos\theta)\right) \cos\theta \, d\theta \tag{5.55}
$$

for the ring-waveguide structure in Fig. 5.7. Equation (5.55) has a closed-form solution given by

$$
B(x) = \pi\, x \exp(-x)\left[I_1(x) + L_{-1}(x)\right] \tag{5.56}
$$

in which $I_1(x)$ is the modified Bessel function of the first kind of order 1 and $L_{-1}(x)$ is the

modified Struve function of the first kind of order -1. Based on the asymptotic behavior of $I_1(x)$ and $L_{-1}(x)$ functions for large x:

$$I_1(x) \approx \frac{\exp(x)}{\sqrt{2\pi x}} \quad , \quad L_{-1}(x) \approx I_1(x) \tag{5.57}$$

it can be verified that

$$B(x) \approx \sqrt{2\pi x}. \tag{5.58}$$

The parameters of function $B(x)$ in Eq. (5.56), and therefore in Eq. (5.58) are given by $x_E = \gamma_E(R + w/2)$ and $x_O = \gamma_O(R + w/2)$. Noting finally that since in general the argument of $\sin(\dots)$ function in Eq. (5.54) is small for coupling of ring to the waveguide, the approximation $\sin(t) \approx t$ can be applied:

$$\kappa \approx \frac{\pi}{\lambda} \left[\frac{a_E}{\gamma_E} e^{-\gamma_E d} B(x_E) + \frac{a_O}{\gamma_O} e^{-\gamma_O d} B(x_O) \right]. \tag{5.59}$$

Equation (5.54) is sufficient to fulfill our initial goal of evaluating κ as a function of radius and gap. It is worth noting that radius only appears in the $B(x)$ function in Eq. (5.54). In fact, the effect of the non-uniform coupling is solely captured by $B(x)$, and the four parameters a_E, a_O, γ_E, γ_O do not depend on the shape of the coupling region. For example, if a uniform coupling of length L and gap d between two straight waveguides is considered, $B(x) = x$ is obtained where $x_{E,O} = \gamma_{E,O} L$.

A summary of commonly used coupling structures and their curvature functions is presented in Fig. 5.9.

5.7   Model vs. FDTD for 2D Structures

We now aim at validating Eq. (5.54) against finite-difference time-domain (FDTD) simulations. We set up a 2D FDTD simulation of 450 nm wide slab waveguides with 10 nm
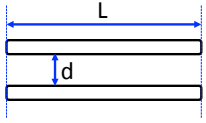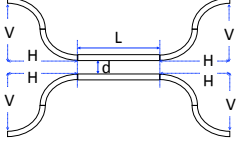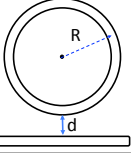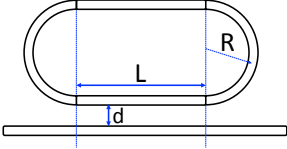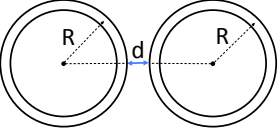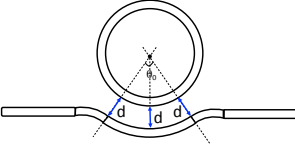
Structure	x	$B(x)$
	γL	x
	γL	$x \left[1 + \dfrac{2H}{L} \exp\left(-\dfrac{V}{L} x \right) I_0 \left(\dfrac{V}{L} x \right) \right]$
	$\gamma \left(R + \dfrac{w}{2} \right)$	$\pi x \exp(-x) \left[I_1(x) + L_{-1}(x) \right]$ $\approx \sqrt{2\pi x}$
	$\gamma \left(R + \dfrac{w}{2} \right)$	$\dfrac{L}{R + \frac{w}{2}} x + B_{\text{ring-wg}}(x)$
	$\gamma \left(R + \dfrac{w}{2} \right)$	$0.5\, B_{\text{ring-wg}}(2x)$
	$\gamma \left(R + \dfrac{w}{2} + \dfrac{d}{2} \right) \theta$	x

Figure 5.9: Summary of the curvature function of the coupling for different structures including uniform and circular coupling regions.

grid resolution in Lumerical FDTD, in order to numerically evaluate the coupling coefficients (t and κ) of the ring-waveguide structure. Fig. 5.10a shows the simulated structure consisting of a straight waveguide and half of a ring resonator. Perfectly matched layers (PML) boundary conditions are applied for the unidirectional transmission. Power monitors are placed at the input and output ports of the bus waveguide, as well as the coupling port. For the 2D example of 5 μm radius at 1550 nm wavelength, the coefficients of the model in Eq. (5.54) are extracted by a nonlinear least mean square error curve fitting as $a_E = 0.141188$, $a_O = 0.092605$, $\gamma_E = 0.012756$ nm^{-1}, $\gamma_O = 0.010761$ nm^{-1}, therefore $x_E = 66.65$, and $x_O = 56.2254$. Fig. 5.10b shows a plot of the curvature function $B(x)$

as a function of x. Figure 5.10c plots Eq. (5.54) against the results of 2D FDTD with 5 nm mesh size for $R = 5$ μm based on the extracted parameters of the modes and the curvature function. A maximum error of 3% is observed for 50 nm gap size while the error is progressively decreasing as the gap size is increased. This further advocates the validity of our model since fabricating gap sizes less than 100 nm is typically difficult and avoided.

(a)

(b)

(c)
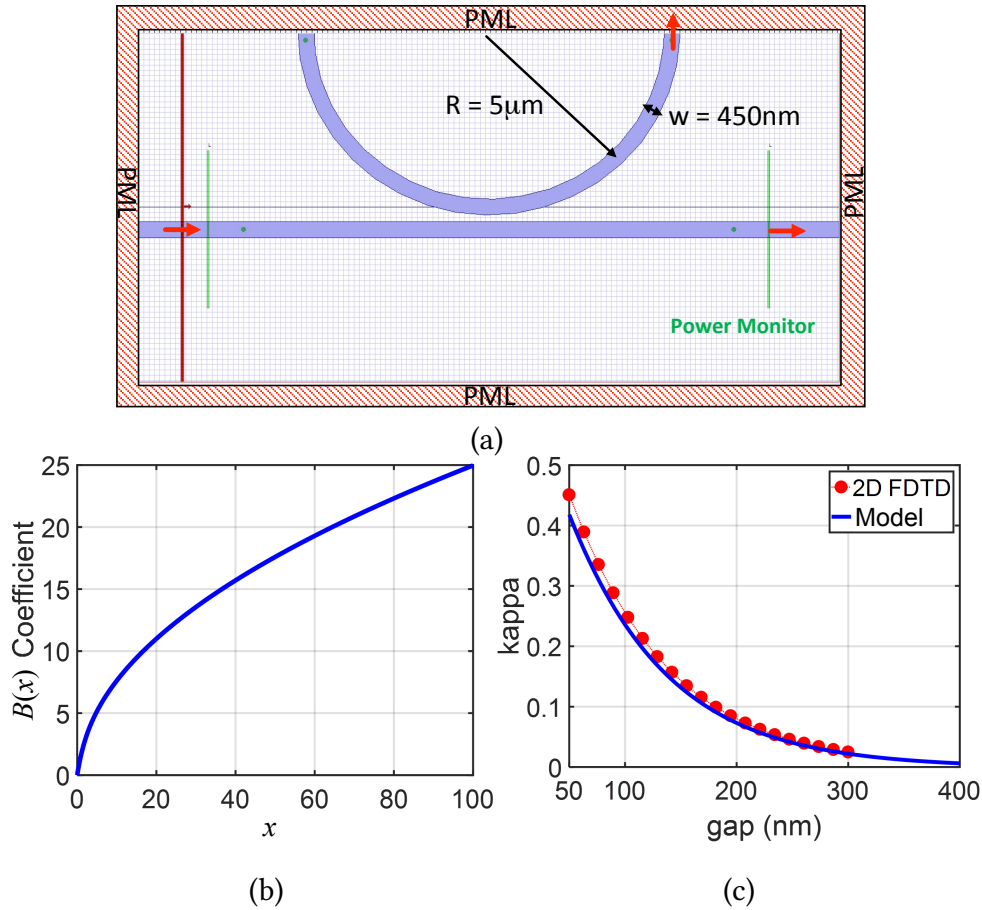
Figure 5.10: (a) FDTD simulation consisting of a straight waveguide and a half-circle ring. (b) Numerical evaluation of the circular curvature function $B(x)$ given in Eq. (5.54). (c) Comparison of results from model and 2D FDTD for $R = 5$ μm, $w = 450$ nm and various coupling gaps.

5.8 Model vs. FDTD for 3D Structures

Next, we examine our analytical approach for a 3D case where the rings and waveguides are all made of silicon strip waveguides with $450{\times}220$ nm cross-section. The modes of

Table 5.1: Parameters of the coupling model for the fundamental quasi-TE$_{00}$ mode of silicon strip waveguides at λ = 1550 nm.

Parameter	400$\times$220	450$\times$220	500$\times$220
a_E	0.242422	0.177967	0.132273
a_O	0.077526	0.049910	0.033840
γ_E	0.010687	0.011898	0.012783
γ_O	0.006129	0.006601	0.006911

a single waveguide and the supermodes of two coupled waveguides are calculated from COMSOL software as discussed for two identical waveguides. For this example, the parameters of the model are extracted as $a_E = 0.177967$, $a_O = 0.049910$, $\gamma_E = 0.011898$ nm^{-1}, $\gamma_O = 0.006601$ nm^{-1}, $x_E = 62.1671$, $x_O = 34.49$ for R = 5 μm at λ = 1550 nm, and then the curvature coefficients are calculated as $B(x_E) = 19.64$, and $B(x_O) = 14.57$. Table 1 summarizes the parameters of the model for 400$\times$220 nm, 450$\times$220 nm, and 500$\times$220 nm strip waveguides at λ = 1550 nm. It is seen that a_E and a_O have a stronger dependence on the width of the waveguide than γ_E and γ_O. Figure 5.11(a) shows the result of estimating t and κ with Eq. (5.54) at 1550 nm wavelength for 5 μm, 10 μm, and 20 μm radii. As can be observed, the curvature of the coupling region (i.e. the radius of the ring) has a noticeable effect on the coupling coefficients. Figure 5.11(b) plots the contours of kappa (cross coupling) as a function of gap and radius of the ring. Even though they approach a vertical line for very large radii, these contours are clearly non-vertical for smaller radii, demonstrating the significant effect of the radius on the coupling strength in that regime.

Next, we perform a series of 3D FDTD simulations with 10 nm grid size with Lumerical FDTD software [224]. To show the effect of three main physical parameters of the structure, i.e. gap, radius, and width of the waveguide, the wavelength of operation was set to the fixed value of 1550 nm. The structure in the simulations is the 3D version of the one shown in Fig. 5.10a. First, the cross-section of the waveguide was set to 400$\times$220 nm and then the FDTD simulations were performed by sweeping the gap size from 50 nm

(a) (b)
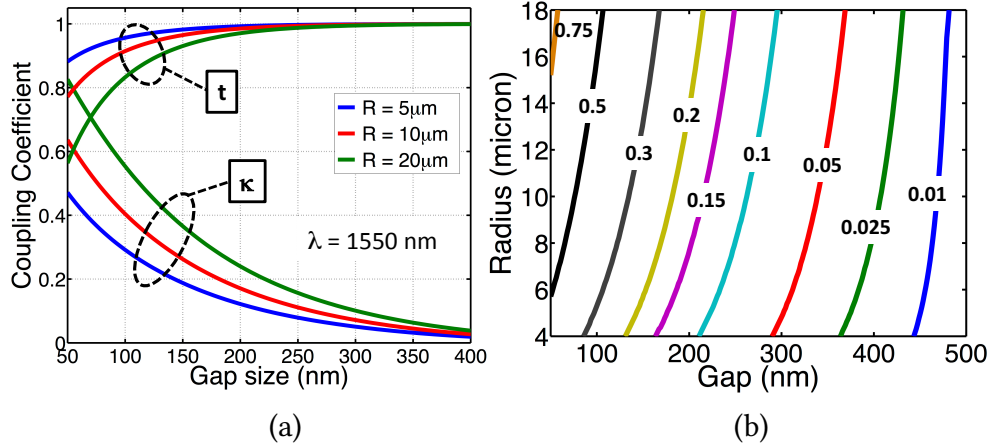
Figure 5.11: (a) Coupling coefficients (κ and t) as a function of coupling gap at 1550 nm with three different radii. (b) Contours of κ for various gap sizes and radii at 1550 nm.

(a) (b)

(c) (d)
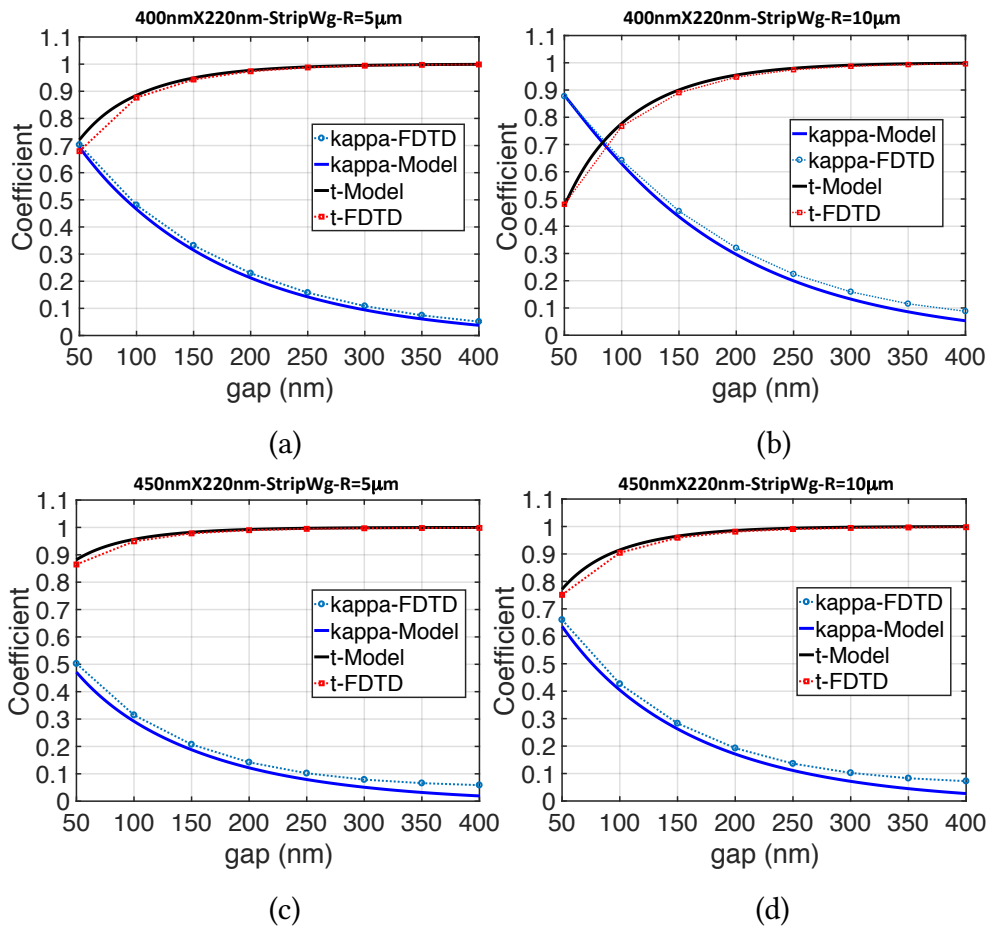
Figure 5.12: (a) Comparison of predicted ring-waveguide coupling coefficients (solid lines) and the 3D FDTD simulations (circled dotted) for $R = 5$ μm and $w = 400$ nm. (b) $R = 10$ μm and $w = 400$ nm. (c) $R = 5$ μm and $w = 450$ nm. (d) $R = 10$ μm and $w = 450$ nm.
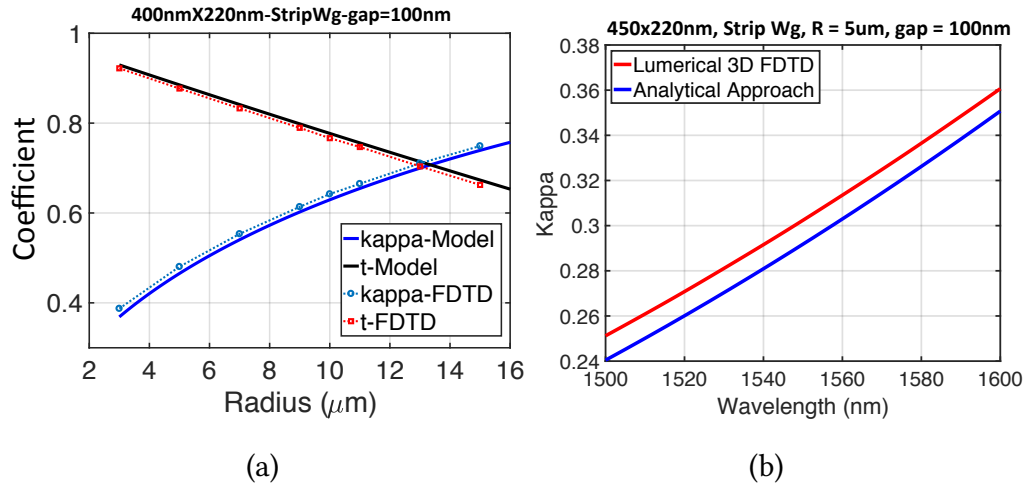
Figure 5.13: (a) Dependence of coupling coefficients on the radius of the ring for gap = 100 nm and w = 400 nm. The proposed model accurately captures the effect of radius as it closely resembles the results of full-wave 3D FDTD simulations. (b) Dependence of the kappa on the wavelength for R = 5μm, w = 400 nm, and gap = 100 nm.

to 400 nm on the step size of 50 nm for various radii. Fig. 5.12a and Fig. 5.12b show the dependence of κ and t as a function of gap size for R = 5 μm and R =10 μm, respectively. A good agreement is observed between our analytical model and 3D FDTD results. As expected, a larger radius results in a stronger coupling between the ring and waveguide, hence a larger value for κ. Next, the cross-section of the waveguide was set to 450×220 nm and the simulations were performed again. Fig. 5.12c and Fig. 5.12d depict the results for R = 5 μm and R = 10 μm, respectively. Compared to Fig. 5.12a and Fig. 5.12b it is seen that both the model and FDTD predict a weaker coupling as the width of the waveguide increases. This is due to the fact that the fundamental optical mode inside a wider waveguide has a higher effective index, hence it is more confined within the silicon core and its exponential tail outside the core is shorter.

Finally, to demonstrate the effect of radius on the coupling strength, we set the gap size to 100 nm and swept the radius from 3 μm to 15 μm. The results are shown in Fig. 5.13a for both FDTD and our model. A very good agreement is observed indicating that our proposed model in Eq. (5.54) fully captures the effect of physical parameters on the coupling strength. Since radius only appears in the curvature function $B(x)$ in Eq. (5.54),

the concluded rule-of-thumb is that ring-waveguide coupling coefficient scales up as the *square-root of the radius.* For example, increasing the radius by a factor of two (e.g. going from 5 μm to 10 μm) while maintaining the same gap size and waveguide dimensions will increase the kappa by a factor of $\sqrt{2}$. This can be easily verified in Fig. 5.13a.

The dependence of kappa on the wavelength is also investigated in Fig. 5.13b for a 450×220 nm strip waveguide with R = 5 μm and gap = 100 nm. It shows that the discretization approach and the FDTD results closely match each other. It is seen that kappa exhibits an *increasing* trend with waveguide. This is expected because the confinement of optical modes decreases with an increase of wavelength (i.e. $dn_{\text{eff}}/d\lambda < 0$), which leads to stronger coupling between waveguides. Moreover, we see that the error between FDTD and analytic approach is almost constant over the entire wavelength range of interest. This may be indicative of the fact that the error is not due to the waveguide or material dispersion and is most likely arises from the discretization procedure.

5.9 Model vs. Experimental Results

After validating our model of ring-waveguide coupling against full-wave FDTD simulations, three batches of test structures based on add-drop configurations were designed and fabricated through AIM Photonics [226] multi-project wafer run. The waveguides were chosen to be 400×220 nm strip waveguides since a 400 nm width will provide the strongest coupling coefficients. We used the "gdspy" open source library [221] to create the layout of the test structures. The first batch of structures have radii of 5 μm while the second and third batches are designed with 7.5 μm and 10 μm radii. Each batch includes five symmetric add-drop structures (i.e. *input gap size = drop gap size*) set to 100 nm, 150 nm, 200 nm, 250 nm, and 300 nm. Figure 5.14 shows the GDS layout of the first batch with 5 μm radius. TE-polarized vertical grating couplers were used to couple the light in and out of the silicon chip (Note that the rings are designed for TE polarization). Each
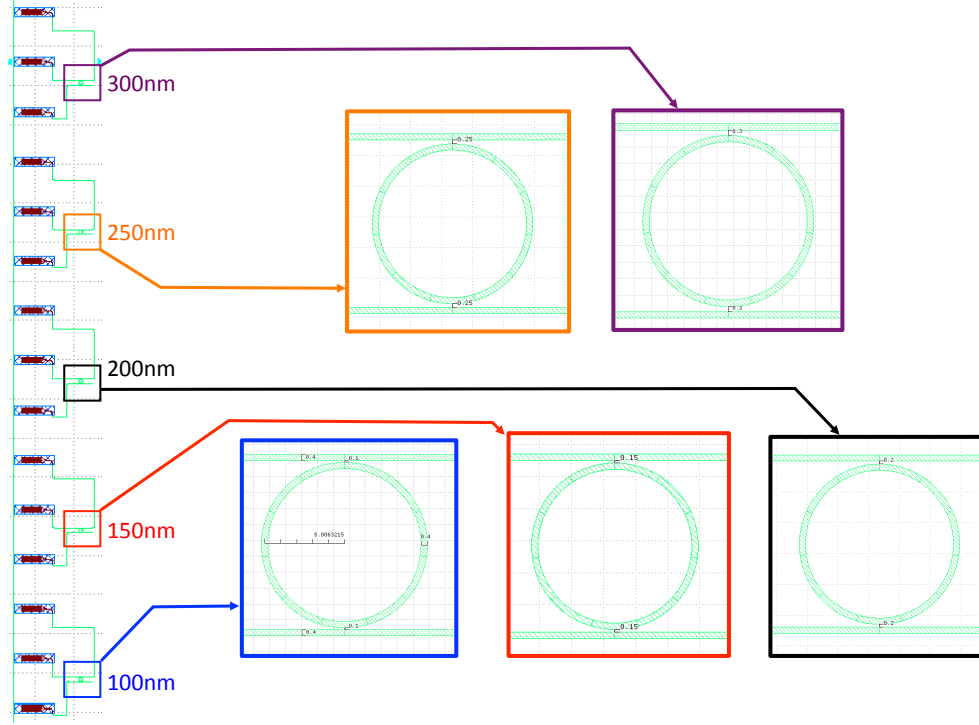
Figure 5.14: Layout of fabricated silicon rings through AIM Photonics MPW run (2016). This figure shows symmetric add-drop structures with 5 μm radius and gap sizes varying from 100 nm to 300 nm in 50 nm steps. The layout also shows the TE-polarized vertical grating couplers used for coupling light in and out. Each add-drop ring is connected to three grating couplers in this design.

add-drop structure is connected to three grating couplers for monitoring both the through spectral response and the drop spectral response.

The through and drop transmissions responses for the add-drop structure are given by [31]

$$TR = \left| \frac{t_{\text{in}} - t_{\text{drp}} \sqrt{L} \exp(-j\delta\phi)}{1 - t_{\text{in}} \, t_{\text{drp}} \sqrt{L} \exp(-j\delta\phi)} \right|^2 \tag{5.60a}$$

$$DR = \left| \frac{\kappa_{\text{in}} \, \kappa_{\text{drp}} \, L^{0.25} \exp(-j\delta\phi/2)}{1 - t_{\text{in}} \, t_{\text{drp}} \sqrt{L} \exp(-j\delta\phi)} \right|^2 \tag{5.60b}$$

where L is the round-trip optical power attenuation inside the ring, t_{in} and t_{drp} are the through coefficient of the input coupling region and the drop coupling region ($t^2_{\text{in,drp}} =$
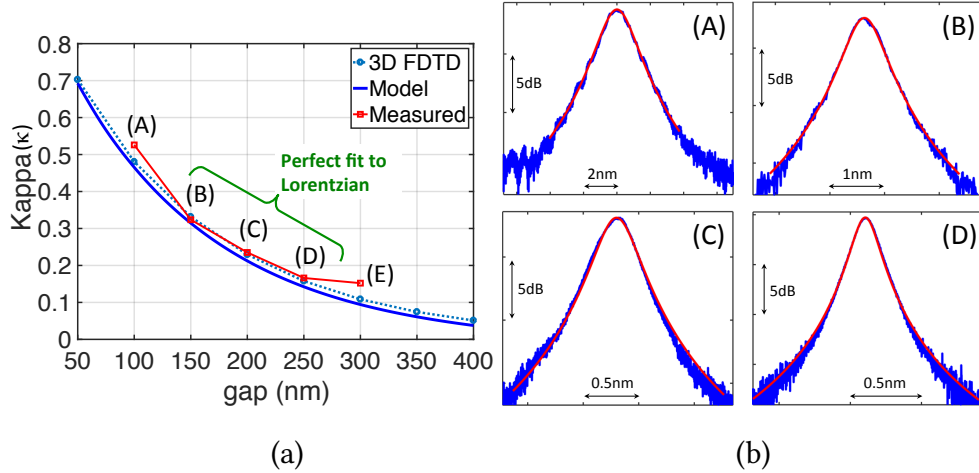
Figure 5.15: (a) Comparison of the predicted coupling coefficient (κ) of the ring-waveguide (solid line), the 3D FDTD simulations (dotted circle), and the extracted values by fitting to the Lorentzian response of the drop port of measured spectra. It can be seen that the model has a good agreement with the measurements. (b) Comparison of the measured spectra and the fitted Lorentzian for cases (A), (B), (C), and (D). The last three cases exhibit very good fit.

$1 - \kappa_{\mathrm{in,drp}}^2$), and $\delta\phi$ is the relative phase shift with respect to the target resonance:

$$\delta\phi \approx -\frac{\delta\lambda}{\mathrm{FSR}_{nm}} \times 2\pi \tag{5.61}$$

and $\delta\lambda = \lambda - \lambda_{\mathrm{res}}$. FSR is the free-spectral range of the resonance spectrum approximated as [31]:

$$\mathrm{FSR}_{nm} \approx \frac{\lambda_{\mathrm{res}(nm)}^2}{2\pi R_{nm}\, n_g}. \tag{5.62}$$

The minimum through power ($\mathrm{TR}_{\min}$) and the maximum drop power ($\mathrm{DR}_{\max}$) is obtained by setting $\delta\phi = 0$ whereas the maximum through power ($\mathrm{TR}_{\max}$) and the minimum drop power ($\mathrm{DR}_{\min}$) at FSR/2 distance from the resonance is obtained by setting $\delta\phi = \pi$. Finally, the half-power bandwidth (full width at half maximum) of the drop spectrum is given by

$$\Delta\lambda_{3\mathrm{dB}} \approx \mathrm{FSR}_{nm} \times \frac{1}{\pi} \cos^{-1}\left(1 - \frac{(1 - t_{\mathrm{in}}\, t_{\mathrm{drp}}\, \sqrt{L})^2}{2\, t_{\mathrm{in}}\, t_{\mathrm{drp}}\, \sqrt{L}}\right). \tag{5.63}$$

Note that due to the symmetry of all the test structures, it is assumed that $t_{\mathrm{in}} = t_{\mathrm{drp}}$ and
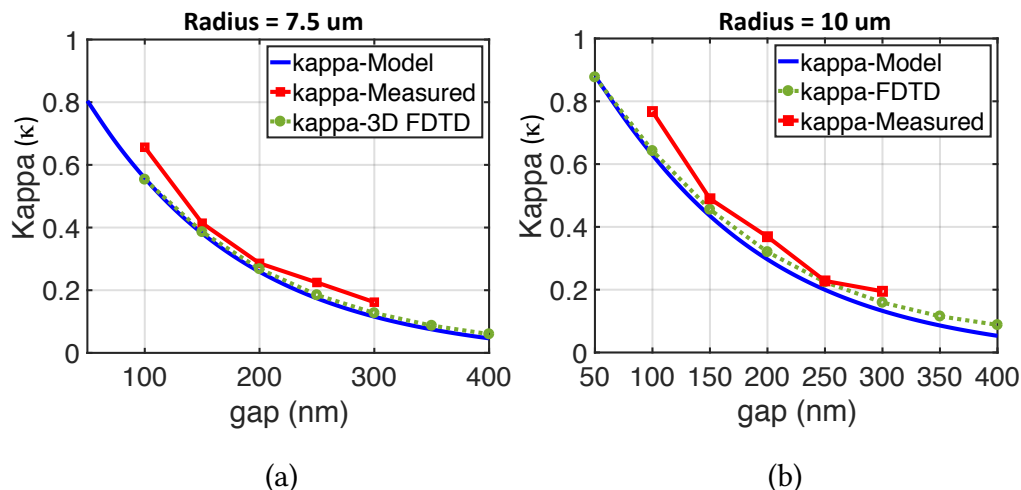
Figure 5.16: Comparison of extracted ring-waveguide coupling with the analytical model and FDTD for (a) 7.5 μm radius, and (b) 10 μm radius. As the radius increases, the coupling coefficient increases.

$\kappa_{\text{in}} = \kappa_{\text{drp}}$.

In order to extract the coupling of ring-waveguide, we use the drop transmission and perform a least mean square fitting of Eq. (5.60b) to the measured spectra. Note that in general determining the state of the ring resonator (under-coupled or over-coupled) is not possible without having knowledge of the phase response of the resonator. As pointed out in [120], this leads to an ambiguity in distinguishing between through coupling coefficients and the round-trip loss for all-pass structures (a ring coupled to a single bus waveguide). However, *since the test structures were designed to have equal input and drop coupling gaps, the ring resonators will never reach the critical coupling state and should stay under-coupled for any given gap size.* The following steps were taken to uniquely extract the coupling coefficients:

- **Step 1**: extract the FSR and 3dB bandwidth of the spectrum by fitting the normalized drop spectrum to the normalized measured spectrum (least-squares fitting).
- **Step 2**: use the extracted bandwidth to find the parameter $\xi = t^2\sqrt{L}$ from Eq. (5.63).
- **Step 3**: use the measured drop loss at the resonance to establish another relation between t and L.

- **Step** 4: solve the two found equations to uniquely determine t, κ, and L.

Fig. 5.15a shows the extracted ring-waveguide coupling coefficients for 5 μm radius batch. Fig. 5.15b depicts the quality of the spectral fitting to the measured spectra for cases (A), (B), (C), and (D). It is seen that in case (A) the measured spectrum is not smooth around the resonance while in cases (B), (C), and (D) the measured spectra are reasonably smooth. One possible explanation is the strong scattering of light due to the roughness of the sidewalls of the ring in the coupling region between the ring and the waveguide when the gap is 100 nm. The extracted coupling coefficients for these four cases are very close to what our model and 3D FDTD predicts. However, in case (E), corresponding to a gap of 300 nm, the extracted value ($\kappa \approx 0.16$) exhibits a noticeable error when compared to the predicted values ($\kappa \approx 0.1$).

As mentioned earlier, the curvature function in Eq. (5.54) includes the impact of the radius on the coupling coefficients. The predictions of the model show that larger radii will result in a stronger cross coupling between the ring and the waveguide. To verify this, the result of the extraction of kappa for R = 7.5 μm and R = 10 μm from our fabricated devices are plotted and compared against the predictions of the model and FDTD in Fig. 5.16a and Fig. 5.16b, respectively. For any given coupling gap size, the larger radius clearly provides a stronger coupling coefficient between the ring and the waveguides.

Impact of Back Scattering Inside the Ring

Figure 5.17a depicts the drop spectrum measured in the (E) case (300 nm gap) of Fig. 5.15a. As known in the literature [104], the measured spectrum indicates a clear resonance splitting with two peaks when the coupling of the ring to the adjacent waveguide is weak. The red curve on this figure is the fitted spectrum based on Eq. 5.60b. Although our fitting procedure appears to be able to decently estimate the 3dB bandwidth of the measured spectrum, it clearly fails to provide a good overall match.
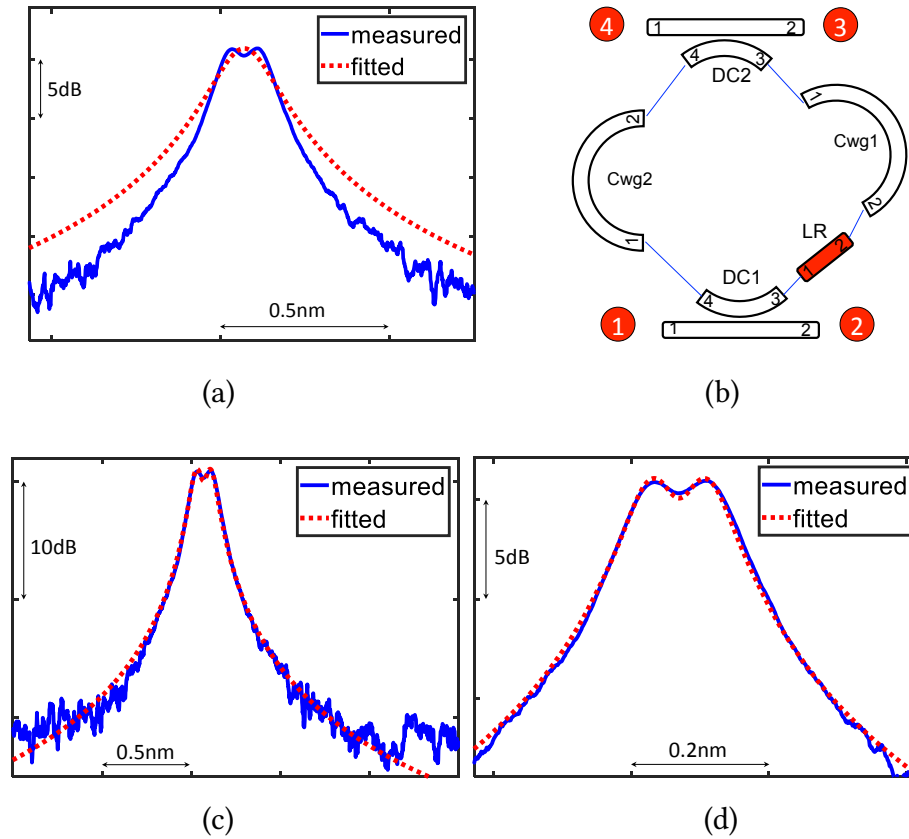
Figure 5.17: (a) Comparison of spectral fitting (solid red) to the measured spectrum (solid blue) for R = 5 μm and gap = 300 nm. A clear splitting of resonance is observed. (b) Modeling the backscattering inside the ring by including a lumped reflector. (c) Comparison of the spectral fitting with lumped reflector to the measured spectrum. A very good agreement is observed. (d) Close-up on a small region around the resonance. A symmetric splitting of resonance is observed.

The phenomenon of resonance splitting in microring resonators has been investigated in the literature and its origins have typically been associated with the roughness of the sidewalls of the ring [104, 93, 123, 124] and possible polarization errors. Recently, the presence of the coupling section has been highlighted as a contributor of reflection in rings with smaller gaps [94]. Such roughness causes backscattering of the optical mode that is circulating in one direction inside the ring. The backscattering results in the excitation of a degenerate optical mode circulating in the opposite direction inside the ring. The work presented in [94] has formulated a correction on Eq. (5.60b) to include the backscattering effects based on the temporal coupled mode theory, while the work in [104] has proposed an interpretation based on a lumped reflectivity inside the ring. In [94] asym-

metric heights of the split resonances were linked to additional reflections in the coupling section. However, in our data the prevalence of symmetric split resonances was observed. Therefore, we choose to model the backscattering effects by a lumped reflector. Fig. 5.17b shows an abstract representation of the structure to which we assimilate the ring. It consists of two lossless directional couplers (DC1 and DC2), two curved waveguides of equal length (Cwg1 and Cwg2) and a lumped reflector (LR) of null length with reflectivity r. Considering this compact model, the scattering parameter from the input port (port 1) to the drop port (port 4) is given by

$$S_{41} = \left(-\kappa_1 \kappa_2 L^{0.25} \exp(-j\phi/2)\right) \frac{t - O}{1 - 2tO + O^2} \tag{5.64}$$

where ϕ is the round-trip phase inside the ring, O is the loop gain of the ring given by

$$O = t_1 t_2 \sqrt{L} \exp(-j\phi) = \xi \exp(-j\phi). \tag{5.65}$$

$\kappa_1, \kappa_2, t_1, t_2$ are the coefficients of the input and output couplers (DC1 and DC2) and L is the round-trip power attenuation inside the ring without a reflector. The splitting of the resonance into two peaks can be characterized by defining $t = \cos(\phi_t)$ and then

$$\frac{t - O}{1 - 2tO + O^2} = -\frac{1}{2}\left(\frac{1}{O - e^{j\phi_t}} + \frac{1}{O - e^{-j\phi_t}}\right). \tag{5.66}$$

By expanding the round-trip phase around the resonance frequency as

$$\phi(\omega) = 2\pi R \left(\beta(\omega_0) + \frac{d\beta}{d\omega}(\omega - \omega_0)\right), \tag{5.67}$$

Eq. (5.66) is given as the sum of two Lorentzian terms:

$$\frac{t - O}{1 - 2tO + O^2} = \frac{A}{2}\left(\frac{1}{(\omega - \omega_+) + \frac{\Delta\omega_0}{2}} + \frac{1}{(\omega - \omega_-) + \frac{\Delta\omega_0}{2}}\right). \tag{5.68}$$

where $A = FSR/\xi$ and $\Delta\omega_0$ is the original 3 dB bandwidth of the ring without the reflector. The frequencies of the two peaks are given by

$$\omega_\pm = \omega_0 \pm \frac{FSR}{\xi}\sqrt{1 - t^2}. \tag{5.69}$$

The coupling coefficients (t and κ) of the directional couplers and the field reflectivity r of the lumped reflector were then subject to a least mean square fitting to the measured spectrum shown in Fig. 5.17a. The result of the curve fitting is shown in Fig. 5.17c, which exhibits an almost perfect fit to the measured spectrum. With this correction, the coupling coefficient is now estimated to be $\kappa \approx 0.12$, which is much closer to the predicted value of 0.1. Fig. 5.17d presents a close-up of the fitting result indicating the clear symmetric splitting of the resonance. The backscattering reflectively was extracted to be r = 0.0124. This is well aligned with the conclusion made by Little *et al.* [104] that when $r \geq \kappa^2$, a clear resonance splitting is observed.

5.10   Modeling the Optical Loss of the Ring

The second compact model relates the loss inflicted to signals transiting inside the MRR to its size. The loss of the optical mode inside the ring has three contributors: 1) Loss due to the material absorption and surface state absorption, 2) scattering, mainly due to sidewall roughness and 3) radiation loss due to the curvature of the ring. The first contributor applies to any waveguide and is independent of the MRR radius. The last contributor radiation loss, in contrast, is widely characterized by the waveguide bending (and for this reason called bending loss). Scattering loss, finally, is present in straight waveguides but can be exacerbated in bent ones as the optical mode is shifted closer to the waveguide boundaries. Radiation loss has been extensively studied from a theoretical perspective [85, 157]. Numerical FDTD models have also been used to express the bending loss as a function of the ring radius [26, 82, 95]. However, none of these approaches can directly
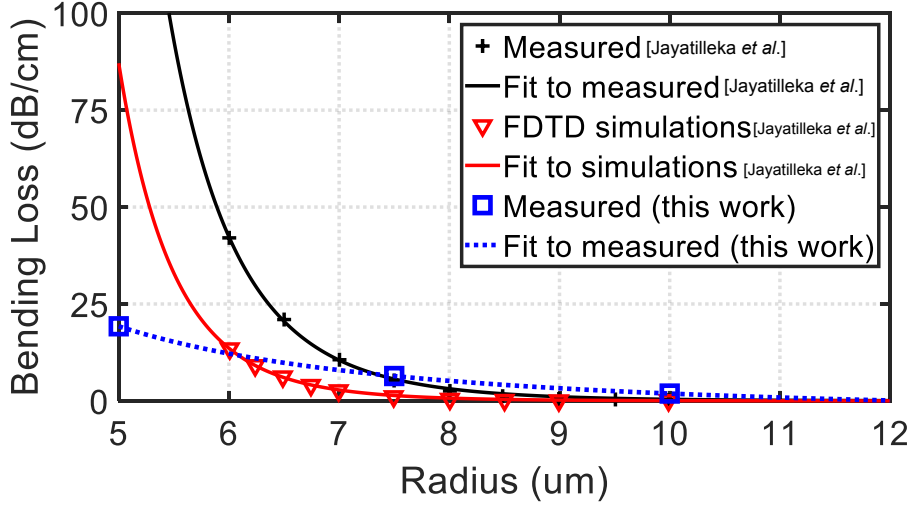
Figure 5.18: Bending loss of ring resonators as a function of radius based on the simulations and measurements presented in [82] and measurements presented in this work.

relate to fabrication defects, which play a major role in defining scattering losses.

Acknowledging that a method taking into account fabrication related imperfections is hard to construct, we consider the bending loss as a fabrication-platform dependent relationship, to be obtained experimentally and fitted. Such experimental measurements of the ring loss (compared to the loss of a straight waveguide) have been recently reported in [82]. We fitted these measurements with a power law

$$\alpha_{[dB/cm]} = a \left(R_{\mu m} \right)^{-b} + c \tag{5.70}$$

where R is the ring radius (in micron units) and a, b and c are constant parameters. Hence, we assume the bending loss to be infinite for null radius and a constant for infinite radius (i.e. rectilinear waveguide). As shown in Fig. 5.18, the fit agrees reasonably well with two sets of measurements, one provided by authors of [82] (with parameters $a = 4.5323 \times 10^8$ and $b = 9.0334$, $c \approx 0$), the other collected from our fabricated structures (with parameters $a = 2096.3$, $b = 2.9123$, and $c = 0$). The power law fit is also in good agreement with FDTD simulations realized in [82] (with parameters $a = 1.1452 \times 10^9$ and $b = 10.1848$, $c = 0$). As indicated in [82], an additional radius-independent propagation loss must be added

to the bending loss to include the effects of material absorption and standard sidewall roughness. Our baseline compact model for ring loss (black curve in Fig. 5.18) considers the aforementioned a and b values, plus a constant loss of $c = 2$ dB/cm [4]. Note that the rings studied in [82] are made of ridge waveguides with 90 nm of slab thickness while the rings in this work are made of strip waveguides.

5.11   Design Space for Add-Drop Filters

Add-Drop ring structures can serve as wavelength selective filters and demultiplexers for WDM-based optical links [14]. The design space for each individual ring can be constructed as shown in the plots of Fig. 5.19. The main assumptions that we apply are that the ring is made of 450×220 nm strip waveguides and is operating at critical coupling. This condition is satisfied when $t_{in}^2 = L \times t_{drp}^2$. For the loss of the ring, we use the power-law model that fits the measured ring losses in Fig. 5.18. Satisfaction of critical coupling condition requires that $\kappa_{drp} < \kappa_{in}$, and hence the drop gap (*output gap*) should be slightly bigger than the input gap. We take the output gap as the independent variable and find the input gap such that the critical coupling condition is held.

Since the design space depends on the loss of the ring, we consider two cases one of which is for the loss model fitted to measured data in [82] and the other one is for the loss model fitted to the data from our own measurements.

Figure 5.19 shows the exploration of the design space for the first case. Fig. 5.19a shows the contours of drop insertion loss at the resonance. By limiting the insertion loss of the drop path to better than 1 dB, the design space is divided such that a larger radius is accompanied by a larger gap size. Figure 5.19b describes the spectral attenuation of the drop path at FSR/2 detuning from the resonance. This is also describing the extinction of each resonance in the spectrum. We set the requirement for the extinction of resonance to be better than -30 dB to provide at least -20 dB crosstalk suppression in a WDM link
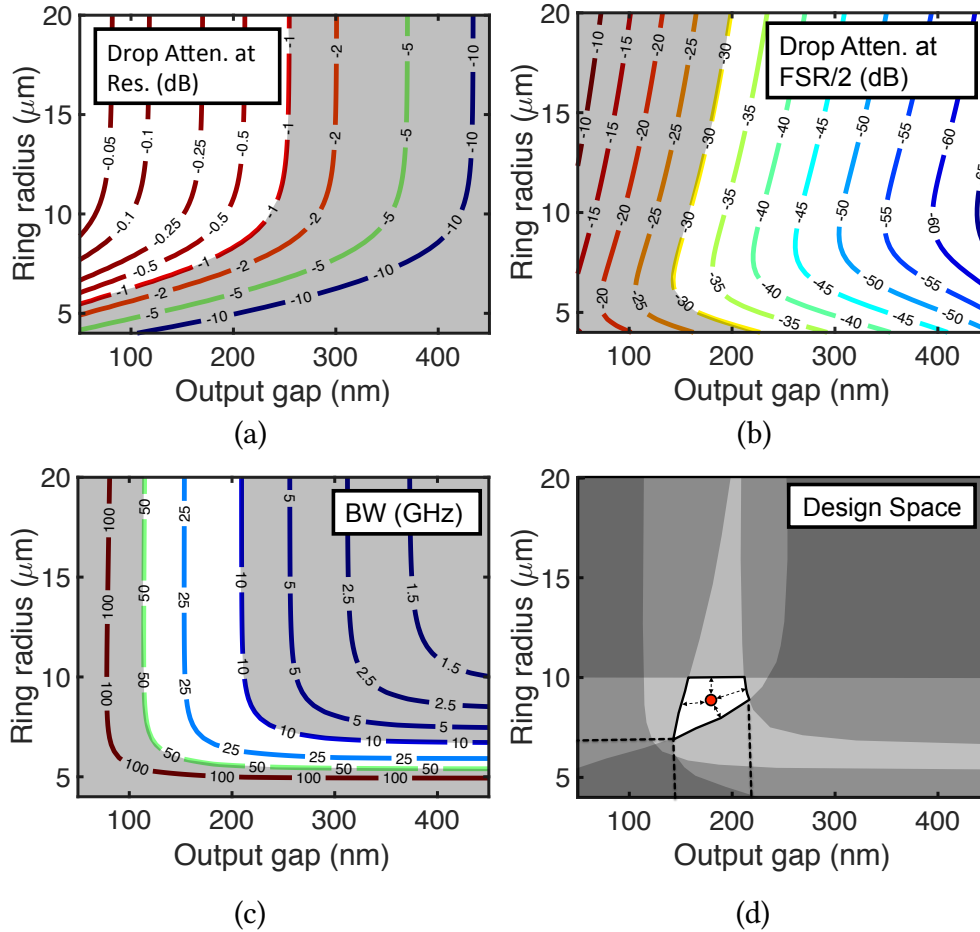
Figure 5.19: Characterization of design space for microring Add-drop filters based on strip waveguides and the baseline loss model [82]. (a) Contours of attenuation at the resonance. The white area corresponds to less than 1 dB attenuation. (b) Contours of attenuation at half FSR. White area corresponds to attenuation better than 30 dB. (c) Contours of 3 dB optical bandwidth. White area corresponds to a bandwidth greater than 10 GHz and less than 50 GHz. (d) Overall design space of add-drop ring filters.

[15]. Therefore, the left side of the design space is grayed out in Fig. 5.19b. Figure 5.19c shows the 3dB optical bandwidth of the drop filter. Considering a minimum signaling rate of 10Gbps-per-λ for a WDM-based link, we require that the bandwidth of the filter be greater than 10 GHz yet less than 50 GHz, and gray out the undesired regions. Finally, we demand that each add-drop ring filter should provide a minimum FSR of 10 nm to allow at least 10×10Gbps channels spaced 1 nm apart from each other. This sets the upper-bound of radius to about 10 μm. The combination of all these constraints results in a design space, shown in Fig. 5.19d, limiting the radius to 7 μm − 10 μm and the output
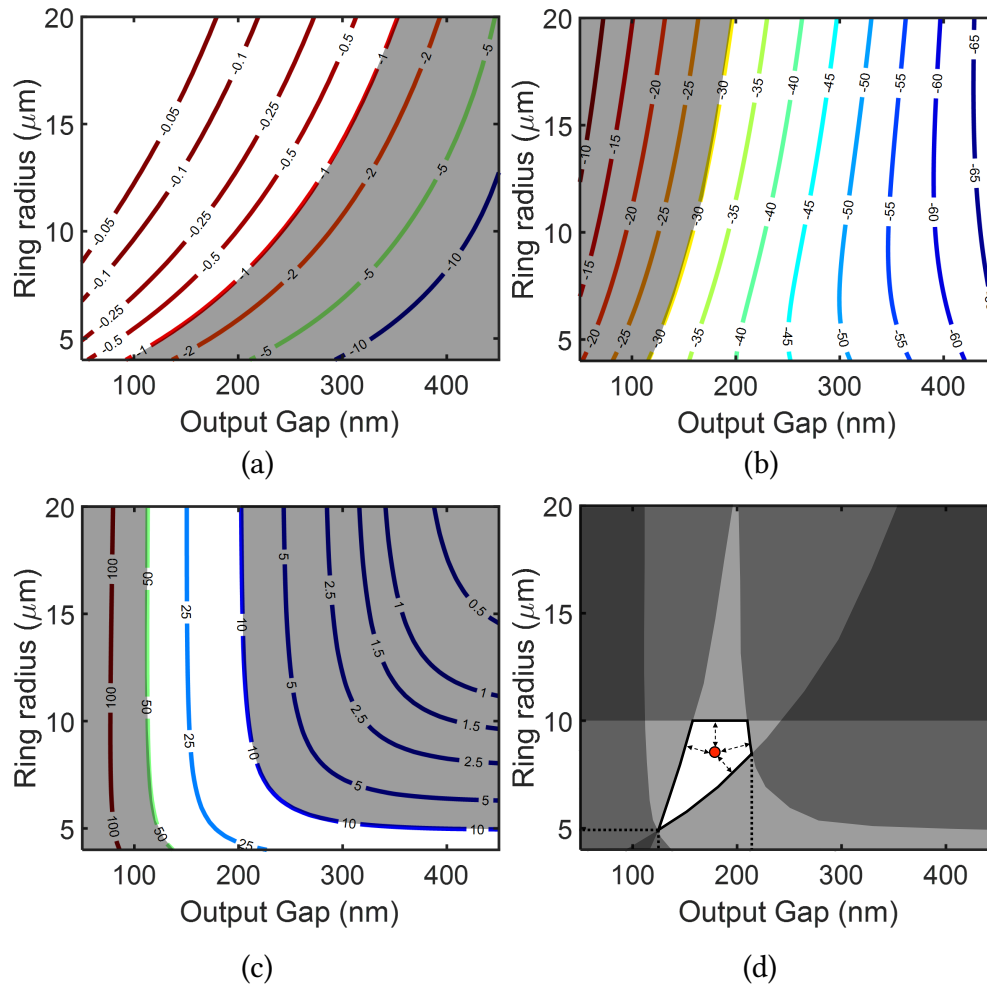
Figure 5.20: Characterization of design space for microring Add-drop filters based on strip waveguides and the loss model from our direct measurements. (a) Contours of drop attenuation at the resonance. The white area corresponds to less than 1 dB attenuation. (b) Contours of attenuation at half FSR. White area corresponds to attenuation better than 30 dB. (c) Contours of 3 dB optical bandwidth. White area corresponds to a bandwidth greater than 10 GHz and less than 50 GHz. (d) Overall design space of add-drop ring filters.

gap to 150 nm – 210 nm. The ideal design point is shown with a red circle on Fig. 5.19d. This point is at the center of the design space corresponding to a radius of about 9 μm and output gap of 180 nm. This choice of design point will to some extend be immune to variations on the radius and gap size since the variations on the width of the waveguides are typically within 5 nm [90].

Fig. 5.20 shows the exploration of the design space for the second case. Same conditions are applied to constrain the design space in Fig. 5.20a for the drop loss at resonance,

Fig. 5.20b for the extinction of resonance, and Fig. 5.20c for the 3 dB optical bandwidth of the ring. Fig. 5.20d shows the optimal design space for this case. Compared to the previous case, smaller radii (down to 5 μm) are supported in the design space. The optimal design point in Fig. 5.20d is characterized by 8.6 μm radius and an output gap of 178 nm. As can be seen, the optimal point in this case is very close to the previous case even though the loss models were quite different for small radii.

5.12   Other Types of Coupling

As we pointed out, Eq. (5.54) can be viewed as a general equation that can estimate the coupling between two identical waveguides with any arbitrary coupling region. The curvature and nonuniformity of the coupling region is entirely lumped inside the curvature function $B(x)$. Here we extend the analysis of the ring-waveguide coupling to provide the curvature function $B(x)$ for different types of coupling structures.

Two Straight Waveguides

For two identical straight waveguides of length L coupled to each other (Fig. 5.9), the gap is constant

$$g(z) = d \tag{5.71}$$

and the curvature function is given by

$$B(x) = x \tag{5.72}$$

where $x_{E,O} = \gamma_{E,O}L$.

Directional Coupler with S-bends

S-bends are usually used to bring two straight waveguides close to each other in the coupling region and then taking them away from each other (Fig. 5.9). S-bends are parametrically defined by a cosine function

$$Z(t) = H\,t\,,\ Y(t) = \frac{V}{2}(1 - cos(\pi t)) \tag{5.73}$$

where $0 < t < 1$ is the parameter of the curve, H is the horizontal offset of the S-bend and V is the vertical offset. In this case, the gap is given by

$$g(z) = \begin{cases} d & |z| \leq L/2 \\ d + V(1 - \cos(\pi \frac{|z|-L/2}{H})) & L/2 \leq |z| \leq L/2 + H \end{cases} \tag{5.74}$$

and the curvature function is calculated as

$$B(x) = x\left[1 + \frac{2H}{L}\exp(-Vx/L)I_0(Vx/L)\right] \tag{5.75}$$

where $x_{E,O} = \gamma_{E,O}L$. Here, $I_0(x)$ is the modified Bessel function of the first kind of order zero.

Ring-Ring Coupling

In the case of two identical ring resonators coupled to each other as shown in Fig. 5.9, the curvature function is given by

$$B_{\text{Ring-Ring}}(x) = 0.5\,B_{\text{Ring-wg}}(2x) \tag{5.76}$$

where $x_{E,O} = \gamma_{E,O}(R + w/2)$. It can be easily verified that the $B(x)$ function for ring-ring coupling is always smaller than the $B(x)$ function for the ring-waveguide coupling.
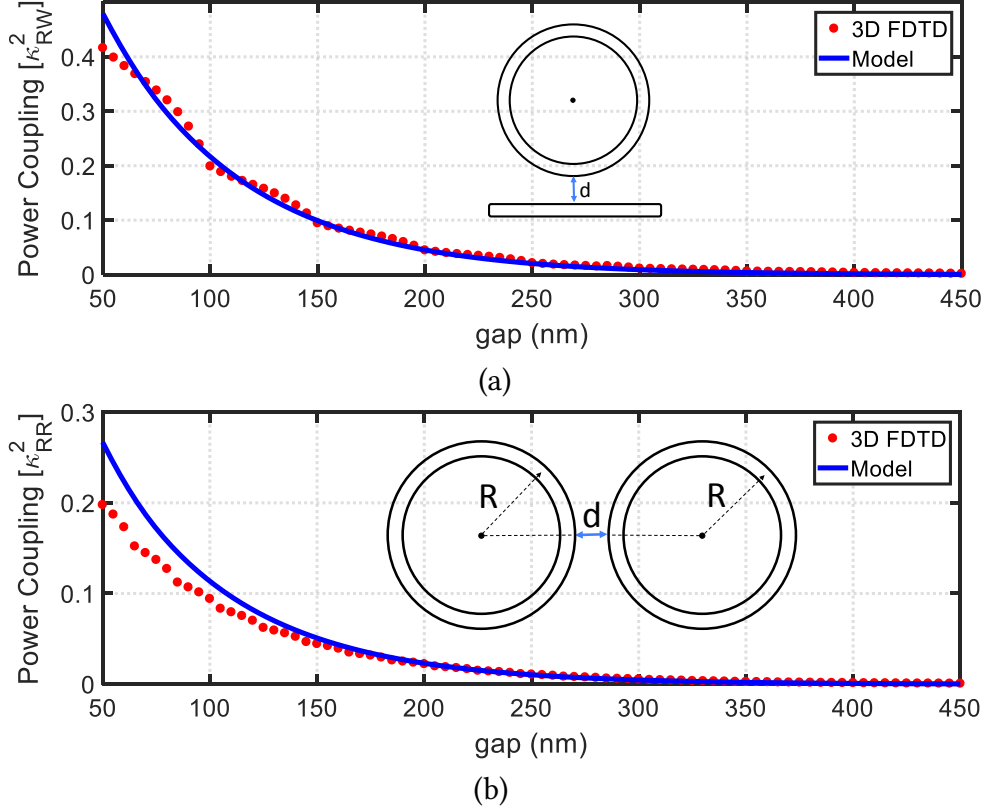
Figure 5.21: (a) Plot of the ring-waveguide power coupling (κ_{RW}^2) for 400×220 nm strip waveguides and 5μm radius. Solid blue curve is estimated from our model and circles are from a 3D FDTD simulation (RSoft). (b) Plot of the ring-ring power coupling coefficient (κ_{RR}^2) for 400×220 nm strip waveguides and 5μm radius. Solid blue curve is estimated from our model and circles are from a 3D FDTD simulation (RSoft). The ratio between case (a) and case (b) is close to 2 as predicted by our model.

More precisely, based on the approximation of $B_{\text{Ring-wg}}(x) \approx \sqrt{2\pi x}$, it can be seen that

$$\frac{B_{\text{Ring-Ring}}(x)}{B_{\text{Ring-wg}}(x)} \approx \frac{1}{\sqrt{2}} = 0.71 \tag{5.77}$$

for large values of x. Considering that $a_E, a_O, \gamma_E, \gamma_O$ in Eq. (5.54) do not depend on the curvature of the coupling region, Eq. (5.77) results in the following relation

$$\frac{\kappa_{\text{Ring-Ring}}}{\kappa_{\text{Ring-wg}}} \approx \frac{1}{\sqrt{2}} = 0.71 \tag{5.78}$$

To show this, we performed an FDTD simulation in RSoft software [225] for 400×220 nm strip waveguides and 5 μm radius of curvature. The calculated power couplings (κ^2) are plotted in Fig. 5.21. A good agreement is observed.

Race-track ring structure

For the case of a race-track ring structure as shown in Fig. 5.9, the gap function is given by

$$g(z) = \begin{cases} d & |z| \le L/2 \\ d + (R + w/2)(1 - \cos\theta) & L/2 \le |z| \end{cases} \tag{5.79}$$

and the curvature function is calculated as

$$B_{\text{Race-track}}(x) = \frac{L}{R + w/2} x + B_{\text{Ring-wg}}(x) \tag{5.80}$$

where $x_{E,O} = \gamma_{E,O}(R + w/2)$. Note that in general, the curvature function of the race-track structure can be written as a superposition of the curvature functions of waveguide-waveguide and ring-waveguide structures

$$B_{\text{Race-track}} = B_{\text{wg-wg}}(x_{\text{wg}}) + B_{\text{Ring-wg}}(x_{\text{ring}}) \tag{5.81}$$

where $x_{\text{wg}} = \gamma L$ and $x_{\text{ring}} = \gamma(R + w/2)$ for even and odd modes.

Ring resonator with circularly shaped bus waveguide

In this case as shown in Fig. 5.9, the bus waveguide has a circular part around the ring over an angle of θ_0 (shown in Fig. 5.22a). This makes the coupling region approximately a uniform region whose gap function is $g(z) = d$ and the curvature function is given by $B(x) \approx x$ where $x_{E,O} = \gamma_{E,O}(R + w/2 + d/2)$. Figure 5.22b shows a comparison between the coupling of a ring resonator to a straight waveguide and the coupling of a ring to a

(a)

(b)
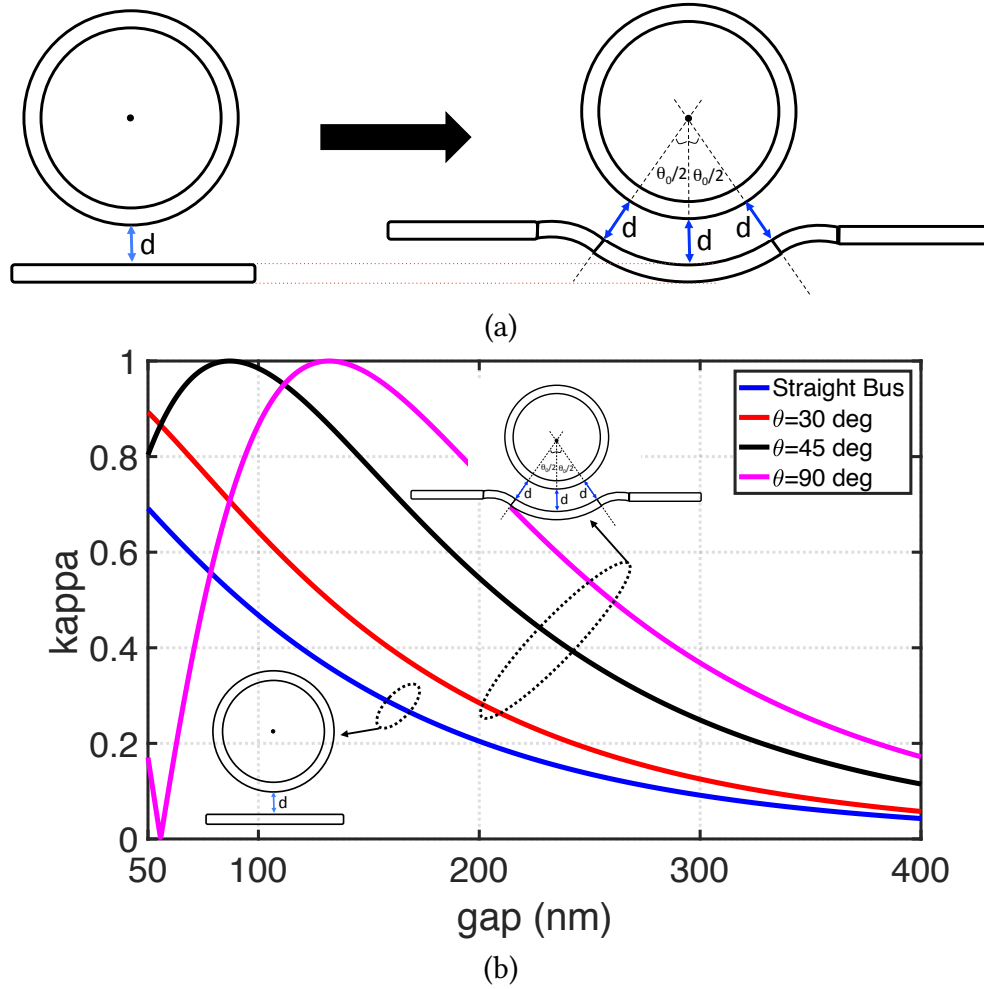
Figure 5.22: (a) Transforming the straight bus waveguide into a circular bus. (b) Comparison of ring-waveguide coupling between the case where the bus waveguide is straight and the case where the bus waveguide is circular with the coupling region angle set to 30 degree, 45 degree, and 90 degree. It can be seen that the circular bus provides stronger coupling under similar conditions (same radius and gap size).

circularly shaped waveguide (based on 400×220 nm strip waveguides and a radius of 5 μm).

5.13   Summary of the Chapter

We introduced compact models for coupling coefficients and bending loss of ring resonators in SOI platform. The model for coupling coefficients was first validated by full-

wave 3D FDTD simulations and then through direct measurements of spectral response of fabricated devices. The model for loss established a power-law relation between the loss of the ring and its radius. It was shown that this model can be reasonably well fitted to the measurements and simulated bending loss of ring resonators.

The proposed compact models were used to characterize a realistic design space for ring resonators that are performing as add-drop demultiplexers in WDM applications. It was concluded that the design space for add-drop ring filters results in a range of radii from 5 μm to 10 μm and gap sizes from 120 nm to 210 nm. The center point of the design space was chosen as the optimal design point whose radius is about 9 μm, output gap is about 180 nm, and is operating at critical coupling. This design will provide better than 30 dB of extinction for the resonance, a drop loss less than 0.5 dB and a resonance bandwidth of about 20 GHz, with an FSR greater than 10 nm.

Chapter 6

Thermo-Optic Efficiency and Thermal Rectification of Integrated Heaters

Abstract– *In this chapter, we characterize the thermo-optic efficiency and transient thermal response of integrated microheaters and develop models for silicon photonic devices equipped with such heating elements. To demonstrate the usefulness and practicality of the proposed models, we investigate pulse width modulation (PWM) drive scheme for microheaters and experimentally demonstrate its use in stabilizing microring resonators by digital driving signals (two-level voltage). Requirements on the drive frequency and duty cycle of the PWM signal are discussed based on the accurately modeled thermal frequency response of the heater-ring system.*

Thermal effects significantly change the optical behavior of silicon photonic devices due to the strong thermo-optic coefficient of silicon material (dn_{Si}/dT = $1.8{\times}10^{-4}$ K^{-1}) [87]. This high thermo-optic susceptibility, on one hand, can be an opportunity as it allows for effective manipulation of the light in thermo-optic modulators [126, 64, 196] and optical switches [182, 95, 107], for example. On the other hand, for very narrow-band optical devices such as ring resonator filters, thermal susceptibility is detrimental, and requires accurate monitoring and control of the temperature to obtain desired behavior and performance [65, 218, 112, 133, 136]. For example, the resonant frequency of a silicon ring resonator is shifted by $\sim$9 Ghz ($\sim$0.07 nm) for each degree Kelvin change in the temperature of the waveguide [118, 139]. A single degree of temperature variation is therefore sufficient to create significant spectral distortion of on-off keying (OOK) signals

for dense wavelength-division multiplexing (DWDM) systems that employ add-drop ring filters at the receiver [15, 17, 181].

Different mechanisms of thermal control and compensation have been proposed to improve tolerance of narrow bandwidth devices to ambient temperature variations. The use of polymer cladding that counteracts the thermo-optic effect of silicon has been proposed [69]; however, this approach requires a specific width of the waveguide (∼306 nm [133]) and some post fabrication processes that are not CMOS compatible. Polymers have also been used in the design of low-power thermo-optic switches [131]. The most popular solution leverages active thermal control and consists of sending an electrical current through external but closely integrated metallic [11] or doped-silicon based micro-heaters to induce ohmic heating [139]. Fabrication constraints require metallic heaters to be located above the waveguide and doped-silicon heaters to be located next to the waveguide. In both cases, the generated heat diffuses mainly through the silica (SiO_2) cladding layer to reach the silicon waveguide [118, 131]. Solutions where the waveguide itself is turned into a heater by tapering it to a wider width and doping it locally have also been proposed in [198, 49], and [199]. An approach consisting of a PN-doped waveguide driven with a reversed voltage to its breakdown to produce Ohmic heat right inside it has also been proposed by Li *et al.* [100]. Heat diffusion by means of external heaters, however, permits a separation between the design of the waveguide (choice of optical mode and polarization) and the design of the heating element.

Designs based on external heaters must take into account the fact that heat diffusion is not instantaneous – it takes time for the heat to propagate from the heater to the waveguide. Moreover, the heat might reach the waveguide at different strengths and different delays. The optical operation of the photonic device depends, therefore, on the frequencies of the electrical excitation signal applied to the heater. Understanding this transient thermal response is crucial to finely adjust and efficiently counter-balance the thermal noise that impacts the response of silicon photonic devices. Current discussion

of transient thermal response is reduced to experimental considerations, specifically a posteriori characterizations of rise and fall times (the time required for the temperature of the waveguide to evolve from 10% to 90% and 90% to 10% of its asymptotic values, respectively), while the stationary thermal response is straightforward to model accurately [197]. Observing the importance of thermal stabilization in the total power budget of a photonic link [18], we conclude that the understanding of thermal effects in silicon photonics platform deserves careful study.

Micro-heaters are generally incorporated into a feedback control loop that identifies the required heater driving conditions [66]. These control loops have been implemented with analog circuits and dithering signals in the work of Padmaraju *et al.* [134, 137] showing wide thermal ranges. Unfortunately, solutions based on analog signaling are susceptible to electrical noise as well as thermal noise and have limited flexibility, unlike digital control loops that provide a more robust and, if required, programmable signaling option [35]. Digital-to-analog converters with 16-bit or more inputs [66] can be used to convert the discrete digital output into a continuous voltage for the heater with sufficient accuracy. This approach, however, is not applicable in cases where a large number of heaters must be driven simultaneously [24]. A scalable solution investigated by Aguinaldo *et al.* in 2014 [2], Zecevic *et al.* in 2015 [210], and recently improved upon by Mansuura *et al.* [119], proposes digital pulse width modulation (PWM) signaling to drive the heaters, an approach that eliminates the need for DACs. In this design, the limited thermal bandwidth of the heater-waveguide system acts as a rectifying low-pass filter, resulting in the net effect of the PWM excitation signal resembling that of a DC voltage excitation. In this approach, frequency and duty cycle of the PWM excitation must be carefully examined.

6.1 Thermal Sensitivity of Silicon Waveguides

Temperature can affect the phase of optical mode inside the waveguide by changing the effective index of the mode. A small change in the effective index is related to the change in temperature through the following equation:

$$\Delta n_{\text{eff}} \approx \frac{dn_{\text{eff}}}{dT} \times \Delta T \tag{6.1}$$

where dn_{eff}/dT is the sensitivity of the optical mode to the temperature. Since the sensitivity of the optical mode to the waveguide parameters is already known, it is sufficient to know how to the temperature affects the parameters of the waveguide.

- **Linear Thermal Expansion**

Temperature affects the physical dimensions of the waveguide by linearly expanding them:

$$\text{Linear Thermal Expansion:} \quad \Delta L \approx \frac{\partial L}{\partial T} \times \Delta T. \tag{6.2}$$

The linear thermal expansion describes the relative expansion sensitivity of a material in vacuum to the temperature:

$$\frac{\Delta L}{L} \approx \alpha \times \Delta T \Rightarrow L \approx L_0(1 + \alpha \Delta T). \tag{6.3}$$

In general, since the silicon core of a waveguide is surrounded by silica, the expansion of the width and height of the core is a function of linear thermal expansion of both silicon and silica. These values in room temperature (27°C, 300°K) are given by

$$\text{Silicon:} \quad \frac{dL_{si}/L}{dT} = \alpha_{si} \approx 2.6 \times 10^{-6} \ K^{-1} \tag{6.4}$$

$$\text{Fused Silica:} \quad \frac{dL_{sio_2}/L}{dT} = \alpha_{sio_2} \approx 5.6 \times 10^{-7} \ K^{-1}. \tag{6.5}$$

Since the thermal expansion of silicon is stronger than silica, we assume that the thermal expansion of the core of the waveguide is only due to the expansion of silicon core. Therefore:

$$\frac{\Delta w}{w} \approx \alpha_{si}\Delta T \quad , \quad \frac{\Delta h}{h} \approx \alpha_{si}\Delta T \tag{6.6}$$

• **Change of Material Index**

The index of silicon and silica can change due to the change in temperature. The thermo-optic coefficients of silicon and silica at 1550 nm wavelength and room temperature are given by [87]

$$\text{Silicon:} \quad \frac{dn_{si}}{dT} \approx 1.8 \times 10^{-4} \ K^{-1} \tag{6.7}$$

$$\text{Silica:} \quad \frac{dn_{sio_2}}{dT} \approx 8.66 \times 10^{-6} \ K^{-1} \tag{6.8}$$

which indicates that the change in the index of silicon is two orders of magnitude stronger than then change in the index of silica. This is expected because silicon is a semiconductor material while silica is an insulator.

• **Change of Effective index**

The thermo-optic effect of the mode inside the waveguide (consider a strip waveguide) can be linearly related to the change in temperature using the thermal sensitivity coefficient dn_{eff}/dT where

$$\frac{dn_{\text{eff}}}{dT} = \frac{\partial n_{\text{eff}}}{\partial n_{si}}\frac{dn_{si}}{dT} + \frac{\partial n_{\text{eff}}}{\partial n_{sio_2}}\frac{dn_{sio_2}}{dT} + \frac{\partial n_{\text{eff}}}{\partial w}\frac{dw}{dT} + \frac{\partial n_{\text{eff}}}{\partial h}\frac{dh}{dT} \tag{6.9}$$

and

$$\frac{dw}{dT} \approx \alpha_{si} \times w, \tag{6.10}$$

$$\frac{dh}{dT} \approx \alpha_{si} \times h. \tag{6.11}$$
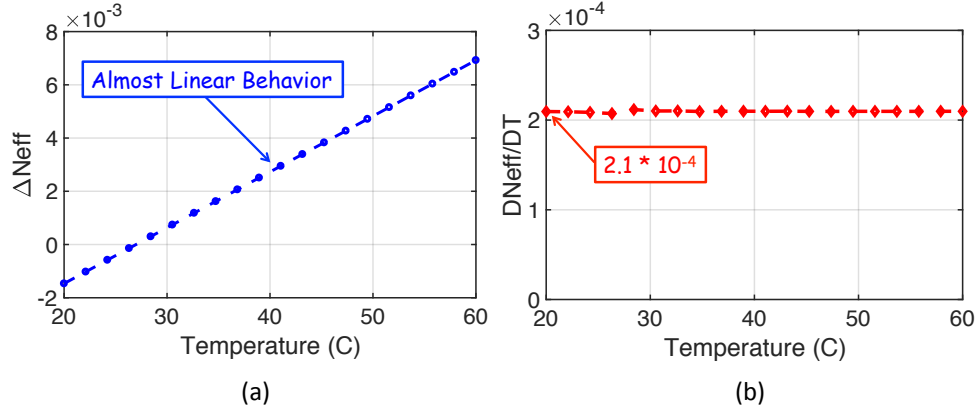
Figure 6.1: Thermo-Optic analysis of quasi-TE$_{00}$ mode of the strip waveguide. (a) Δn_{eff} as a function of temperature. The original waveguide is assumed to be at room temperature. (b) Thermo-optic sensitivity of the strip waveguide as a function of temperature.

For the standard strip waveguide with dimensions 450 nm$\times$220 nm, the thermo-optic sensitivity of the fundamental quasi-TE mode (TE$_{00}$) can be calculated as

$$
\begin{aligned}
\frac{dn_{\text{eff}}}{dT} &= (1.127)(1.8 \times 10^{-4}) \\
&+ (0.2269)(8.66 \times 10^{-6}) + (0.0022)(2.6 \times 10^{-6} \times 450) \\
&+ (0.0039)(2.6 \times 10^{-6} \times 220) = 2.1 \times 10^{-4} \ K^{-1} \quad (6.12)
\end{aligned}
$$

Therefore

$$
\text{quasi-TE}_{00} : \quad \frac{dn_{\text{eff}}}{dT} \approx 2.1 \times 10^{-4} \ K^{-1}. \quad (6.13)
$$

Next, we change the temperature of a 450 nm$\times$220 nm strip waveguide and find the effective index of the mode as shown in Fig. 6.1a. The slope of the Δn_{eff} versus ΔT is the thermo-optic sensitivity of the optical mode. Figure 6.1b shows that the thermo-optic sensitivity of the mode is almost constant over a wide range of temperature.
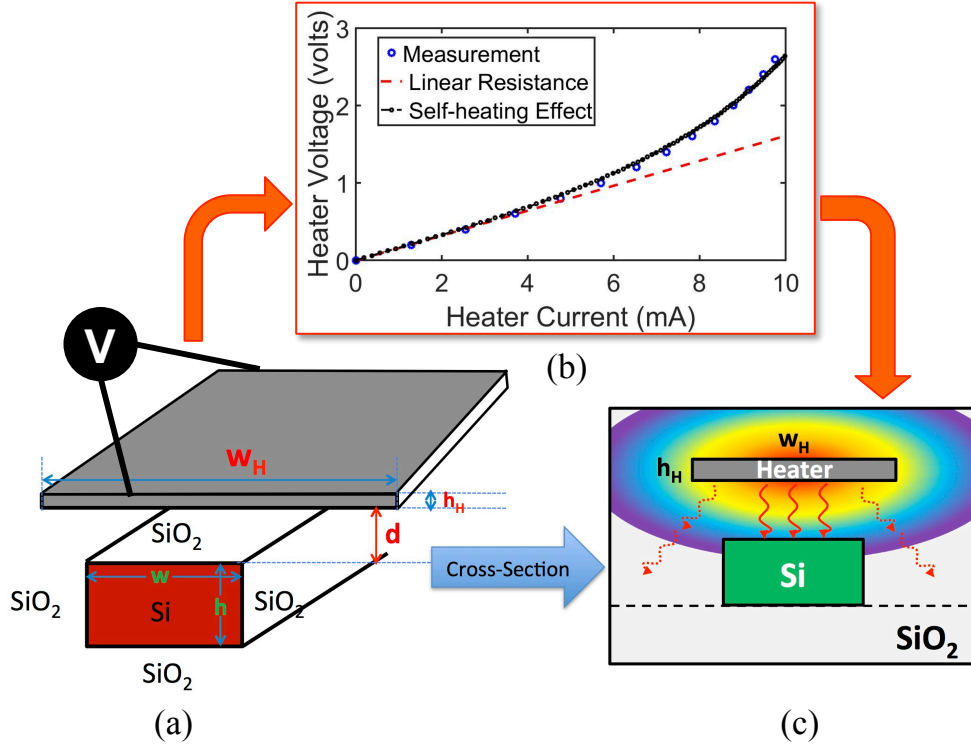
Figure 6.2: (a) Structure of strip waveguide with a metallic heater situated at a distance d above it. (b) V-I characteristic of the resistive heater. The nonlinear behavior is observed. (c) Schematic of heat generation at the heater and its distribution in the cladding. Dimensions are not necessarily to scale.

6.2 Heater-Waveguide System

The schematic of a strip waveguide with a micro-heater situated at a distance d above it is shown in Fig. 6.2a. By applying a voltage to the heater, $V(t)$, and running electrical current, $I(t)$, through it ohmic power $P(t) = V(t) \times I(t)$ is dissipated by the heater, and results in a rise of the heater temperature (ΔT_H).

Heater Thermal Response

We are interested in expressing the temperature rise ΔT_H as a function of the applied voltage $V(t)$. We first note that would the resistance of the heater R stay constant, the power dissipated would be strictly a quadratic function of the voltage as $I(t) = V(t)/R$ and thus $P(t) = V^2(t)/R$. However, as the heater temperature increases, the resistivity

of the heater increases according to $R = R_{\text{linear}}(1 + \alpha_H \Delta T_H)$ and therefore

$$\frac{V}{I} = R(V) \approx R_{linear}(1 + \alpha_H \Delta T_H(V)) \tag{6.14}$$

where α_H is the temperature coefficient of the resistance of the heater, which is usually on the order of 10^{-3}. The value of α_H for some metals are given by

$$
\begin{aligned}
\text{Nickel:} \quad & \alpha_H = 5.866 \times 10^{-3} \ K^{-1} \\
\text{Iron:} \quad & \alpha_H = 5.671 \times 10^{-3} \ K^{-1} \\
\text{Tungsten:} \quad & \alpha_H = 4.403 \times 10^{-3} \ K^{-1} \\
\text{Aluminum:} \quad & \alpha_H = 4.308 \times 10^{-3} \ K^{-1} \\
\text{Copper:} \quad & \alpha_H = 4.041 \times 10^{-3} \ K^{-1} \\
\text{Silver:} \quad & \alpha_H = 3.819 \times 10^{-3} \ K^{-1} \\
\text{Platinum:} \quad & \alpha_H = 3.729 \times 10^{-3} \ K^{-1} \\
\text{Gold:} \quad & \alpha_H = 3.715 \times 10^{-3} \ K^{-1} \\
\text{Zinc:} \quad & \alpha_H = 3.847 \times 10^{-3} \ K^{-1}
\end{aligned}
$$

Here, R_{linear} is the resistivity of the heater at low voltage and $R(V)$ is the resistivity after the steady-state for self-heating is reached. The increased resistivity leads in turn to a decrease of the current and dissipated ohmic power. The self-heating situation of the heater finally regulates itself at a steady sate of temperature and supplied current. We assume that the self-heating of heater is almost instantaneous, i.e. the resistivity of the heater instantaneously changes with a change in the applied voltage. In that case, the rate of generating ohmic energy is equal to the rate with which heat is escaping from the boundaries of the heater. Therefore

$$P_H(t) \approx - \oint k_{sio_2} \, \nabla T \cdot \hat{n} \, dS \tag{6.15}$$

where k_{sio_2} = 1.38 W/(m.K) is the thermal conductivity of SiO$_2$. Schafft [162] and Hida *et al.* [75] have shown that solving the Laplace equation for the temperature distribution in SiO$_2$ at the steady-state results in a linear relation between ohmic power and the temperature at the heater surface $P_H(t) = \chi_{eq} \Delta T_H(t)$. In the case of metallic heater shown in Fig. 6.2a the χ_{eq} factor is given by

$$\chi_{eq} \approx L_H \, W_H \, k_{sio_2} \left(\frac{1}{Y_0} + \frac{0.88}{W_H} \right) \tag{6.16}$$

where L_H is the length of heater, and Y_0 is the distance from heater to the Si substrate. Combining this with the thermal dependence of the resistivity, $R = R_{linear}(1 + \alpha_H \, \Delta T_H(t))$, results in the following equation for the temperature of the heater:

$$\Delta T_H(t) = \frac{1}{\chi_{eq} \, R_{linear}(1 + \alpha_H \, \Delta T_H(t))} V^2(t). \tag{6.17}$$

The temperature of heater then follows the voltage of heater according to

$$\Delta T_H = \frac{1}{2\alpha_H} \left(-1 + \sqrt{1 + K_v V^2(t)} \right), \tag{6.18}$$

where K_v represents the thermal nonlinearity of the ohmic resistance:

$$K_v = \frac{4\, \alpha_H}{R_{linear} \, \chi_{eq}}. \tag{6.19}$$

The result of such self-heating phenomenon is a nonlinear DC voltage-current characteristic of the resistive heater given by

$$I(t) = \frac{V(t)}{R_{linear}} \times \frac{2}{1 + \sqrt{1 + K_v V^2(t)}}. \tag{6.20}$$

Note that if $\alpha_H \to 0$ then $K_v \to 0$, and Eq. (6.18) and (6.20) reduce to the simple forms for a linear resistance. Fig. 6.2b shows an example of experimentally measured V-I data

points (blue circles) along with an expected linear V-I curve (dashed red line). By fitting Eq. (6.20) to the measured data points, we find out that with parameters R_{linear} = 161 Ω and K_v = 0.6 V^{-2}, the model matches the measurements quite well. Fig. 6.2c schematically depicts the heat distribution after the self-heating of heater is reached.

6.3   Waveguide Thermal Impulse Response

Once we know how the heater temperature is affected by the voltage applied to it, we analyze how this temperature change propagates to the waveguide. We first note that performing a full analysis of the heater-waveguide system generally requires a 3D numerical solution of the heat transfer equation in the cladding [11]:

$$\frac{\partial \Delta T}{\partial t} = \xi_{\text{SiO2}} \left(\frac{\partial^2 \Delta T}{\partial x^2} + \frac{\partial^2 \Delta T}{\partial y^2} + \frac{\partial^2 \Delta T}{\partial z^2} \right) \tag{6.21}$$

where ξ_{SiO2} = 8.7$\times$10^{-7} m^2/s is the thermal diffusivity of silica. However, by assuming that *i*) the structure is uniform in the z-direction (cross-section is in the xy-plane) and, *ii*) the width of the heater is large enough compared to the width of waveguide ($w_H \gg w_{wg}$) and/or the width of the heater is large enough compared to the distance from waveguide ($w_H \gg d$), we can reduce this equation to only one direction (assumed to be y-direction). If the heater is situated at y = 0 and waveguide is situated at y = $-d$ as shown in Fig. 6.3, the equation of a 1D Fourier analysis of heat conduction in the region between heater is given by

$$\frac{\partial \Delta T}{\partial t} = \xi_{sio_2} \frac{\partial^2 \Delta T}{\partial y^2}. \tag{6.22}$$

In order to solve this problem, we first apply a Fourier transform in the time domain

$$\Delta T(y,t) = \int_{-\infty}^{\infty} \tilde{\Delta T}(y,\omega) e^{j\omega t} d\omega / 2\pi \rightarrow \frac{\partial^2 \tilde{\Delta T}}{\partial y^2} = j\omega / \xi_{sio_2} \tilde{\Delta T} \tag{6.23}$$
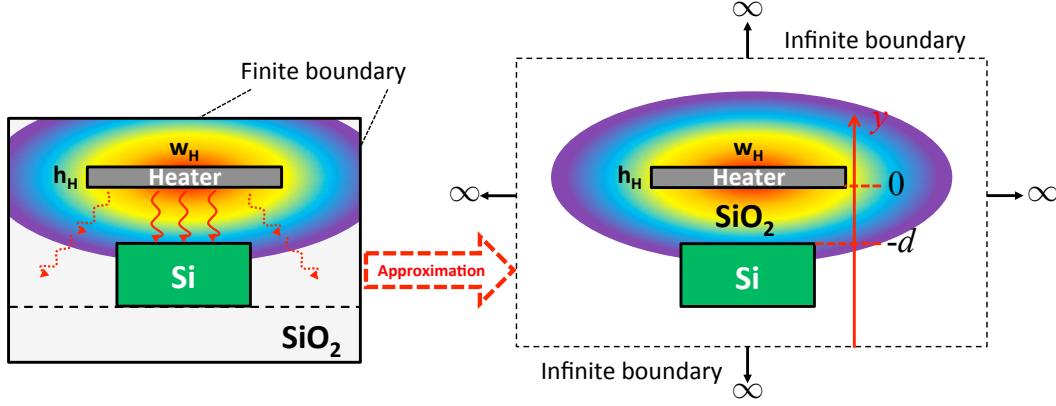
Figure 6.3: Approximation of extending the boundaries to the infinity and considering only 1D heat propagation in the y direction.

whose eigen values can be easily found:

$$k(\omega) = \pm\sqrt{j\omega/\xi_{sio_2}} = \pm\frac{1+j}{\sqrt{2}}\sqrt{\frac{\omega}{\xi_{sio_2}}} \qquad (6.24)$$

and since we are considering propagation of heat in the $-y$ direction, only one of the signs is acceptable. Therefore

$$k(\omega) = \begin{cases} k_+(\omega) & \text{for} \quad \omega > 0 \\[2mm] k_-(\omega) & \text{for} \quad \omega < 0 \end{cases} \qquad (6.25)$$

where

$$\text{if} \quad \omega > 0 \quad : \quad k_+(\omega) = \frac{1+j}{\sqrt{2}}\sqrt{\frac{|\omega|}{\xi_{sio_2}}} \qquad (6.26)$$

$$\text{if} \quad \omega < 0 \quad : \quad k_-(\omega) = \frac{1-j}{\sqrt{2}}\sqrt{\frac{|\omega|}{\xi_{sio_2}}} \qquad (6.27)$$

It can be seen that $k_-(\omega) = k_+^*(\omega)$ or in other words $k(-\omega) = k^*(\omega)$. The solution in the frequency domain is then given by

$$\tilde{\Delta T}(y, \omega) = C(\omega) \times e^{k(\omega)y} \qquad (6.28)$$

in which $C(\omega)$ can be determined by the initial conditions. The solution in the time domain is

$$\Delta T(y, t) = \int_{-\infty}^{\infty} C(\omega)e^{k(\omega)y}e^{j\omega t}d\omega/2\pi \tag{6.29}$$

which can be written as the convolution of two functions as

$$\Delta T(y, t) = C(t) \circledast I(y, t) \tag{6.30}$$

where

$$C(t) = \int_{-\infty}^{\infty} C(\omega)e^{j\omega t}d\omega/2\pi \tag{6.31}$$

$$I(y, t) = \int_{-\infty}^{\infty} e^{k(\omega)y}e^{j\omega t}d\omega/2\pi. \tag{6.32}$$

In order to find $C(t)$, we apply the boundary condition at $y = 0$ which corresponds to the change of temperature of the heater as a function of time $\Delta T(0, t) = \Delta T_H(t)$. On the other hand, we can see that $I(0, t) = \delta(t)$. Therefore

$$I(y = 0, t) = \delta(t) \tag{6.33}$$

$$C(t) = \Delta T_H(t) \tag{6.34}$$

and

$$\text{for} \quad y < 0 \ : \ \Delta T(y, t) = \Delta T_H(t) \circledast I(y, t). \tag{6.35}$$

The function $I(y, t)$ can thus be interpreted as the impulse response of the propagation of the heat from the heater to the surrounding. The change of temperature of the silicon waveguide is calculated by substituting $y = -d$:

$$\Delta T(-d, t) = \Delta T_H(t) \circledast I(-d, t), \tag{6.36}$$

and hence

$$\Delta T_{wg}(t) = \Delta T_H(t) \circledast I_{wg}(t), \tag{6.37}$$

where we call $I_{wg}(t)$ the *thermal impulse response* of the waveguide given by

$$I_{wg}(t) = 2 \int_0^\infty \exp\left(-\sqrt{\frac{f}{f_0}}\right) \cos\left(2\pi f t - \sqrt{\frac{f}{f_0}}\right) df \tag{6.38}$$

which has a closed-form solution for $t > 0$ as

$$I_{wg}(t) = \frac{1}{\sqrt{4\pi^2 f_0 t^3}} \exp\left(-\frac{1}{4\pi f_0 t}\right), \tag{6.39}$$

and is zero for $t < 0$. To show this, we consider the initial condition problem for the heat equation given by

$$\begin{cases} (\partial_t - \xi \partial_{yy})u(y,t) = 0 \\ u(y=0,t) = \delta(t) : \text{boundary condition} \\ u(y,t=0) = 0 \end{cases} \tag{6.40}$$

The goal here is to use Laplace transform to calculate the solution. If $u(y,t)$ is the solution, then the Laplace transform is defined as $U(y,s)$ such that

$$U(y,s) = \int_0^\infty u(y,t)\, e^{-st} dt \tag{6.41}$$

and it can be easily shown that

$$L\left\{\frac{\partial}{\partial t}u(y,t)\right\} = s\, U(y,s) - u(y,0). \tag{6.42}$$

Now we apply the Laplace transform to the heat equation and keep one of the solutions for the $y > 0$ region:

$$(s - \xi \partial_{yy} U(y, s)) = 0 \rightarrow U(y, s) = U(0, s) e^{-\sqrt{\frac{s}{\xi}} y} \tag{6.43}$$

Clearly $U(0, s)$ is the Laplace transform of $u(0, t)$, hence $U(0, s) = 1$. Finally, we need to perform the inverse Laplace transform:

$$u(y, t) = L^{-1}\{U(y, s)\} = \frac{1}{2\pi j} \int_{\gamma - j\infty}^{\gamma + j\infty} U(y, s) e^{st} ds \tag{6.44}$$

where γ is chosen such that the integral contour is on the right side of all the singularities of $U(y, s)$. In this particular case, $U(y, s)$ does not have any singularities, therefore we set $\gamma = 0$. The integral can then be evaluated as

$$u(y, t) = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} e^{-y\sqrt{s/\xi}} e^{st} ds \tag{6.45}$$

and a chagne of variable $z = \sqrt{t}\, s$ results in

$$u(y, t) = e^{\left(-\frac{y^2}{4\xi t}\right)} \frac{1}{j\pi t} \int_{-j\infty}^{j\infty} e^{(z - y/2\sqrt{\xi t})^2} z \, dz. \tag{6.46}$$

Note that the domain of integral can be remapped to $-j\infty \rightarrow j\infty$ after the change of variable $z = \sqrt{t}\, s$ because the integrand is zero at any point $x_0 \pm j\infty$. Finally, the change of variable $w = z - y/2\sqrt{\xi t}$ and then the change of variable $w = j\nu$ will yield the following result for $t > 0$:

$$u(y, t) = \frac{y}{\sqrt{4\pi \xi t^3}} \exp\left(-\frac{y^2}{4\xi t}\right). \tag{6.47}$$

By setting $y = -d$ and defining $f_0 = \xi_{\text{SiO2}}/(\pi d^2)$, function $I_{wg}(t)$ is obtained. We call f_0 as the *thermal frequency* of the system (note that the units of f_0 is 1/sec or Hz). For example,
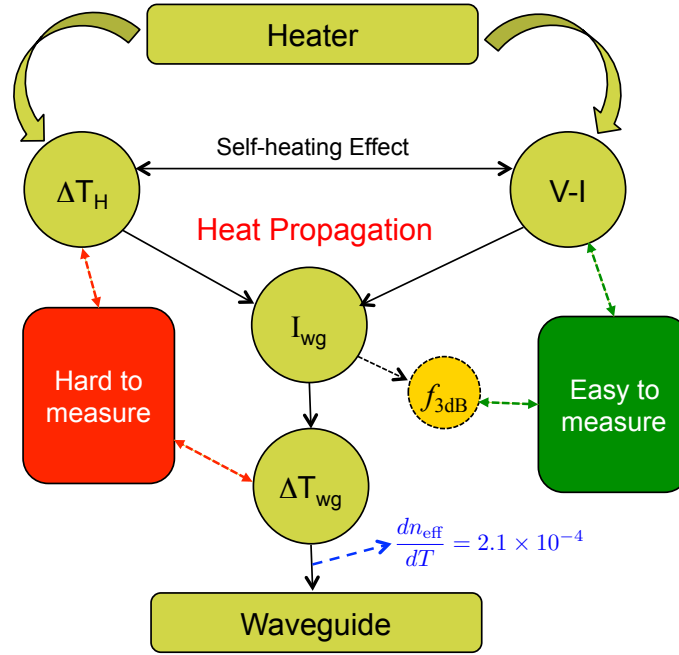
Figure 6.4: Block diagram of related parameters of the heater and the waveguide in order to estimate the transient thermo-optic response of the waveguide. Through V-I characteristic of heater $\Delta T_H(t)$ is estimated and through the thermal frequency response, $\Delta T_{wg}(t)$ is calculated.

if d = 1 μm and ξ_{SiO2} = 8.7$\times$10^{-7} m^2/s, then f_0 is about 275 kHz. Fig. 6.4 shows a plot diagram of the inter-relations and the procedure to estimate the transient thermal behavior of heater-waveguide system. Although direct measurement of the temperature is hard, Eq. (6.18) combined with Eq. (6.39) provides an approximation of how the temperature is changing at the heater and at the waveguide. Fig. 6.5a plots the impulse response calculated from Eq. (6.39) for f_0 = 30 kHz. Based on Eq. (6.39) one can see that the peak of the thermal impulse response occurs at t_{peak} = $1/(6\pi f_0)$. Fig. 6.5b plots the thermal frequency response of the system for f_0 = 30 kHz based on Eq. (6.38). As can be seen, the behavior of the system resembles that of a lowpass filter whose 3dB bandwidth is approximately 15 kHz. In cases where the assumption $w_H \gg d$ causes a significant error on the thermal impulse response, a correction term can be added to Eq. (6.39) [23].
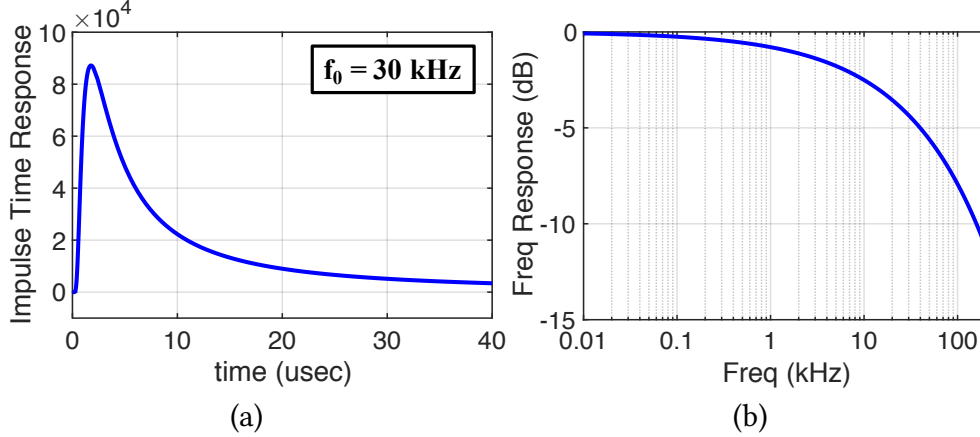
Figure 6.5: (a) Plot of impulse response function for f_0 = 30 kHz. (b) Plot of thermal frequency response of the heater-waveguide system for f_0 = 30 kHz.

Effect of a Thermal Wall on the Thermal Impulse Response

Even though Eq. (6.37) and Eq. (6.39) enable us to predict the temperature of a waveguide surrounded by a heater to which a prescribed voltage is applied, for the complete analysis, the model needs to take into account thermal obstacles at the immediate waveguide proximity. We thus now consider the case in which there is a thermal wall in the direction of heat diffusion such that the temperature of the wall is almost constant at all times, and thus perfectly absorbs and pipes out the incurring heat. For example, in the case of the metallic heater as shown in Fig. 6.6, the boundary between the buried oxide (BOX) layer and silicon substrate can typically be considered as a thermal wall [29, 162] due to the fact that the thermal diffusivity of silicon (substrate) is two orders of magnitude stronger than silica (BOX layer). Since temperature is continuous across the boundary, the change of temperature of the substrate at the boundary should be very small. In order to incorporate the effect of such hit sink into the equations, we start with the heat transfer equation and take the Fourier transform:

$$\Delta T(y,t) = \int_{-\infty}^{\infty} \tilde{\Delta T}(y,\omega) e^{j\omega t} d\omega/2\pi \rightarrow \frac{\partial^2 \tilde{\Delta T}}{\partial y^2} = j\omega/\xi_{sio_2} \tilde{\Delta T} \qquad (6.48)$$
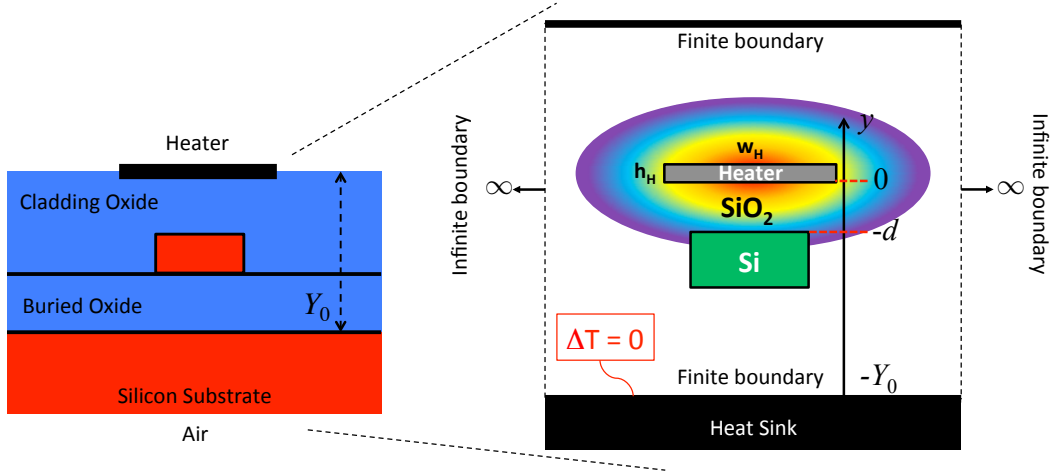
Figure 6.6: (left) Cross-section of the SOI waveguide. (right) Approximate 1D model for the heat-transfer in the $-\hat{y}$ direction. The heat-sink denotes the boundary at which the temperature is not changed at all times.

and the solution in the frequency domain now consists of both forward and backward terms:

$$\tilde{\Delta T}(y, \omega) = C(\omega)e^{k(\omega)y} + D(\omega)e^{-k(\omega)y} \tag{6.49}$$

which should satisfy the boundary condition at $y = -Y_0$:

$$\tilde{\Delta T}(y = -Y_0, \omega) = 0 \rightarrow \tilde{\Delta T}(y, \omega) = C(\omega)e^{k(\omega)y}\left(1 - e^{-2k(\omega)(y+Y_0)}\right) \tag{6.50}$$

and taking the inverse Fourier transform will yield

$$\Delta T(y, t) = \int_{-\infty}^{\infty} C(\omega)e^{k(\omega)y}\left(1 - e^{-2k(\omega)(y+Y_0)}\right)e^{j\omega t}d\omega/2\pi \tag{6.51}$$

which can be written in terms of time-domain convolution:

$$\Delta T(y, t) = C(t) \circledast I(y, t) \tag{6.52}$$

181

where $I(y,t)$ is the thermal impulse response of the system. At $y = 0$, the boundary condition is given by

$$\Delta T(0,t) = \Delta T_H(t) \rightarrow \tilde{\Delta T}_H(\omega) = C(\omega)\left(1 - e^{-2k(\omega)Y_0}\right) \tag{6.53}$$

which finally results in the following expression:

$$\Delta T_{wg}(t) = \int_{-\infty}^{\infty} \tilde{\Delta T}_H(\omega)e^{-k(\omega)d}\frac{1 - e^{-2k(\omega)(Y_0 - d)}}{1 - e^{-2k(\omega)Y_0}}e^{j\omega t}d\omega/2\pi. \tag{6.54}$$

This integral can be expressed in the time domain as the following convolution

$$\Delta T_{wg}(t) = \Delta T_H(t) \circledast I_{wg}(t) \tag{6.55}$$

where the impulse response $I_{wg}(t)$ is now given by:

$$I_{wg}(t) = \int_{-\infty}^{\infty} e^{-k(\omega)d}\left(\frac{1 - e^{-2k(\omega)(Y_0 - d)}}{1 - e^{-2k(\omega)Y_0}}\right)e^{j\omega t}d\omega/2\pi \tag{6.56}$$

or equivalently, the frequency response of the waveguide is given by:

$$\mathbf{1}_{\mathbf{wg}}(f) = \exp(-k(f))\left[\frac{1 - \exp(-2(\nu - 1)\,k(f))}{1 - \exp(-2\nu\,k(f))}\right] \tag{6.57}$$

where $\nu = Y_0/d > 1$, and

$$k(f) = (1 + j\,\mathrm{sign}(f))\sqrt{\frac{|f|}{f_0}} \tag{6.58}$$

and f_0 is defined as before ($f_0 = \xi_{\mathrm{SiO2}}/(\pi d^2)$). Parameter ν represents the ratio of the distance of heater to the thermal wall and the waveguide. The 3dB thermal bandwidth is approximated as

$$f_{\mathrm{3dB}} \approx f_0\left[\ln\left(\frac{2\nu}{\nu - 1}\right)\right]^2 \tag{6.59}$$
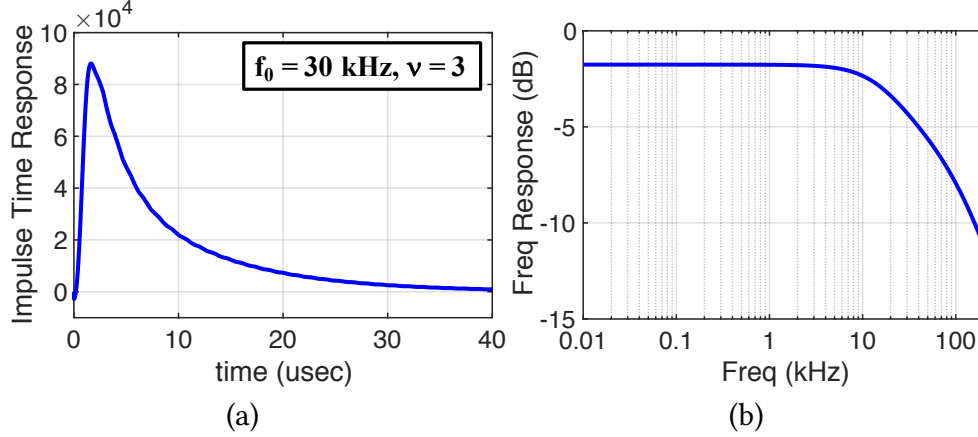
Figure 6.7: (a) Plot of impulse response in the presence of a thermal wall for f_0 = 30 kHz and ν = 3. (b) Plot of thermal frequency response in the presence of a thermal wall.

and the low frequency response ($f \to 0$) is given by $(\nu-1)/\nu$. It can be seen that setting ν = ∞ will result in Eq. (6.38) and Eq. (6.39). As a comparison, the thermal impulse response and frequency response based on Eq. (6.57) for f_0 = 30 kHz and ν = 3 are calculated and plotted in Fig. 6.7a and 6.7b, respectively. Eq. (6.57) exhibits a flatter low frequency response compared to Eq. (6.38).

Since the parameter ν is proportional to the distance from the thermal wall, increasing ν is equivalent to increasing the thickness of BOX layer in Fig. 6.6, which results in a decrease of thermal bandwidth according to Eq. (6.59) and an increase in the rise/fall time of the step response. This is in agreement with the numerical results presented in [11].

The 3dB thermal bandwidth of the heater-waveguide system depends on f_0 and ν parameters, therefore, knowing only $f_{3\text{dB}}$ will not uniquely determine f_0 and ν. For a given value of $f_{3\text{dB}}$, reducing ν on the $f_{3\text{dB}}$ contour given by Eq. (6.59) will reduce f_0 and make the frequency response sharper so that higher frequencies get more suppressed. In order to uniquely determine f_0 and ν, we can also define another frequency bandwidth such as 7dB given by

$$f_{7\text{dB}} \approx f_0 \left[\ln \left(\frac{5\nu}{\nu - 1} \right) \right]^2. \tag{6.60}$$

183

Knowledge of both f_{3dB} and f_{7dB} will uniquely determine the thermal frequency response.

As the first example, the 3dB and 7dB thermal bandwidths of a metallic heater in Atabaki *et al.* [11] are given as 132 kHz and 284.5 kHz, respectively. The contours of these two bandwidths are plotted in Fig. 6.8a as functions of f_0 and ν. The intersection of the two curves determines f_0 = 30 kHz and ν = 1.325. A comparison of the normalized frequency response from Atabaki *et al.* [11] and our proposed model given in Eq. (6.57) is plotted in Fig. 6.8b.

As the second example, the 3dB and 7dB thermal bandwidths of a doped-silicon heater in Harris *et al.* [70] are given by 130 kHz and 500 kHz, respectively. Contours of these two frequencies are plotted in Fig. 6.8c with an intersection occurring at f_0 = 141.6 kHz and ν = 4.262. Fig. 6.8d plots our proposed frequency model (normalized) against the measurements presented in [70]. A good agreement is again observed. Note that the intersection of 3dB and 7dB contours occurs at lower values of f_0 and ν in Fig. 6.8a, whereas in Fig. 6.8c it occurs at large values of f_0 and ν. This means that the thermal frequency response of the first example is sharper than the second example.

6.4   Thermo-Optic Transient Response of Waveguide

The change in the waveguide temperature will cause an instantaneous change in the phase of the light propagating inside the waveguide as $\Delta\phi(t) = H_\phi \times \Delta T_{wg}(t)$ where H_ϕ is given by

$$H_\phi = \frac{L}{\lambda}\left(2\pi\frac{dn_{\text{eff}}}{dT}\right) \approx 0.0013\,\frac{L}{\lambda} \tag{6.61}$$

for a 450 nm$\times$220 nm strip waveguide. L is the length of the waveguide exposed to the thermal effects and λ is the wavelength of light.
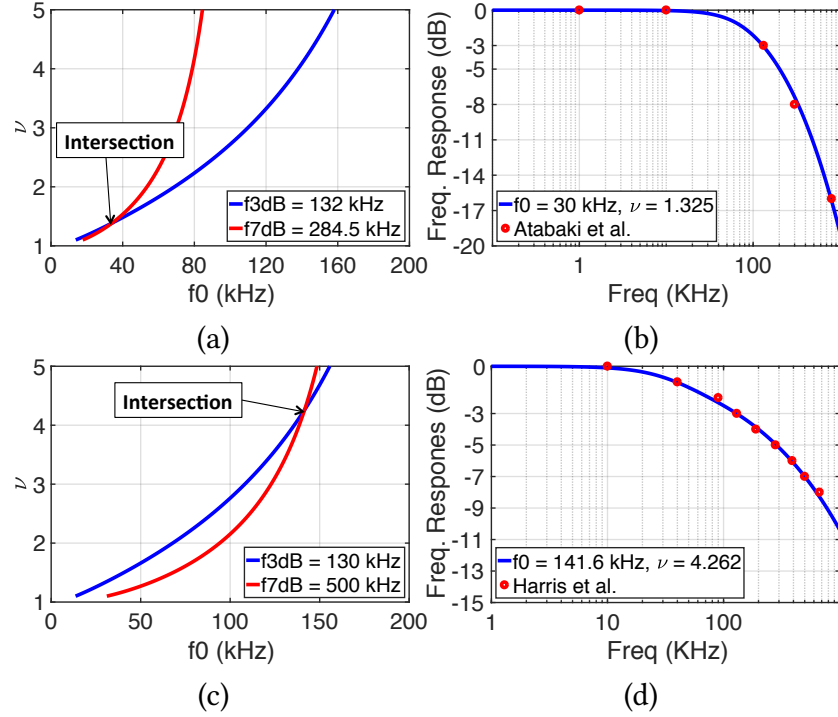
Figure 6.8: (a) Contours of 3dB (blue curve) and 7dB (red curve) thermal bandwidth of a metallic heater given in [11]. The intersection provides the values of f_0 and ν. (b) Comparison of frequency response in [11] (red circles) and our proposed model (solid) by matching 3dB and 7dB bandwidths. A good agreement is observed. (c) Contours of 3dB and 7dB bandwidth of a doped-silicon heater given in [70]. (d) Comparison of measured frequency response from [70] and our proposed model.

6.5 Verification with Finite Element Method

In this section, COMSOL, a finite element analysis tool is used to estimate the transient thermal response of heater-waveguide system. The width and height of the heater are set to 5 μm and 0.2 μm, respectively. The silicon waveguide is situated 1 μm below the heater and the space between them is filled with silica. The boundaries of the simulation area are located 10 μm from each edge of the heater and are set to thermal insulation boundary condition. The initial values are set to zero and the heat flow is modeled by setting the temperature of heater as a function of voltage according to Eq. (6.18) with parameters K_v = 0.6 V^{-2} and α_H = 4.5×10^{-3} K^{-1}. Note that the choice of metal for the heater affects α_H and R_{linear} parameters (and hence K_v) even if all the heaters have the same geometrical

design. However, the goal here is to show how the temperature of the waveguide changes as a function of temperature of the heater. Therefore, the boundaries of the heater are set to a fixed temperature according to Eq. (6.18).

Thermal Impulse Response

In order to find the transient shape of the thermal impulse response, a very short pulsed voltage (50 nsec duration) with 1V amplitude is applied to the heater. This will set the temperature of the heater to 29.5 °K during the 50 nsec period. Then the heat transfer equation is solved. Fig. 6.9a shows the temperature distribution at the time when the temperature of the waveguide is at its peak. Fig. 6.9b shows the comparison between the analytic model (solid) with 132 kHz of 3dB thermal bandwidth and the result from COMSOL (dotted). A good agreement is observed.

Step Response

The step response can be calculated from the thermal impulse response using Eq. (6.37). If a step voltage is applied to the heater at $t = 0$, its temperature response is also a step according to Eq. (6.18). The step response of the waveguide is then calculated as

$$\Delta T_{wg}(t) = \int_0^t \Delta T_H(t - \tau) I_{wg}(\tau) d\tau = \Delta T_H(0^+) \int_0^t I_{wg}(\tau) d\tau. \tag{6.62}$$

This integral can be calculated with a closed-form solution. First, we use a change of variable

$$u = \frac{1}{4\pi f_0 t} \rightarrow du = -4\pi f_0 u^2 dt \rightarrow I_{wg}(t) \approx \frac{\nu - 1}{\nu} \times 4\sqrt{\pi} f_0 u^{3/2} e^{-u} \tag{6.63}$$

and therefore

$$\int_0^t I_{wg}(x) dx \approx \frac{\nu - 1}{\nu} \times \frac{1}{\sqrt{\pi}} \int_u^\infty z^{-0.5} e^{-z} dz \tag{6.64}$$

now we use another change of variable

$$z = p^2 \rightarrow du = 2pdp \rightarrow \int_u^{\infty} z^{-0.5} e^{-z} dz = 2 \int_{\sqrt{u}}^{\infty} e^{-p^2} dp. \tag{6.65}$$

The last integral is the definition of the complementary error function as

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt. \tag{6.66}$$

Using this definition, the integral of the thermal impulse response is given by

$$\int_0^t I_{wg}(x) dx \approx \frac{\nu - 1}{\nu} \times \text{erfc}(\sqrt{u}) \approx \frac{\nu - 1}{\nu} \times \text{erfc}\left(\frac{1}{\sqrt{4\pi f_0 t}}\right) \tag{6.67}$$

and therefore the impulse response of temperature is

$$\Delta T_{wg}(t) \approx \frac{\nu - 1}{\nu} \Delta T_H(0^+) \times \text{erfc}\left(\frac{1}{\sqrt{4\pi f_0 t}}\right). \tag{6.68}$$

Note that $\text{erfc}(0) = 1$ and $\text{erfc}(\infty) = 0$. Based on this solution, we can easily estimate the 10% to 90% rise time of the system. Using the change of variable $u = 1/(4\pi f_0 t)$ and the fact that $(\nu - 1)/\nu \times \Delta T_H(0^+)$ is the final value of the waveguide response, we can write:

$$\text{erfc}(\sqrt{u_{0.1}}) = 0.1 \rightarrow \sqrt{u_{0.1}} = 1.162 \rightarrow u_{0.1} = 1.3525$$

$$\text{erfc}(\sqrt{u_{0.9}}) = 0.9 \rightarrow \sqrt{u_{0.9}} = 0.087 \rightarrow u_{0.9} = 0.0079. \tag{6.69}$$

Therefore, the rise time is estimated as

$$\Delta t_{10\% \rightarrow 90\%} = \frac{10.0143}{f_0} \approx \frac{10}{f_{3\text{dB}}} \left[ln\left(\frac{2\nu}{\nu - 1}\right) \right]^2. \tag{6.70}$$
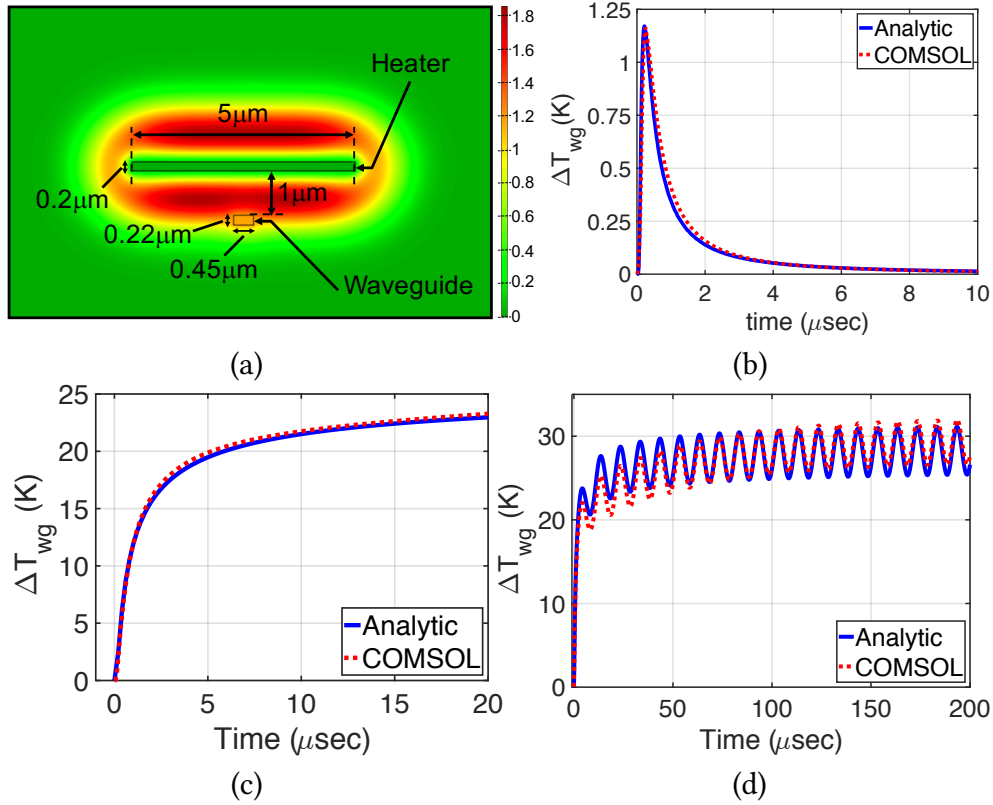
Figure 6.9: (a) Structure of metallic heater and silicon waveguide in COMSOL simulations. The width of the heater is set to 5 μm and the waveguide is 1 μm below the heater. (b) Analytic (solid) and COMSOL results (dotted) of a 50 nsec pulse excitation of the heater. (c) Step response of the waveguide temperature. (d) Sinusoidal response of the waveguide for a 100 kHz excitation of the heater.

As an example, if the 3dB bandwidth is about 132 kHz and $\nu \to \infty$, the rise time is approximated to be 36 μsec.

We applied a 0V to 1V step voltage to the heater in COMSOL and calculated the step response. The result is shown in Fig. 6.9c (dotted) along with the predicted analytic model (solid). Since the boundaries are set at 10 μm distance from the heater, $\nu = 10$ is set in Eq. (6.68). A good agreement is observed between the results.

Steady-State Sinusoidal Response

Now we assume that the heater is excited by a sinusoidal voltage biased with a large DC part, $V_H(t) = V_0 + V_i \sin(\omega_i t)$ where $\omega_i = 2\pi f_i$. Due to nonlinear nature of Eq. (6.18), all

the harmonics are presented in the temperature change of the heater. However, as long as the AC amplitude is small, we can only include the DC and first harmonic. Therefore $\Delta T_H(t) \approx \Delta T_H^{\mathrm{dc}} + \Delta T_H^{\mathrm{ac}} \sin(\omega_i t)$ and the waveguide temperature response is $\Delta T_{wg}(t) \approx \Delta T_{wg}^{\mathrm{dc}} + \Delta T_{wg}^{\mathrm{ac}} \sin(\omega_i(t - \tau_{\mathrm{delay}}))$ where the DC and the AC parts are given by

$$\Delta T_{wg}^{\mathrm{dc}} \approx \frac{\nu - 1}{\nu} \Delta T_H^{\mathrm{dc}} \tag{6.71a}$$

$$\Delta T_{wg}^{\mathrm{ac}} \approx \Delta T_H^{\mathrm{ac}} \exp\left(-\sqrt{\frac{f_i}{f_0}}\right) \tag{6.71b}$$

$$\tau_{\mathrm{delay}} \approx \frac{1}{2\pi\sqrt{f_i \, f_0}}. \tag{6.71c}$$

Fig. 6.9d shows an example of analytic calculation (solid) and COMSOL simulation (dotted) of sinusoidal response with parameters $f_i = 100$ kHz, $V_0 = 1$V, and $V_i = 0.1$V. The delay between the sinusoidal excitation and the waveguide thermal response is estimated to be 0.96 μsec in this example which agrees well with the COMSOL result. Figure 6.10 shows the developed graphical interface based on the mathematical model for calculating the step response of the heater-waveguide system.

6.6   Pulse Width Modulation and Thermal Rectification

The lowpass filtering nature of the thermal impulse response of the heater waveguide system causes the system to not react to very high frequencies. Considering the analysis of sinusoidal respond in the previous section, if the frequency of the excitation voltage is too high then its effect will be almost completely suppressed. This fact brings up the question if it is possible to drive the heater that requires an analog signal with a digital pulse-shaped signal to get any arbitrary step response. Assume a periodic square signal with pulse amplitude V_i, frequency $f_i = 1/T_i$, and duty cycle $D = \tau_i/T_i$ is applied to the heater. This voltage will trigger a thermal response at the heater as plotted in Fig. 6.11a. The amplitude of the resulting thermal pulse, ΔT_i, is given by Eq. (6.18). Figure 6.11b
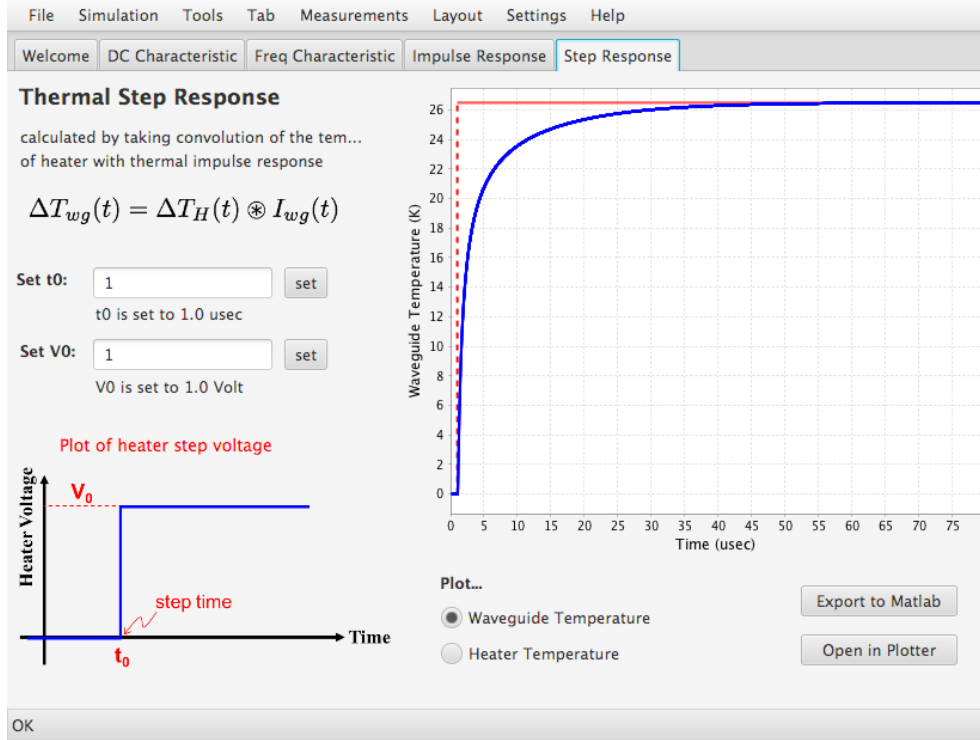
Figure 6.10: A developed graphical interface for calculating the thermo-optic step response of the heater-waveguide system.

describes how the low-pass system will interpret this thermal excitation as a step excitation if the frequency and duty cycle are properly adjusted. To prove this, we look at the Fourier series of the heater temperature:

$$\Delta T_H(t) = D\,\Delta T_i \left(1 + 2 \sum_{m=1}^{\infty} \operatorname{sinc}(mD) \cos(m(\omega_i t - \pi D)) \right) \tag{6.72}$$

Each component of this harmonic series excites a corresponding harmonic of the waveguide temperature with different amplitudes and time delays. In order for this series to result in a step response, we should ensure that all the harmonics (except for the DC component) are adequately attenuated. In this case, the DC component of the waveguide temperature is $\Delta T_{wg}^{\mathrm{dc}} = (\nu - 1)/\nu \times D\Delta T_i$ (see Eq. 6.71a) which is a linear function of
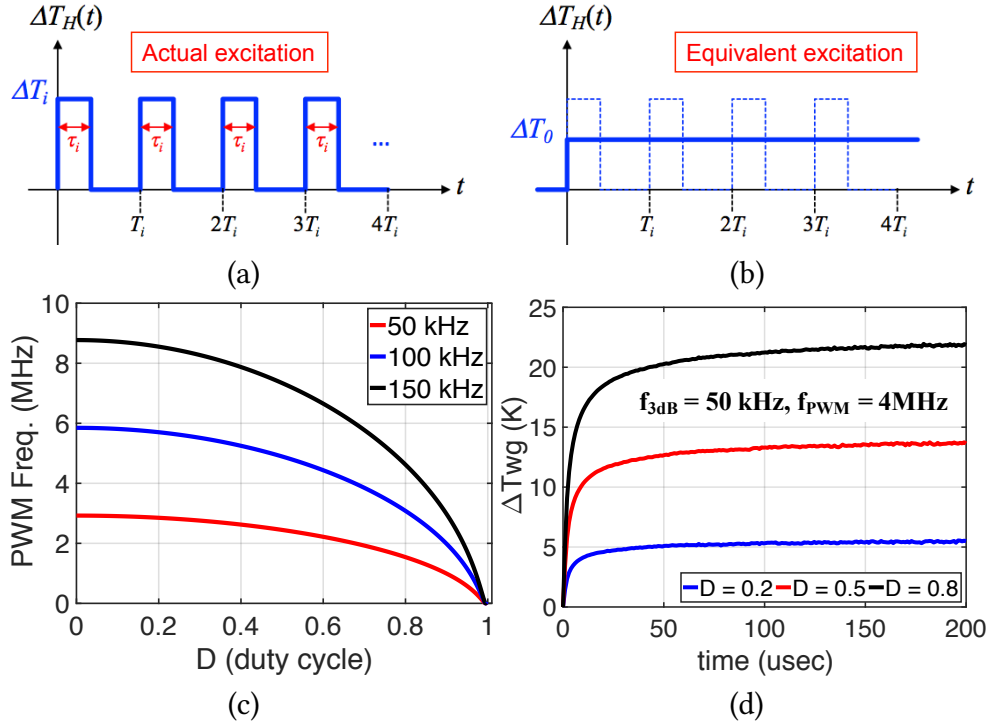
Figure 6.11: (a) PWM temperature of the heater that consists of periodic pulses with peak-to-peak amplitude denoted by ΔT_i. (b) Equivalent thermal step drive of the heater based on the PWM drive. The amplitude of the step drive is $\Delta T_0 = D \, \Delta T_i$. (c) Plot of the required frequency for PWM signal based on the duty cycle. The relative attenuation of the first harmonic (i.e. ripples) is set to 0.01. (d) Thermal response of waveguide for three different duty cycles.

duty cycle. The ratio between the first harmonic and the DC component is given by

$$\frac{\Delta T_{wg}^{\mathrm{ac}}}{\Delta T_{wg}^{\mathrm{dc}}} \approx \frac{2\nu}{\nu - 1} \, \mathrm{sinc}(D) \exp\left(-\sqrt{\frac{f_i}{f_0}}\right) < A \tag{6.73}$$

where A denotes the desired relative attenuation of the first harmonic. This yields the following lower-bound on the frequency of the PWM signal based in its duty cycle, D, and ripple suppression level, A:

$$f_i > f_0 \times \left[\ln\left(\frac{1}{A} \frac{2\nu}{\nu - 1} \, \mathrm{sinc}(D)\right)\right]^2 . \tag{6.74}$$

Fig. 6.11c plots this lower bound as a function of duty cycle for the relative attenuation

191

of first harmonic set to $A = 0.01$ and three cases where the 3dB thermal bandwidth of the system is set to 50 kHz, 100 kHz, and 150 kHz, respectively, and $\nu = \infty$. Thermal parameters of heater are set to $K_v = 0.6$ V^{-2}, and $\alpha_H = 4.5 \times 10^{-3}$ K^{-1}. The required PWM frequency rises as the duty cycle decreases and the maximum frequency occurs at $D = 0$. Figure 6.11d plots the thermal response of the waveguide for three different duty cycles with 4 MHz drive frequency and 50 kHz thermal bandwidth. As can be seen, the thermal response of the waveguide is very close to that of a step excitation.

PWM Thermo-Optic Response of Microring Resonators

Two common structures of microring resonators are the "all-pass" in which one waveguide is coupled to the ring, and the "add-drop" configuration in which two waveguides are coupled to the ring. For an "all-pass" configuration the Lorentzian transmission spectrum [105] at the output port is given by $P(x) = 1 - A/(1 + x^2)$ and for an "add-drop" configuration the spectrum at the drop port is given by $P(x) = A/(1 + x^2)$, where A determines the extinction ratio of resonance and x is the normalized deviation of the input laser from the resonance wavelength given by $x = (2/\Delta\lambda_{3dB}) \times (\lambda_{input} - \lambda_{res})$. A change in the temperature of the waveguide, $\Delta T_{wg}(t)$, causes a change in the resonance, $\Delta\lambda_{res}(t)$ according to

$$\Delta\lambda_{\text{res}} \approx \frac{\Delta\phi(t)}{2\pi} \times \text{FSR}_{nm} = \left(\frac{H_\phi}{2\pi} \times \text{FSR}_{nm}\right)\Delta T_{wg}(t). \qquad (6.75)$$

Due to the Lorentzian lineshape, the behavior of the output optical power depends on the input wavelength. As shown in Fig. 6.12a, if the input wavelength is set close to the resonance of the ring then the change in optical power is linearly related to the change in waveguide temperature. However, if the input wavelength is tuned exactly to the resonance as shown in Fig. 6.12b, the optical power has a quadratic dependence on the temperature of the waveguide. This can be confirmed by noting that the first derivative
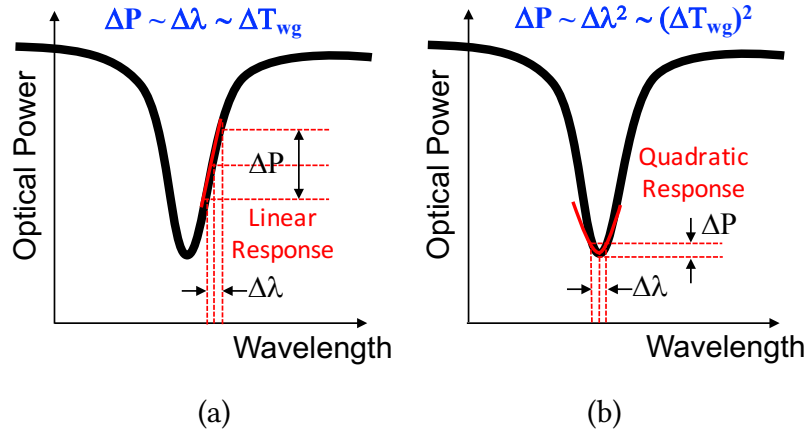
Figure 6.12: (a) Region of linear small-signal thermo-optic response of the microring resonator. In this region, the change in optical power is linearly proportional to the change in the temperature of waveguide. (b) Region of nonlinear small-signal thermo-optic response of microring resonator. In this region, the change in optical power has a quadratic dependence on the change in temperature.

of $P(x)$ with respect to x is zero at $x = 0$. If a PWM signal is applied to the heater and the ring is working in the linear regime, then the amplitude of the first harmonic on the optical output is proportional to $D \operatorname{sinc}(D)|\mathbf{t_{wg}}(f_i)|$ and therefore Eq. (6.73) that describes the ripples on the temperature also applies for the optical power. On the other hand, if the ring is working in the quadratic regime then the steady-state DC level of the output power is proportional to $D^2 \Delta T_i^2 |\mathbf{t_{wg}}(f = 0)|^2$ and the strength of the first harmonic is proportional to $4D^2 \Delta T_i^2 \operatorname{sinc}(D)|\mathbf{t_{wg}}(f_i)|$. In this case the relative suppression of the optical ripples is

$$\frac{\Delta P_{ac}}{\Delta P_{dc}} = \left(\frac{2\nu}{\nu - 1}\right)^2 \operatorname{sinc}(D) \exp\left(-\sqrt{\frac{f_i}{f_0}}\right) \tag{6.76}$$

which can be used to determine the required PWM frequency based on the duty cycle. For example, if $\nu = 3$ and $f_0 = 30$ kHz, a duty cycle of 50% requires at least a PWM frequency of 1.21 MHz for more than 99% suppression of thermal ripples. A more detailed treatment of the first harmonic of the Lorentzian response is provided in [23]. The peak-to-peak voltage of the PWM control signal and its duty cycle will ultimately determine the tuning

range of the resonance. Note that the absolute amplitude of the first harmonic of the optical power is dependent on the peak-to-peak voltage of the heater and therefore a larger peak-to-peak voltage will in general require a higher PWM frequency to keep the amplitude of the first harmonic below a certain value. However, in Eq. (6.73) and Eq. (6.76) we define the suppression of the first harmonic as the relative error with respect to the DC value which is independent of the peak-to-peak voltage.

6.7 Thermo-Optic Efficiency of Microrings

Based on the developed models for the thermal response, the thermo-optic tuning efficiency of the resonance of a microring resonator can be derived by combining Eq. (6.18), Eq. (6.20), Eq. (6.71a), and Eq. (6.75) which results in

$$
\text{eff} = \frac{\Delta\lambda_{\text{res}}}{P_H} \approx \left(\frac{dn_{\text{eff}}}{dT} \frac{\lambda_{\text{res}}}{n_g} \right) \left(N \frac{\nu - 1}{\nu} \right) \left(\frac{1}{4\alpha_H} R_{\text{linear}} K_v \right),
\tag{6.77}
$$

showing a constant value. Therefore, the relation between $\Delta\lambda_{\text{res}}$ and P_H is a linear behavior. Here, N is a factor that describes what fraction of the circumference of the microring is exposed to the heat ($0 < N < 1$). If the entire ring is exposed to heat, then N = 1. Equation (6.77) is useful in determining the upper limit on the thermo-optic efficiency of integrated microheaters. As shown in Table 6.1, assuming that for the three types of heaters (metallic, externally doped, internally doped), the ring is made of 450×220 strip waveguides, we can see that optimistic values for the thermo-optic parameters lead to thermo-optic efficiencies in the range of 1–1.9 nm/mW. By further tweaking the design parameters, it may be possible to increase the efficiency of microring tuning up to 2 nm/mW. *However, going beyond 2 nm/mW seems to be very hard to achieve.* In order to put these efficiency numbers into perspective, a survey of published microrings with heaters is presented in Table 6.2.

We now look at the trade-offs between the heating speed (3dB thermal bandwidth) of

Table 6.1: Optimistic evaluation of thermo-optic efficiency of microrings with three different types of integrated heaters.

Type	Metallic	Doped (external)	Doped (internal)
Waveguide	Strip $450{\times}220$	Strip $450{\times}220$	Strip $450{\times}220$
Mode	TE_{00}	TE_{00}	TE_{00}
$\lambda_{\text{resonance}}$ (nm)	1550	1550	1550
n_{eff}	2.35	2.35	2.35
n_{g}	4.393	4.393	4.393
dn_{eff}/dT (K^{-1})	2.1×10^{-4}	2.1×10^{-4}	2.1×10^{-4}
N	1	0.9	0.5
ν	~ 3	~ 10	large
α_H (K^{-1})	5×10^{-3}	3×10^{-3}	3×10^{-3}
K_v	0.5	0.3	0.6
R_{linear} (kΩ)	1	1	1
eff. (nm/mW)	1.23	1.5	1.853

Table 6.2: Survey of thermo-optic efficiency (nm/mW) of published microring structures with integrated heaters.

Structure	Heater	Efficiency	Reference	Year
Filter	External	0.13	Amatya et al. [6]	2007
Filter	External	0.26	Qiu et al. [150]	2011
Modulator	External	0.32	Zheng et al. [213]	2014
Modulator	External	0.38	DeRose et al. [48]	2014
Filter	Internal	1.08	Su et al. [177]	2015
Modulator	Internal	1.14	Timurdogan et al. [186]	2014
Modulator	Internal	1.14	Zortman et al. [219]	2012
Filter	Internal	1.23	Sun et al. [180]	2016
Modulator	Internal	1.71	Timurdogan et al. [188]	2013
Filter	Internal	1.78	Timurdogan et al. [187]	2012
Filter	Internal	1.81	Watts et al. [198]	2009

a microring and the heating efficiency (thermo-optic shift of resonance) for a metallic microheater. There are four main physical parameters of the heater-ring structure involved
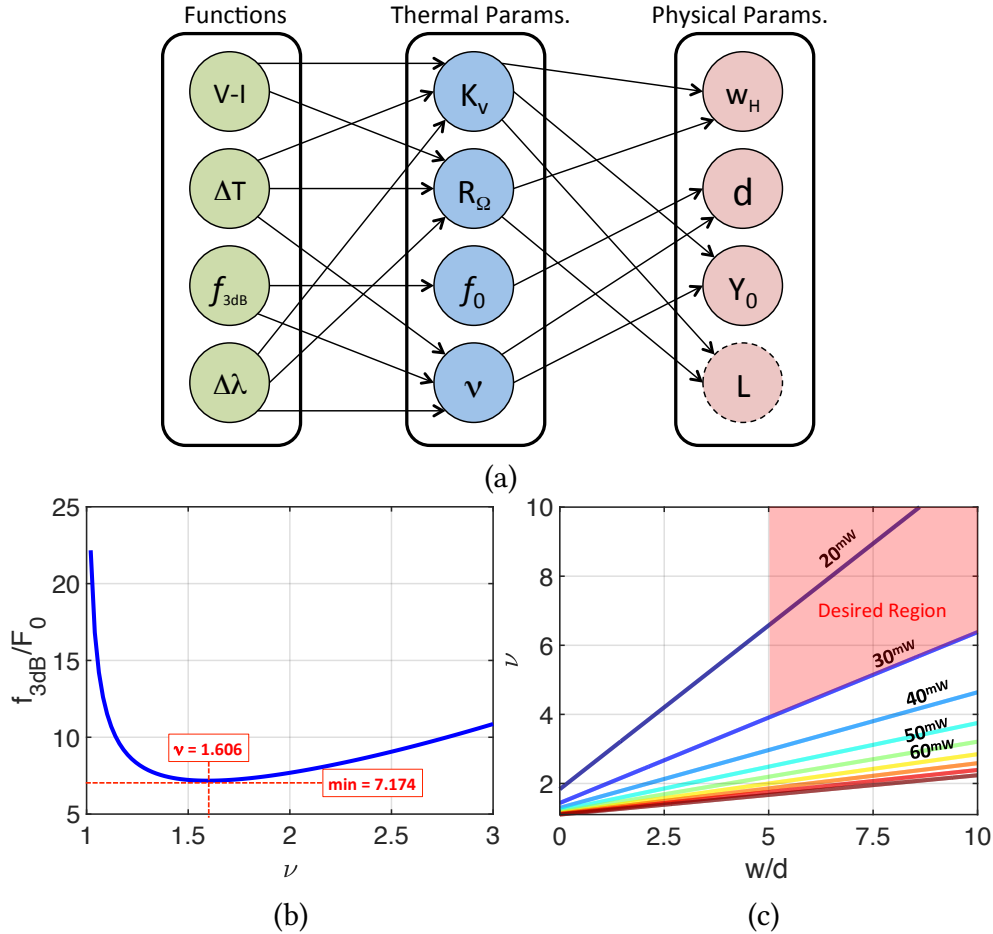
Figure 6.13: (a) breakdown of the relation between thermal variables and the physical parameters of the heater-ring system. For V-I of the heater, L refers to the length of the heater while for $\Delta\lambda$ it refers to the length of the silicon waveguide exposed to the heat. (b) Plot of normalized 3dB thermal bandwidth as a function of ν. The minimum point occurs at $\nu = 1.606$. (c) Contours of required mW ohmic power for one FSR shift of resonance of a microring.

in the design space as shown in Fig. 6.13a: width of the heater, w_H, vertical distance of heater from ring, d, vertical distance of heater from silicon substrate, Y_0, and the length of heater, L. Note that here we assume the thickness of heater is fixed and is not subject to variation and optimization. There are four thermal parameters two of which (K_v, R_Ω) are specifically for the heater and the other two (f_0, ν) are for the thermal response of the microring. Finally, there are four functions that are subject to variation and optimization.

In order to see the trade-offs between thermal speed and efficiency, we consider f_0

and ν as the main optimization parameters. Using Eq. (6.77) and (6.16) it can be shown that the heating efficiency is given by

$$\text{eff} = \frac{\Delta\lambda_{\text{res}}}{P_H} = \left(\frac{dn_{\text{eff}}}{dT}\frac{\lambda_{\text{res}}}{n_g}\right) \times \frac{1}{L_H k_{sio_2}} \times \frac{\nu - 1}{\sqrt{f_0/f_w} + 0.88\,\nu} \tag{6.78}$$

where we defined $f_w = \xi_{sio_2}/(\pi\, w_H^2)$. If we assume that the heater covers the entire microring, then $L_H = 2\pi R$ and therefore:

$$\text{eff} = \frac{\Delta\lambda_{\text{res}}}{P_H} = \left(\frac{dn_{\text{eff}}}{dT}\frac{\text{FSR}}{\lambda_{\text{res}}}\right) \times \frac{1}{k_{sio_2}} \times \frac{\nu - 1}{\sqrt{f_0/f_w} + 0.88\,\nu}. \tag{6.79}$$

Now it is easy to see that

$$\frac{\partial f_{3\text{dB}}}{\partial \nu} < 0 \quad , \quad \frac{\partial f_{3\text{dB}}}{\partial f_0} > 0 \tag{6.80}$$

whereas

$$\frac{\partial \text{eff}}{\partial \nu} > 0 \quad , \quad \frac{\partial \text{eff}}{\partial f_0} < 0 \tag{6.81}$$

and hence a trade-off exists. Since for the application of thermal tuning of a microring, heating efficiency is in general more important than the thermal speed, we give the priority to a larger ν and smaller f_0. Next, we see that the plot of $f_{3\text{dB}}$ as shown in Fig. 6.13b has a minimum point at $\nu = 1.606$ given by

$$f_{3\text{dB}} \geq 7.174 \left(\frac{\xi_{sio_2}}{\pi\, Y_0^2}\right) = 7.174\, F_0. \tag{6.82}$$

Therefore, in order to ensure that the 3dB bandwidth is at least 100 kHz regardless of the position of waveguide, the condition $Y_0 < 4.46\ \mu$m must be satisfied. Finally, we calculate the contours of the required ohmic power (mW) to tune the resonance to a full FSR and plot them in Fig. 6.13c as a function of ν and w_H/d. The power for tuning the resonance

over a full FSR range is given by

$$\text{full FSR:} \quad P_H = k_{sio2} \times \frac{\lambda_{\text{res}}}{dn_{\text{eff}}/dT} \times \frac{\sqrt{f_0/f_w} + 0.88\,\nu}{\nu - 1}. \tag{6.83}$$

which shows that the contours of constant P_H are lines in the $(\nu, w/d)$ plane. Considering that small ring resonators with radii on the order of 5μm or smaller (FSR > 20 nm) are desirable for dense WDM photonic links [14, 18], we put an upper bound of 30 mW on the tuning power to ensure a tuning efficiency of at least 0.67 nm/mW. We also put the constraint that $w_H/d > 5$ to make sure the heat diffusion from heater to waveguide is uniform enough such that the 1D equations hold. The desired region based on these constraints is highlighted in Fig. 6.13c indicating that $\nu > 4$ must be satisfied which results in the condition $d < 1.2$ μm.

6.8 Experimental Demonstrations

In this section, we present the experimental results on the thermo-optic response of microring resonators and show that the thermo-optic models presented in the previous sections agree well with the experimental results. As a practical example, we further present the use of the thermal models for controlling microring resonators with PWM signals. For completeness, we consider two separate devices equipped with integrated micro-heaters: the first one is an "all-pass" configuration and the second one is an "add-drop" configuration.

Device 1: All-Pass Ring Resonator

The band-stop or notch tunable silicon ring filter was designed by Analog Photonics [227] as part of the AIM Photonics [226] process design kit (PDK) and fabricated through a multi-project wafer (MPW) run. The tuning is achieved by an integrated silicon resistor

(a)

(b)

(c)

(d)

(e)
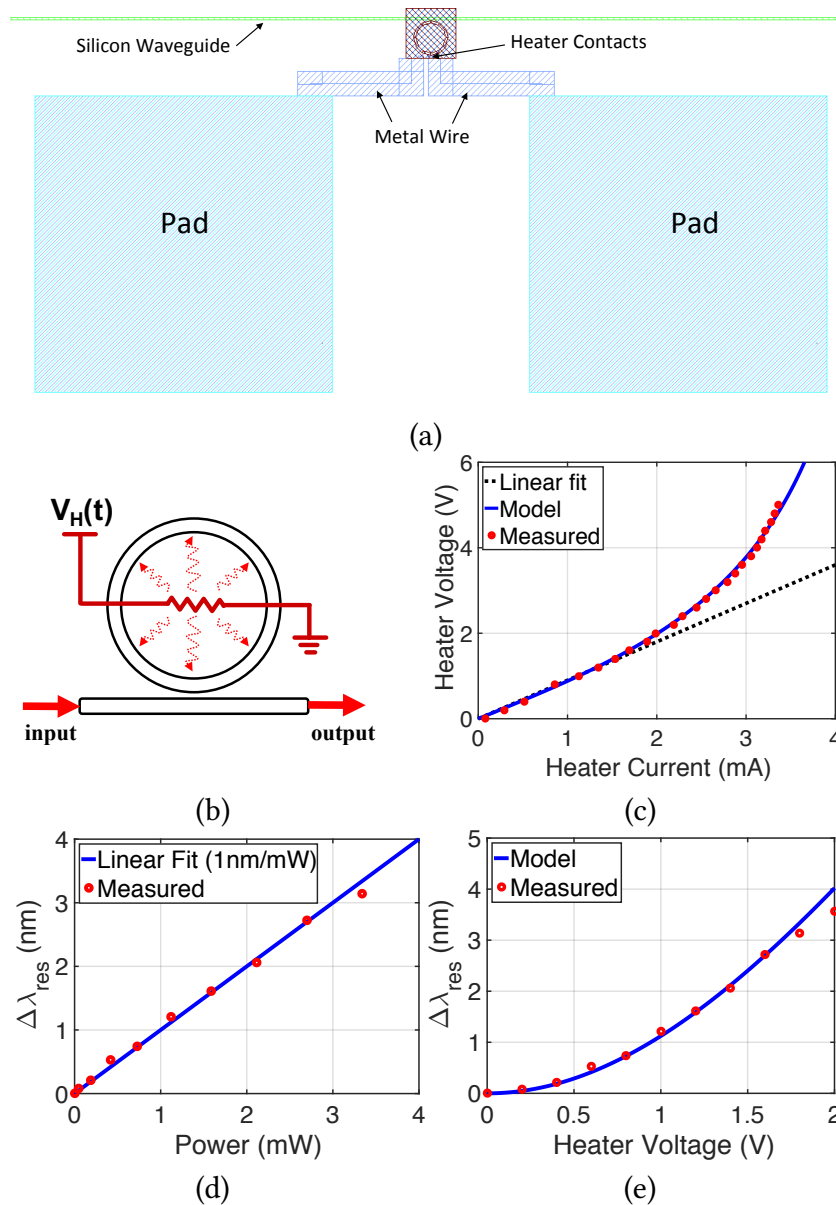
Figure 6.14: (a) A picture of the exact layout of the all-pass ring from the designed GDS file. (b) Schematic of the "all-pass" microring configuration. (c) Measured (red circles) and Modeled DC characteristic of the heater. Extracted parameters are R_{linear} = 850 Ω and K_v = 0.2 V^{-2}. (d) Stationary thermo-optic response of the microring as a function of heater voltage. A linear fit indicates a thermo-optic efficiency of 1 nm/mW. (e) Stationary thermo-optic response as a function of heater voltage.
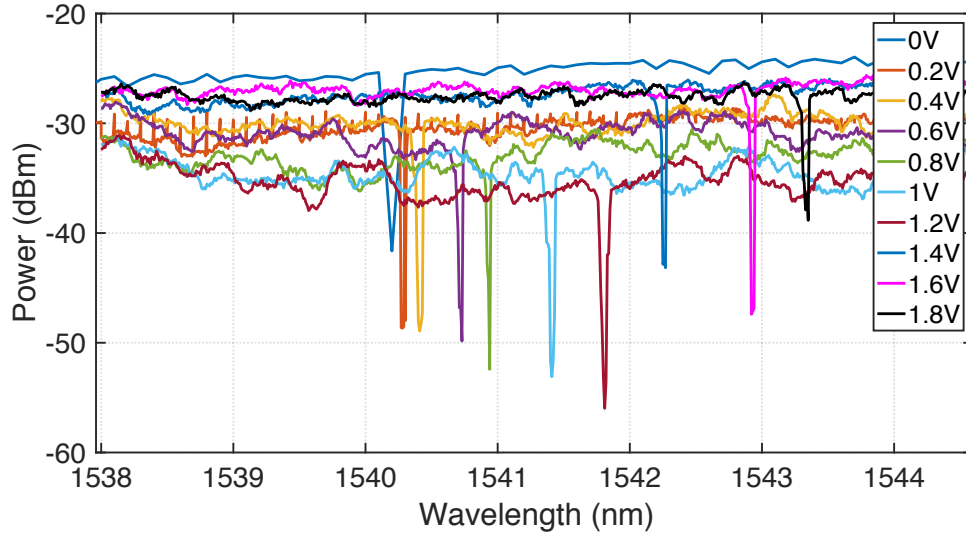
Figure 6.15: Measured spectra of the all-pass microring structure under various bias voltages applied to the heater.

that acts as a joule heater. The resistor is formed by doping the silicon and placed away from the optical mode and contacts to ensure a low optical loss and a high thermal tuning efficiency (external heater). This device has a resonance at λ = 1540.2 nm. A picture of the designed GDS file is shown in Fig. 6.14a. A schematic of this structure is shown in Fig. 6.14b. The DC measured V-I characteristic of the integrated heater is shown in Fig. 6.14c (red circles) along with the linear fit shown with dotted line. The linear resistance is estimated to be R_{linear} = 850 Ω and K_v is set to 0.2 so that Eq. (6.20) matches the measurements as shown in Fig. 6.14c with the solid curve. Next, we examined the stationary thermo-optic response of the microring by tuning the laser wavelength to the resonance while sweeping the heater voltage from 0V to 2V and measuring the shift of resonance as shown in Fig. 6.15. The results are plotted in Fig. 6.14d as a function of heater power (in mW units) and in Fig. 6.14e as a function of heater voltage. A linear curve has been fitted to the data indicating a thermo-optic efficiency of 1 nm/mW.

In order to characterize the transient response, we applied a small-signal sinusoidal voltage with enough DC bias such that it always stays positive, and measured the peak-to-peak amplitude of the optical output after being converted to electrical voltage by a

photodiode. Due to the presence of a DC component in the heater voltage, the output optical power is linearly proportional to the temperature of waveguide in the small-signal approximation. The frequency of the applied voltage was swept from 5 kHz to about 150 kHz and both the amplitude and phase delay of the output were measured. Figure 6.16a shows the normalized measured amplitudes (red circles) in dB units and Fig. 6.16b shows the measured phase delays in radian units. We fitted the frequency response model of Eq. (6.57) to the 3dB and 7dB measured bandwidths by setting f_0 = 30 kHz and ν = 2.223, and then plotted the normalized amplitude and phase delay in Fig. 6.16a and Fig. 6.16b with solid curves. A good agreement is observed validating our frequency model. The 3dB thermal bandwidth is estimated to be 50 kHz. Based on the measured 1 nm/mW thermal efficiency, α_H is extracted from Eq. (6.77) to be 1.7×10^{-3} K^{-1} and then used to calculate the shift of resonance as a function of heater voltage. The results are plotted as a solid curve in Fig. 6.14e, which agree well with the measurements. Fig. 6.16c shows an example of the measured heater voltage and output voltage at 60 kHz. The time delay between the peaks of the two signals is observed to be 3.8 μsec, which is very close to the estimated value of 3.7513 μsec from Eq. (6.71c). Based on Eq. (6.75), we estimate the thermal sensitivity of the resonance to be $\Delta\lambda_{\mathrm{res}}/\Delta T_{wg}$ = 0.073 nm/$^\circ$K or 9.07 GHz/$^\circ$K.

Next, we performed a time-domain small-signal measurement (V_{pp} = 100 mV) of the thermo-optic ripples due to the PWM drive of the heater. The definition of the peak-to-peak ripple amplitude is shown in Fig. 6.17a. As the frequency of the PWM drive increases, the ripples of the optical output get closer to a sinusoidal behavior. We applied a 150 kHz PWM signal to the heater and measured the amplitudes of the optical ripples for 20%, 40%, 50%, 60%, and 80% duty cycles. The measured ripple amplitudes are normalized to their maximum value and plotted in Fig. 6.17b as a function of duty cycle. Since the input laser is tuned to the resonance, the quadratic behavior of thermo-optic response [see Fig. 6.12b] is in effect. If we assume that the ripples are mainly due to the first harmonic of the thermal response considering the small-signal voltage applied to the heater, the

(a)

(b)

(c)
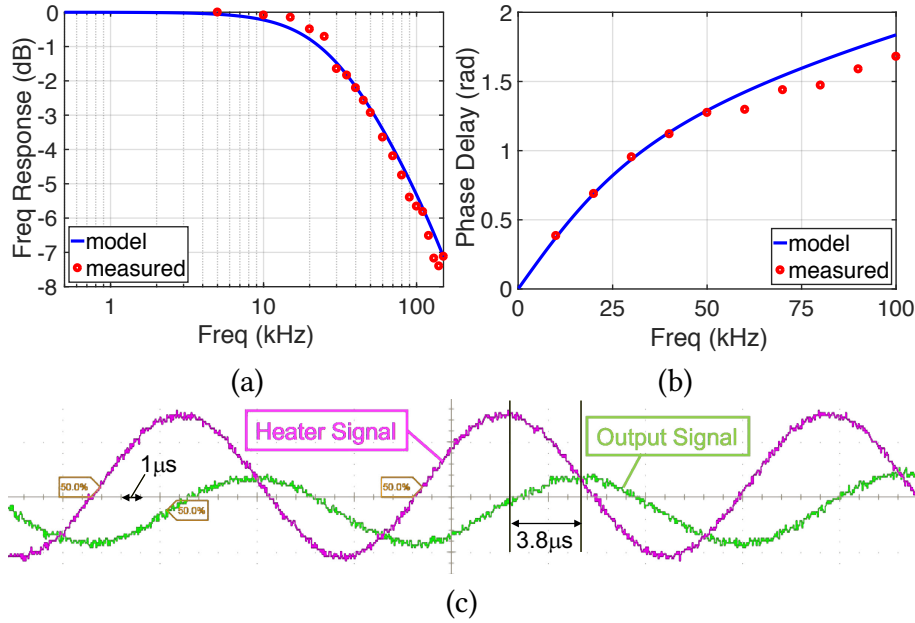
Figure 6.16: (a) Normalized amplitude thermal frequency response. The model fits the measurements by setting f_0 = 30 kHz and ν = 2.223. The 3dB thermal bandwidth is 50 kHz. (b) Phase delay of the thermal frequency response. (c) Measured heater voltage and output optical power (after photodiode) at 60 kHz. The time delay between the peaks is about 3.8 μsec.

normalized behavior of optical ripple amplitude as a function of duty cycle is given by $1.73 \times D \sin(\pi D)$ based on the previous analysis. The maximum amplitude of the ripples is predicted to happen at 65% duty cycle. This curve is plotted in Fig. 6.17b which has a good agreement with the measured data. Note that the asymmetric behavior around 50% is due to the quadratic optical response of the microring. If the laser was set to the linear region of the Lorentzian curve, the ripple amplitude would have behaved symmetric around 50% duty cycle.

Device 2: Add-Drop Ring Resonator

The second device is an "add-drop" configuration equipped with integrated doped-silicon micro-heater (internal heater) as schematically shown in Fig. 6.18a. This device has a resonance at 1545.8 nm and the optical 3dB bandwidth of this resonance is about 0.7 nm [66].
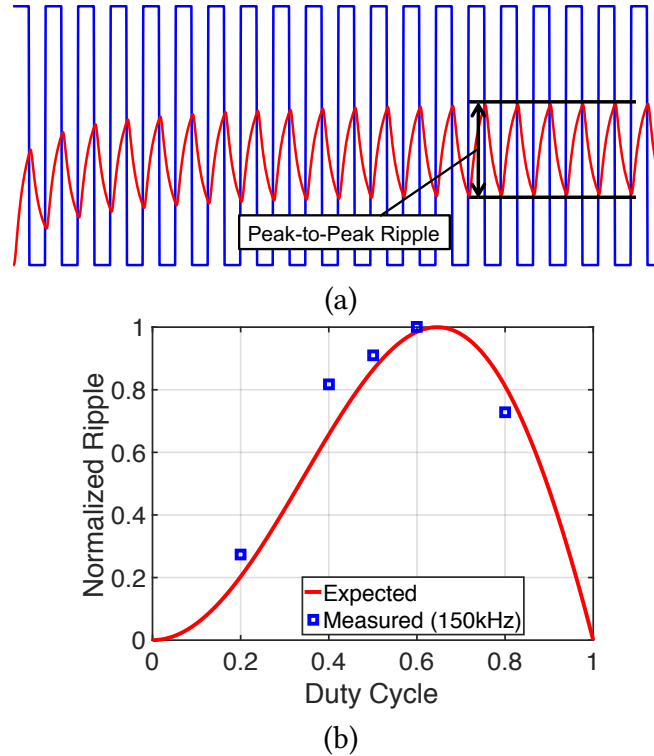
Figure 6.17: (a) Example of ripples on the waveguide thermal response to the PWM excitation of the heater. (b) Comparison of measured (squares) normalized ripple amplitudes for 150 kHz small signal PWM drive of heater as a function of duty cycle. Solid curve shows the ripple amplitude based on the first harmonic of the thermo-optic response. A good agreement is observed.

First, the DC V-I characteristic of the heater was measured and plotted in Fig. 6.18b. The extracted parameters are K_v = 0.6 V^{-2} and R_{linear} = 161 Ω. Then we applied DC voltages from 0V to 4V in 0.2V steps to the heater and measured the shift of resonance. The results are plotted in Fig. 6.18c and 6.18d. The heating efficiency is approximated to be 0.266 nm/mW based on the linear fit in Fig. 6.18c. We measured the 1dB thermal bandwidth of the structure to be about 40 kHz and 3dB thermal bandwidth to be about 350 kHz by tuning the input single-λ laser to the resonance and modulating the DC voltage by a small signal sinusoidal voltage and monitoring the peak-to-peak amplitude of the sinusoidal optical power at the drop port. Based on Eq. (6.57) the contours of 1dB and 3dB thermal bandwidth are plotted in Fig. 6.19a as a function of f_0 and ν. The intersection of these two contours yields f_0 = 723.122 kHz and ν = 203. A comparison of the measured
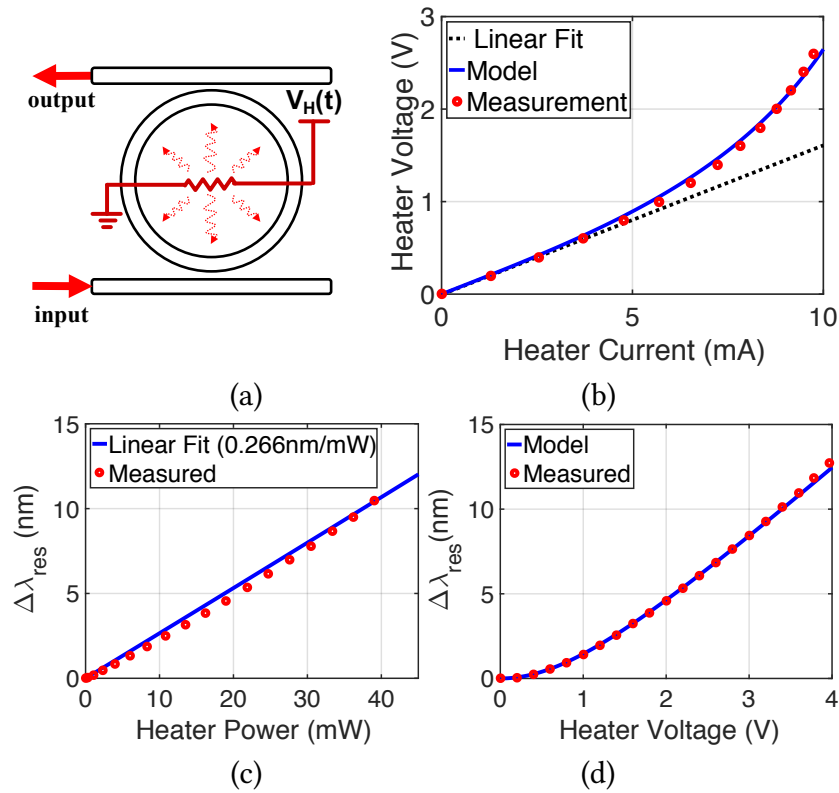
Figure 6.18: (a) Schematic of "add-drop" configuration with micro-heater. (b) Measured DC V-I characteristic of the heater. (c) Measured stationary thermo-optic response of the microring as a function of heater power. The heating efficiency is about 0.266 nm/mW. (d) Measured stationary thermo-optic response of the microring as a function of heater voltage.

thermal frequency response (red circles) and the fitted model (solid blue curve) is plotted in Fig. 6.19b indicating a good agreement. Finally, by substituting the extracted parameters in Eq. (6.77), α_H is estimated to be 6.6×10^{-3} K^{-1} and the thermal sensitivity of the resonance is estimated to be 0.0729 nm/°K or 9.05 GHz/°K.

Next, we examined the PWM response of this structure. In order to experimentally determine the minimum PWM drive frequency, we injected a broadband laser source in the input port and monitored the spectral response of the drop port as the PWM frequency increased [see Fig. 6.20]. As an example, a PWM signal with 2V peak-to-peak and 50% duty cycle was applied. We tracked the resonance shift for upper and lower thermal ripple amplitudes by increasing the PWM frequency. The two branches are plotted in Fig. 6.21a. When the frequency of the PWM drive is low, the upper and lower ripples are close to
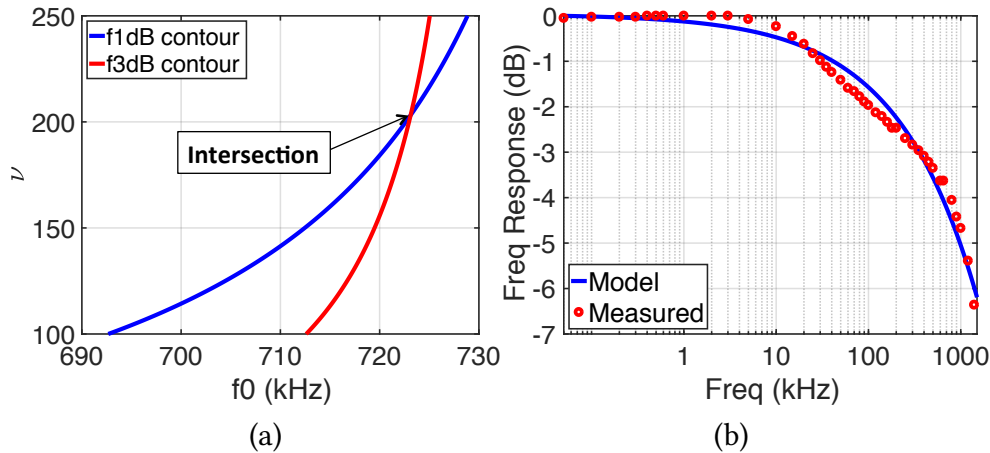
Figure 6.19: (a) Contours of 1 dB and 3dB thermal bandwidth to extract f_0 and ν. (b) Comparison of measured thermal frequency response and fitted model.
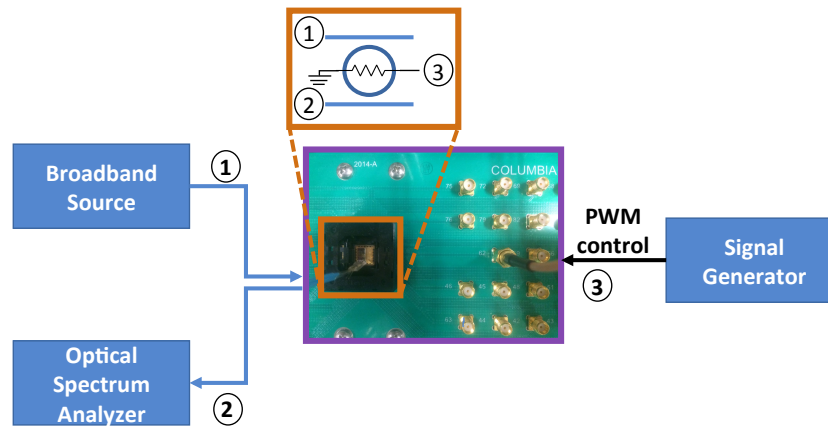
Figure 6.20: Experimental setup for PWM excitation of integrated microheater. The SiP chip is mounted on top of a testing socket that is soldered on a printed circuit board. A broad-band source is injected and the response is monitored with a spectrum analyzer through grating couplers. PWM control signal to the heater is applied through SMA connectors on the board [66].

0V and 2V; hence the shift of the resonance is close to its DC response. This is shown in Fig. 6.21b by markers "A" and "B". As the frequency of the PWM increases, the two branches get closer and closer to each other in the wavelength domain. This corresponds to a smaller peak-to-peak ripple in the time domain as shown in Fig. 6.21c by markers "C" and "D". Finally, at about 4 MHz of the drive frequency the two branches collide and a completely rectified thermo-optic response is achieved.

(a)

(b)   (c)
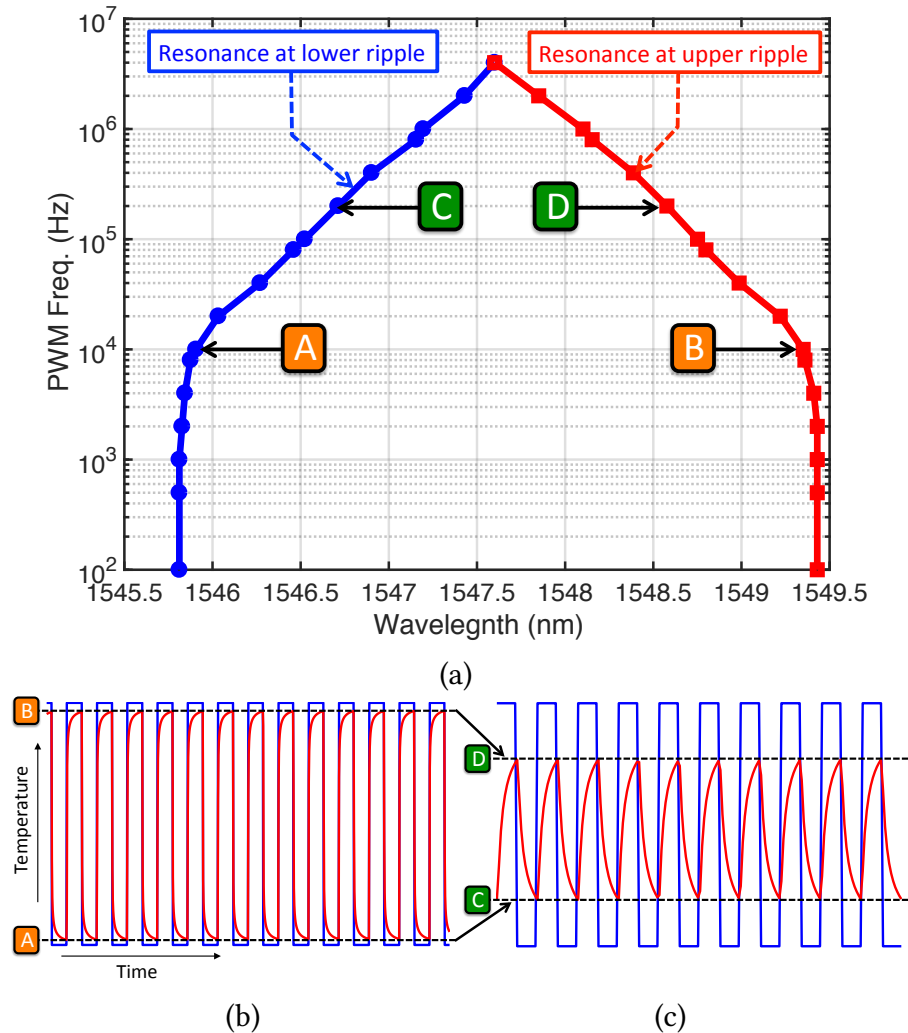
Figure 6.21: (a) Experimental measurement of the steady-state resonance wavelength of the add-drop ring resonator for upper and lower PWM ripple levels with 2V peak-to-peak and 50% duty cycle. The increase of the PWM frequency will bring the two branches closer together until they collide at about 4 MHZ. This indicates an experimental way to determine the minimum PWM frequency for a stable thermal response. (b) Large ripple levels due to the low frequency of PWM drive. (c) Small ripples due to the high frequency of PWM drive.

6.9   Summary of the Chapter

In this chapter, a compact modeling approach for characterizing the stationary and transient thermo-optic responses of microring resonators was presented. We showed that both temperature and electrical current of the resistive heater have a nonlinear dependence on the voltage applied to the heater. We also showed that our analytic models for

DC response and frequency response agree well with our experimental measurements.

A key parameter discussed in this chapter was the thermal impulse response of the heater-waveguide system that relates the change of temperature of the heater to the change of temperature of the waveguide. We showed how the 3dB thermal bandwidth of the system depends on physical parameters such as the distance between the heater and waveguide (inverse square relation), and we verified the good fit of the frequency model with experimental results in both amplitude and phase.

Finally, we characterized the thermal rectifying behavior of the microring systems for PWM excitation of the heater. The peak-to-peak amplitude of the optical ripples were estimated based on the first harmonic of the thermo-optic response. We showed that our estimations for the PWM response agree with the measurements in both the time domain and wavelength domain.

The results of this chapter can be utilized to estimated the performance of integrated heaters in terms of thermal efficiency and speed, especially when metallic heaters are concerned. Furthermore, in the case of PWM signaling for thermal tuning of microring resonators [210] or switching of Mach-Zehnder interferometers [119], the required amplitude and frequency of the signal can be predetermined based on our proposed model for thermal frequency response of a heater-waveguide system.

Chapter 7

Crosstalk Penalty in Microring-Based Silicon Photonic

Interconnects

Abstract– *In this chapter, We examine inter-channel and intra-channel crosstalk power penalties between WDM channels for microring-based silicon photonic interconnects. We first propose a model that relates the crosstalk power penalty to the interfering signals power, the extinction ratio of the non-return-to-zero on-off keying (NRZ-OOK) modulated "victim" channel, and finally the bit-error-ratio (BER) that the power penalty is referenced to. As for inter-channel crosstalk, the proposed model agrees well with the experimental measurements. We leverage this model to quantify crosstalk induced power penalties in a microring based WDM receiver. We also propose an optimization procedure to equilibrate the power penalty across channels. We then compare our model with intra-channel crosstalk measurements, where two NRZ channels are at the same wavelength and are simultaneously routed to different paths by two cascaded ring resonators. We remark that intra-channel crosstalk is very sensitive to the data rate of NRZ channels.*

In high-capacity WDM links, channels tend to correlate and imprint negative effects on each other [204]. Various factors such as strong spectral overlap of adjacent channels and polarization mode dispersion [204, 55, 127] can contribute to such impairments of an optical link. These crosstalk effects are especially prominent where microrings are employed to add or drop individual channels to or from a WDM signal (Inter-channel crosstalk) [145, 80, 53, 170], or to actively switch and spatially route several channels of the same wavelength (intra-channel crosstalk) [81, 83, 172]. In all of these situations,
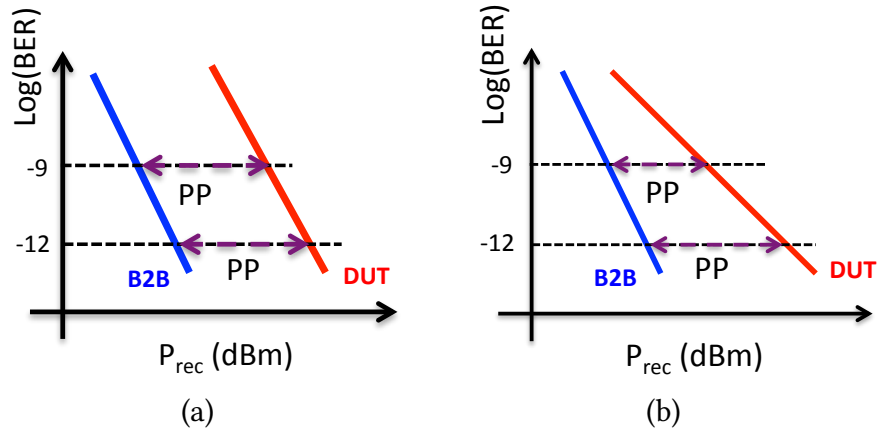
Figure 7.1: Examples of the BER power penalty extraction from the measurements. (a) The back-to-back case (B2B) and the device under test (DUT) case are almost parallel. (b) B2B case and the DUT case are not in parallel.

optical crosstalk tends to distort and close the eye diagram of the non-return-to-zero on-off keying (NRZ-OOK) modulated channel and affect the performance of the link in terms of bit error ratio (BER).

The effect of the crosstalk in an optical link can be experimentally quantified by comparing BER curves measured with and without the crosstalk at the receiver. The difference between the two curves then reveals the power penalty associated with the crosstalk effect [81, 80, 83]. Typically, if the signal impairment is related to optical loss or spectral filtering the back-to-back (B2B) measurement and the device under test (DUT) measurement are almost parallel curves indicating the nature of power penalty is independent of the noise in the system. This is shown in Fig. 7.1a. However, crosstalk impairments contribute to the noise, therefore the power penalty varies from one BER level to another. This is shown in Fig. 7.1b where the two BER lines are not in parallel.

In general, predicting the impact of the crosstalk noise through signal modeling and statistical analysis is less straightforward, and although several studies in the past have proposed different ways of estimating the inter-channel crosstalk power penalties [55, 164, 184, 54, 152], none are well suited for MRR based WDM links with high-speed modulation and/or high channel density. Recent analytical studies on the crosstalk in MRR-

based links have shown the potential impact of the crosstalk effect on the performance of the photonic networks-on-chips (NoCs) [56, 57, 207], but these studies disregard the role of the modulation speed on the severity of the crosstalk effect.

Being able to correctly estimate crosstalk is critical not only to ensure proper functioning of projected WDM links, but also to guarantee low power consumption, since unaccounted crosstalk can oblige the initial laser power to rise, which is clearly detrimental to power consumption. Therefore, our objective in this paper is to thoroughly capture the effect MRRs (and their Lorentzian filter response) have on crosstalk, and to express these effects in power penalty terms.

One difficulty in detailed mathematical modeling of crosstalk effects stems from the fact that, at the receiver, the *main* optical signal (the victim) and all crosstalk signals (the aggressors) go through a square-law detector; a nonlinear operation. As a result, photocurrents corresponding to the main signal, and the crosstalk fluctuations may experience a rather strong cross-correlation with each other, even if the crosstalk and main signals are not correlated at any point along the optical path. Therefore, a thorough analysis should not only include the impact of the "aggressor" channel on a "victim" channel, but also the impact of the intermodulation products between any pair of channels on the "victim" channel [205, 152]. This demands for estimation of joint probability density function (pdf) of all crosstalk channels.

Provided that a rigorous and computationally intensive statistical analysis may not be of much practical use for system designers, especially when dealing with dense WDM links, we aim in this chapter at introducing a lighter weight formulation to estimate the impact of crosstalk on microring-based links. Our approach is based on recognizing the fluctuations introduced by crosstalk on the victim channel as Gaussian noise that accompanies the NRZ-OOK channel, and determining the BER as a function of the signal-to-noise ratio. Although crosstalk effects for optical filters have been studied for Gaussian-like transmission spectrum [183], the spectral response of microring filters tends to devi-

ate from a perfect Gaussian shape and conforms to a Lorentzian shape. We specifically include this behavior in our study. The model proposed here shows good agreement with various measurements of the inter-channel crosstalk power penalty for microring resonators. We also use our analytical model to investigate how the speed of modulation and the bandwidth of ring resonators contribute to the crosstalk power penalty.

7.1 Inter-Channel and Intra-Channel Crosstalk

Multiplexing/Demultiplexing (Mux/DeMux) of optical channels can be done in the wavelength domain or space dimensions. The former is the concept of WDM and the latter is related to the spatial routing of optical wavelengths (spatial switching or space division multiplexing). Figure 7.2a shows an example of wavelength domain DeMux, where an array of microrings performs demultiplexing on the WDM input. Each ring drops a specific wavelength (channel), but also collects some residual power from neighboring channels. This is both due to the Lorentzian shaped filtering response of the ring whose transitions are not sharp, and to the potentially very broad spectrum of modulated signals. An example of three channels is shown in Fig. 7.3a where the spectrum of the middle channel (characteristic of NRZ-OOK modulations) is almost fully dropped while the two adjacent channels have some residual power leaking into the drop port as well. Since this type of residual power is, in this case, not at the same wavelength as the main dropped channel, we refer to it as the *inter-channel* crosstalk. It is worth noting that the part of this crosstalk resulting from imperfect filtering of ring resonators can be filtered out later, before the photodiode, with an optical filter offering sharper response, or after the photodiode with an electrical filter.

On the other hand, in space multiplexing, distinct segments of an optical transmission medium are used by two or more channels. This results in *intra-channel* crosstalk if the same wavelength is used over neighboring segments (in addition to *inter-channel*
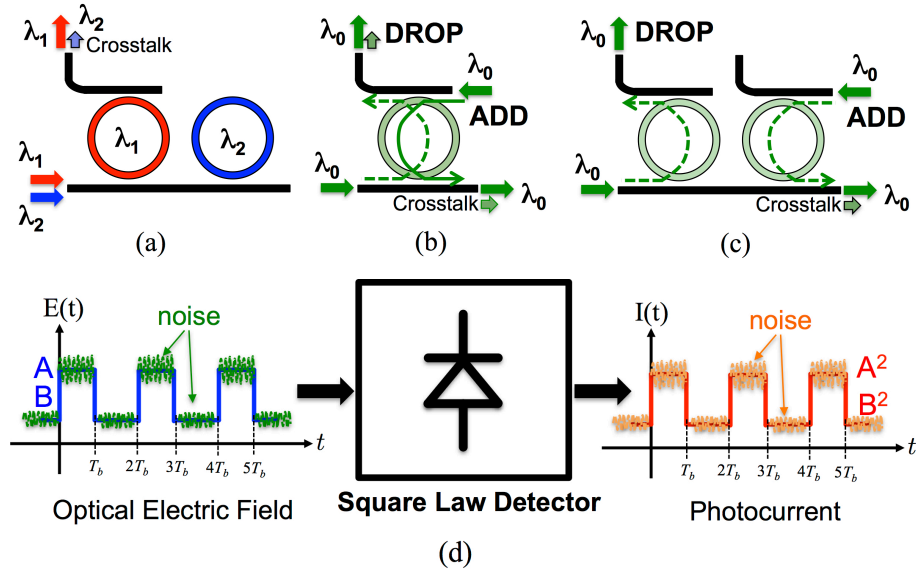
Figure 7.2: (a) Schematic of *inter-channel* crosstalk between two NRZ-OOK WDM channels. The residual power of the λ_2 channel causes crosstalk power penalty at the drop port of the λ_1 ring. (b) and (c) schematic of *intra-channel* crosstalk between two identical NRZ-OOK channels when ADD and DROP operations are performed. The leakage of the DROP channel to the through port of the ADD channel causes power penalty. (d) Conversion of the optical field into the photocurrent by an ideal square-law detector.

crosstalk if multiple wavelengths are used). Situations involving intra-channel crosstalk are illustrated in Fig. 7.2b, where one single ring routes two channels simultaneously [99] and in Fig. 7.2c where two separate rings route each channel separately [81]. Ideally, there should be no leakage of power from each channel to the other in Fig. 7.2c, but the non-idealistic Lorentzian-shaped filtering response of the ring makes some leakage of power inevitable. The power spectrum of two channels and their leaked distorted spectrum at both through and drop ports are schematically shown in Fig. 7.3b. To mathematically estimate the power penalty associated with this leakage, we compare the expressions of BER for two cases, without and with the crosstalk noise. The power penalty is the difference in the optical power required to compensate for the crosstalk effect and maintain the same BER level.

(a)

(b)
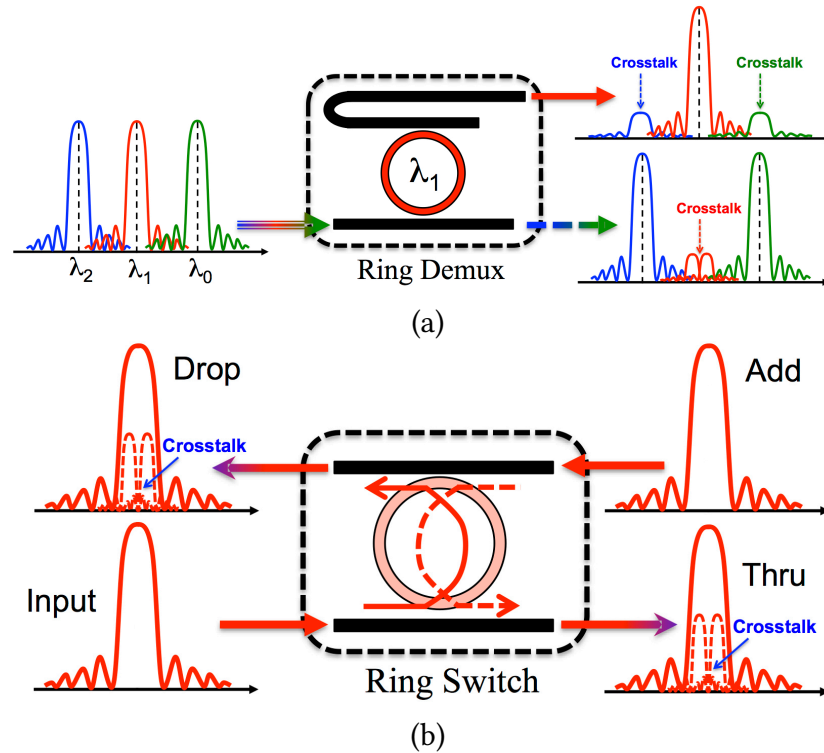
Figure 7.3: (a) Illustration of inter-channel crosstalk with three NRZ-OOK channels. The residual powers in both the drop port and the through port are not on the same wavelength as the dropped or passed channels. (b) Illustration of the intra-channel crosstalk where some power leaks due to the imperfect spatial switching functionality of the ring. Since the spectrum of the leaked channel is severely distorted by the ring, the residual power can be considered effectively as noise.

7.2 Estimation of Inter-Channel Crosstalk

Analytical Modeling Approach

We consider a WDM receiver composed of an array of ring resonators. Each ring acts as a demultiplexer for a single wavelength. Once demultiplexed, the main signal and all the inter-channel crosstalk signals directly go through the square-law detector (shown in Fig. 7.2d) and add to the photocurrent. The total electric field at the input of the first ring in Fig. 7.2a from the main channel, named $E_0(t)$, and all the crosstalk channels, $E_i(t)$ [i = 1, 2, ..., N], can be written as $E_{in}(t) = E_0(t) \exp(j\omega_0 t) + E_1(t) \exp(j\omega_1 t + \phi_1(t)) + E_2(t) \exp(j\omega_2 t + \phi_2(t)) + \cdots + E_N(t) \exp(j\omega_N t + \phi_N(t))$ where $\phi_i(t)$ accounts for possible

temporal phase difference between i^{th} crosstalk channel and the main channel [152]. The total average spectral optical power (and the photocurrent) at the drop port of the ring tuned to the main channel (ω_0) is

$$
\begin{aligned}
P_{drop}^{av} \approx P_0^{av} + \sum_{i=1}^{N} \rho_i P_i^{av} + \\
2\sqrt{P_0^{av}} \sum_{i=1}^{N} \sqrt{\rho_i \ P_i^{av}} < \cos(\Delta\omega_i t + \phi_i(t)) >_t + \\
2\sum_{i>j} \sqrt{\rho_i \ \rho_j \ P_i^{av} \ P_j^{av}} < \cos(\Delta\omega_{ij}t + \Delta\phi_{ij}(t)) >_t
\end{aligned} \tag{7.1}
$$

where $< ... >_t$ means averaging over time. The first term is the spectral power of the main ("victim") channel. The second term is the sum of spectral powers from all the crosstalk channels. ρ_i's refer to the average attenuation of the Lorentzian tail of the drop path of the ring for the ith NRZ channel, which can be calculated based on the power spectral density of the NRZ modulation [13]. The third term in Eq. (7.1) accounts for the beating of each crosstalk channel against the main channel and the last term refers to the beating of crosstalk channels against each other. Since the photodiode directly converts optical power into electrical current, the high frequency variations of the input light will be suppressed by the limited bandwidth of the photodiode (PIN diode) and the electronic front-end of the receiver. Hence, for the case of inter-channel crosstalk, if the spacing between channels is large enough, both $\Delta\omega_i = \omega_i - \omega_0$ and $\Delta\omega_{ij} = \omega_i - \omega_j$ will be large enough (e.g. > 100 GHz) and the time average of $\cos(...)$ function results in zero. The only remaining term that reflects the crosstalk effect is the second term, meaning that as long as the frequency spectra of the base-band signals fall within the electrical bandwidth of the receiver, each crosstalk channel directly contributes some optical power at the receiver. Since we are considering the inter-channel crosstalk as a negative side effect of imperfect filtering of ring resonators, in the simplest modeling approach this crosstalk power is considered to entirely add up to the thermal and electrical shot noise fluctuations

at the receiver. It is also assumed that the square-law detector does not contribute any noise to the signals.

The thermodynamic processes that are involved in generating noise at the receiver are random in nature (hence thermal noise is presumably characterized by an additive Gaussian noise). The crosstalk mechanism, in contrast, is not purely random due to the deterministic response of the ring filter and photodetector. However, if the NRZ-OOK is generated by a sequence of purely random '0' and '1' bits with equal probabilities, and the strength of the crosstalk power is low enough (i.e. all ρ_i's are small enough in Eq. (7.1)), we may regard the inter-channel crosstalk power as a quasi-random phenomenon with a relatively white (wideband) behavior. By making this assumption, i.e. considering simple Gaussian distribution for the *total* (thermal + shot + crosstalk) noise at the receiver, the BER of the NRZ-OOK modulation is now given by the complementary error function whose argument is called the Q-parameter [151]. The expressions for Q-parameter without the crosstalk effect (Q_{BER}) and with the crosstalk (Q'_{BER}) for a receiver with optimal decision threshold that minimizes the BER are given by

$$Q_{\text{BER}} = R\frac{P_A - P_B}{\sigma_{n,A} + \sigma_{n,B}}, \ Q'_{\text{BER}} = R\frac{P_A - P_B}{\sigma_{n,A} + \sigma_{n,B} + \sigma_{n,XTalk}} \tag{7.2}$$

where P_A and P_B are the optical power of high and low levels of the modulated electric field, R is the responsivity of the photodiode, $\sigma_{n,A}$ and $\sigma_{n,B}$ are the standard deviations of the Gaussian distribution of the high and low photocurrents, and $\sigma_{n,crosstalk}$ is the standard deviation of the crosstalk noise current. With these two expressions, the Q-penalty (or power penalty) resulting from the crosstalk, defined as the extra power factor required to maintain the same BER, can be written as PP $= -10\log_{10}(Q'_{\text{BER}}/Q_{\text{BER}})$, where

$$\frac{Q'_{\text{BER}}}{Q_{\text{BER}}} = 1 / \left(1 + \frac{\sigma_{n,XTalk}}{\sigma_{n,A} + \sigma_{n,B}}\right) \approx 1 - \frac{\sigma_{n,XTalk}}{\sigma_{n,A} + \sigma_{n,B}}. \tag{7.3}$$

Using Eq. (7.2), the original noise at the receiver (in the absence of crosstalk) is given by

$$\sigma_{n,A} + \sigma_{n,B} \approx R \, \frac{r-1}{r+1} \times \frac{P_A + P_B}{Q_{\mathrm{BER}}}, \qquad (7.4)$$

where r is the extinction ratio of the NRZ-OOK modulation defined as $r = P_A/P_B$ (note that $r > 1$). To evaluate $\sigma_{n,crosstalk}$, we consider the spectral power of the NRZ crosstalk channel that we denote by $P^{\mathrm{av}}_{crosstalk}$. Since the perturbing photocurrent generated by the crosstalk power is proportional to the spectral power of the NRZ crosstalk channel, i.e. $I_{crosstalk} \approx R P^{\mathrm{av}}_{crosstalk}$, and the variance (strength) of the crosstalk electrical noise current at the receiver is given by $\sigma^2_{n,crosstalk} = <I^2_{crosstalk}>$, we assume that $\sigma_{n,crosstalk} \approx R P^{\mathrm{av}}_{crosstalk}$. By combining these relations with Eq. (7.3) and replacing $(P_A + P_B) = 2P^{\mathrm{av}}_{NRZ}$, power penalty takes the following form:

$$\mathrm{PP}_{XTalk} \approx -10 \, \log_{10} \left(1 - \frac{1}{2} Q_{\mathrm{BER}} \, \frac{P^{av}_{XTalk}}{P^{av}_{NRZ}} \, \frac{r+1}{r-1} \right). \qquad (7.5)$$

In this equation the power penalty mainly depends on the ratio $P^{av}_{crosstalk}/P^{av}_{NRZ}$, which is indicative of the relative strength of the spectral crosstalk power that reaches the victim channel. Furthermore, the power penalty also depends on the Q_{BER} which takes the approximate value of 6 for BER = 10^{-9} and 7 for BER = 10^{-12}. For the same crosstalk power, dependence on Q_{BER} means that power penalty is higher for a lower BER (e.g. power penalty at 10^{-12} is larger than at 10^{-9}). This is congruent with the observation that in the experimental measurements the BER curves are not in parallel (i.e. power penalty varies if the reference BER is changed) [81, 80, 83]. Finally, the dependence of power penalty on the extinction ratio and the fact that $(r+1)/(r-1) > 1$ shows that NRZ channels with lower extinction ratio (i.e. a more closed eye) are more prone to crosstalk effects and exhibit higher penalties.

In order to capture the full impact of the inter-channel crosstalk in a WDM receiver based on a demultiplexing array of rings, we first need to evaluate the strength of all

signals contributing to $P_{crosstalk}$, i.e. the second term of Eq. (7.1). Figure 7.4a illustrates the general case of an array of N rings employed at the receiver of a silicon photonic link [73, 215]. The input is a WDM light with N wavelengths such that $\lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_N$ (or equivalently $f_1 < f_2 < f_3 < \cdots < f_N$). Using the approximate Lorentzian lineshape of the through and drop paths of each ring, the fraction of crosstalk power present at the drop port of the ith ring from the spectrum of the jth crosstalk channel is given by,

$$\Gamma_{i,j} = \int_{-\infty}^{\infty} \frac{\text{sinc}^2(F)\ dF}{1 + \left(\frac{F+(j-i)F_\Delta}{\xi_i}\right)^2} \left[\prod_{k=1}^{i-1} \frac{\left(\frac{F+(j-k)F_\Delta}{\xi_k}\right)^2}{1 + \left(\frac{F+(j-k)F_\Delta}{\xi_k}\right)^2}\right], \tag{7.6}$$

where the product in brackets signifies the filtering effect (limited bandwidth) of the through path of cascaded rings from 1 to $i-1$. Note that for the first ring, this product does not exist and the value of the term inside the bracket should be set to unity. This integral is evaluated numerically for a large array of rings. In Eq. (7.6), $F = f/r_b$ is the frequency normalized to the bit rate, $F_\Delta = f_\Delta/r_b$ is the frequency spacing between adjacent channels normalized to the bit rate (hence $f_i - f_j = (i-j)f_\Delta$ for a uniform grid), and $\xi_i = \text{FWHM}_i/(2r_b)$ is the 3dB bandwidth of the ith ring normalized to the bandwidth of the NRZ modulation (assuming all the incoming channels have the same bit rate). Finally, the total crosstalk power at the ith drop port is

$$(P^{av}_{XTalk})_{i^{\text{th}}\ Ring} \approx P^{av}_{NRZ} \times \sum_{j=1, j\neq i}^{N} \Gamma_{i,j}\ , \tag{7.7}$$

which is then introduced into Eq. (7.5). *The appearance of the r_b in $\Gamma_{i,j}$ coefficients underscores the impact of the modulation rate in our crosstalk modeling approach*; a fact that has been generally disregarded in previous studies of crosstalk for micro-rings [56, 57, 207].

(a)

(b)
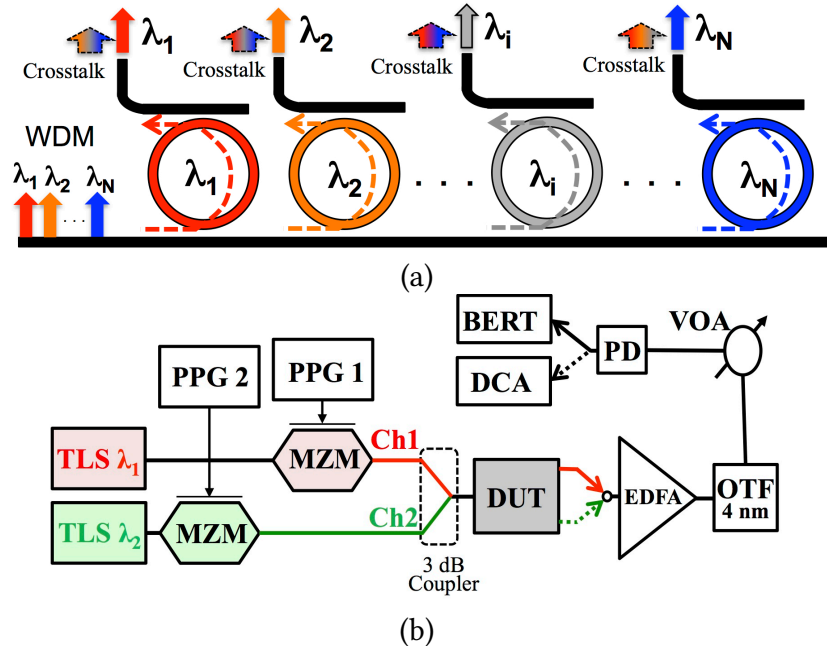
Figure 7.4: (a) An array of microring demultiplexers at the receiver of a silicon photonic link. (b) Experimental setup for measuring the BER curves of inter-channel crosstalk between two cascaded rings with various channel spacings. Tunable laser sources (TSL) are used to control the channel spacing.

Comparison with Measurements

Now we compare our model for inter-channel crosstalk against the experimental data. The experiment is carried out with only two cascaded rings (as illustrated in Fig. 7.2a). The radius of the both fabricated silicon rings is 8 μm in a silicon-on-insulator platform with a waveguide height of 220 nm and a slab thickness of 90 nm[1]. Wavelength tuning is accomplished via thermal tuning by integrating doped resistive regions in the silicon ring (internal heater). The measured spectra of the fabricated devices have been presented in [81, 80, 83], showing a full-width-at-half-maximum bandwidth of approximately 30 GHz. The free spectral range (FSR) of the rings was measured to be around 12.4 nm. The experimental setup for the inter-channel crosstalk is presented in Fig. 7.4b. Two separate pulse pattern generators (PPGs) with $2^{31} - 1$ pseudo random binary sequences

---

[1]This device was designed, fabricated, tested and measured at the University of British Columbia, Canada.

at 25Gb/s were employed to drive two LiNbO$_3$ Mach-Zehnder modulators (MZMs) for NRZ-OOK modulation. The optical power in both channels was kept on the same level by adjusting the output power of the two tunable lasers. The modulated signals were combined using a simple 3dB coupler. This WDM signal creation results in no tangible inter-modulation crosstalk. The optical signal at the drop port of the ring (device-under-test, DUT) was amplified and then filtered using an optical tunable filter (OTF). A variable optical attenuator (VOA) controlled the received power at the photo-detector (PD). The bit-error-rate (BER) of the drop port of the first ring (Ch1) and the second ring (Ch2) was measured for different channel spacings. A few of the measured BER curves are presented in Fig. 7.5a (channel spacings 112.5 Ghz, 75 Ghz, 50 Ghz, 37.5 Ghz). Here, Ch1 is the "victim" and Ch2 is considered the "aggressor". To support the assumption of crosstalk behaving as Gaussian noise at the receiver, we fit a Gaussian BER curve (BER $\approx 1/2 \ \mathrm{erfc}(Q_{\mathrm{BER}}/\sqrt{2})$) to the measured data. The fitted curves are presented as dotted black curves in Fig. 7.5a where they agree well with the measurements (colored markers). A graphical interface was developed to perform the power penalty extraction. An example of this interface (for the case with the crosstalk channel at 75 Ghz spacing) is shown in Fig. 7.6.

For a given channel spacing, the power penalty is extracted by performing a Gaussian extrapolation of the measured BER for Ch1 and finding the received power that corresponds to a BER of 10^{-12} ($Q_{\mathrm{BER}} \approx 7$) with and without the presence of Ch2 (i.e. the aggressor). This will ensure that the estimated power penalty at the drop port of each ring is mainly due to the crosstalk and not due to any spectral distortions of the victim channel (Ch1) [13, 20]. The results are presented in Fig. 7.5b as red circles. The solid blue curve represents the calculated power penalty from the combination of Eq. (7.5), (7.6), and (7.7). The extinction ratio of the referenced NRZ-OOK channel at the drop port of the rings was measured to be about 9 dB (hence $r \approx 8$ and $(r + 1)/(r - 1) \approx 1.23$). A good agreement is observed between the measurements and the model, indicating that a
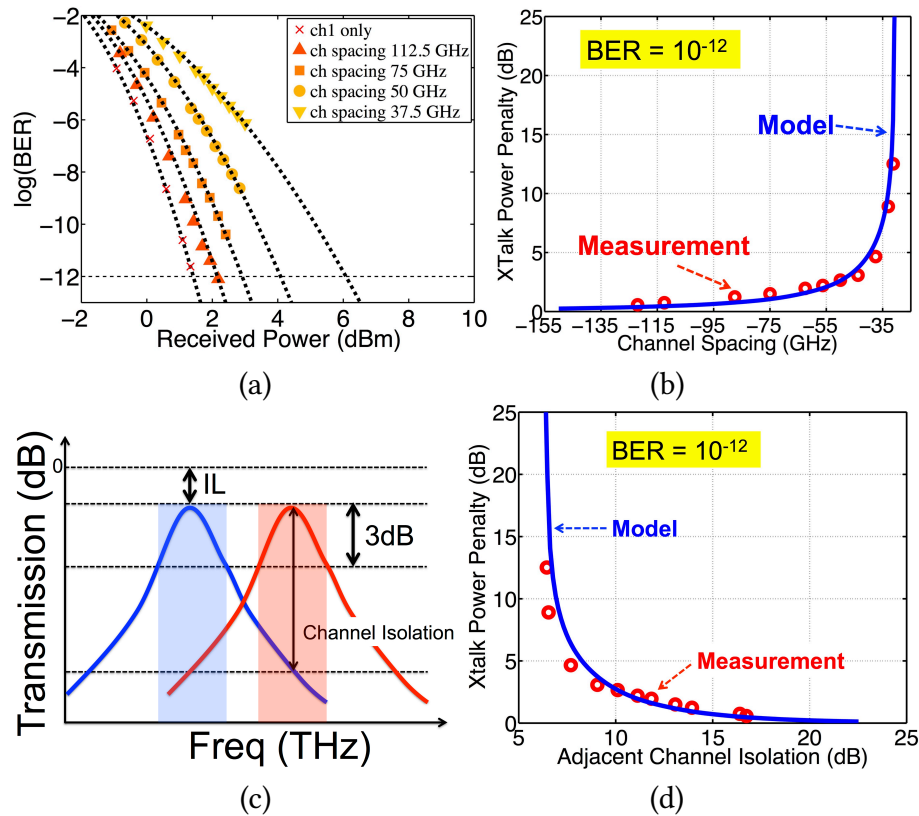
Figure 7.5: (a) Measured BER curves (markers) and Gaussian fitting (dotted curves) for inter-channel crosstalk between two WDM channels. (b) Comparison of the measurements (circles) and the model (solid) for inter-channel crosstalk. (c) Illustration of the drop spectrum of the two DeMux rings with highlighted 3dB drop bandwidths. (d) Comparison of the model and measurements for the crosstalk power penalty as a function of the adjacent channel isolation due to the Lorentzian tail of the spectrum of the rings.

minimum channel spacing of 80 GHz is essential to keep the crosstalk penalty for the first ring below 1 dB (for this 25 Gb/s case).

Since the two ring filters have the same Lorentzian spectra (although tuned to different wavelengths), another possible way of expressing the power penalty is based on the attenuation of the crosstalk channel due the Lorentzian tail of the drop path response, which is typically referred to as the channel suppression or channel isolation. A schematic plot of the Lorentzian responses of the two *isolated* ring filters tuned to Ch1 and Ch2 is shown in Fig. 7.5c indicating the FWHM and channel isolation. The Lorentzian response
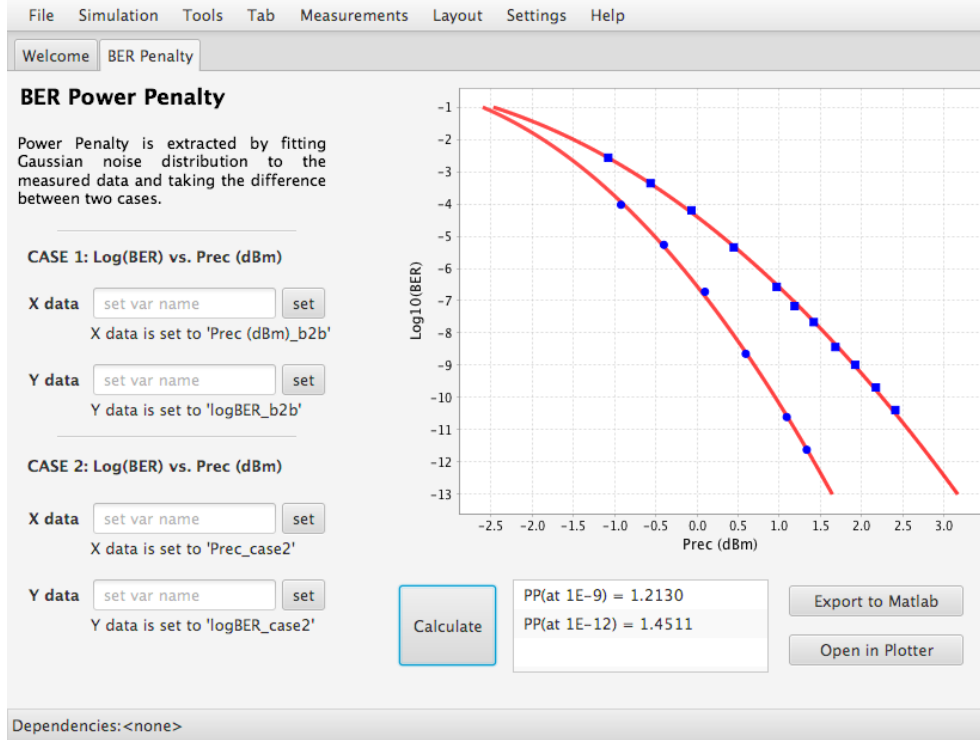
Figure 7.6: A developed interface for extracting BER power penalty from the isolated channel and the crosstalk case.

can be written as

$$T(f) \approx \frac{q_0}{1 + \left(\frac{2}{\text{FWHM}}(f - f_0)\right)^2} \tag{7.8}$$

where q_0 corresponds to the insertion loss of the drop path [IL = $-10\log_{10}(q_0)$] at the resonance frequency (f_0). We calculate the crosstalk power penalty versus channel isolation both from the model and the actual measured spectra of the ring resonators. The results are plotted in Fig. 7.5d where a good agreement is again observed between them. A minimum channel spacing of 80 GHz ($\sim$1 dB power penalty) in this case corresponds to a minimum channel suppression of 15 dB. This value mainly depends on the 3dB optical bandwidth of the microring resonators.
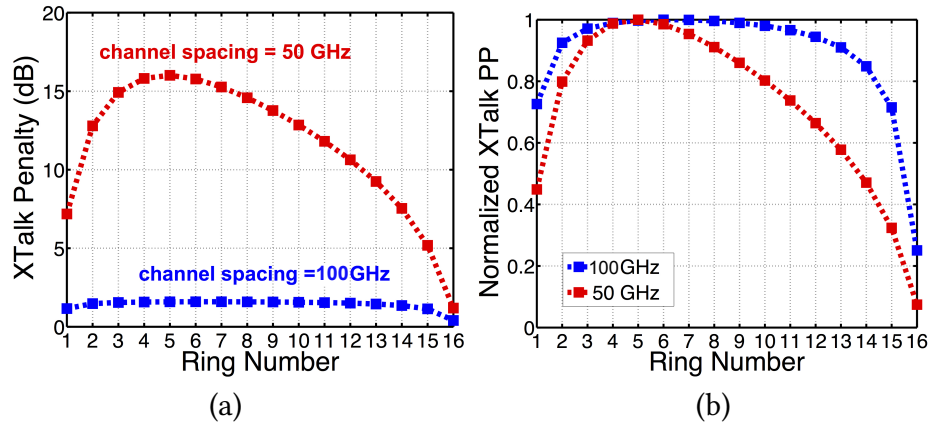
Figure 7.7: (a) Estimation of crosstalk power penalty for an array of 16 cascaded rings. Each ring has a bandwidth of 30 GHz and the incoming WDM channels are at 25 Gb/s. It is clear that the spacing between channels plays an important role in the crosstalk power penalty. (b) Same as (a) but each curve is normalized to its maximum value.

7.3   Optimization of Demultiplexing Array

Crosstalk Estimation

We now show how the crosstalk model can be used for optimizing a ring-based demultiplexing array. In the case of two cascaded rings, the first ring always experiences a higher crosstalk penalty than the second ring. However, our model shows that this is not always the case if the number of channels is large. As an example, we use the parameters of the devices under test in the previous section to investigate an array of 16 cascaded ring resonators. As discussed in [81], the choice of 16 wavelengths comes from the FSR of the rings (12.4 nm) and the requirement of about 100 GHz spacing between channels to keep the crosstalk penalty at a low level (< 0.5 dB).

Fig. 7.7a shows the estimated crosstalk penalty at the drop port of each ring for two channel spacings (50 GHz and 100 GHz). As can be clearly observed, neither the first nor the last ring experiences the worst-case of crosstalk. The general shape can be explained by two facts:

1. Channels located on the sides of the spectrum are less affected as they receive

222

crosstalk from one side only. This premium then vanishes progressively as more central channels are considered.

2. Because signals are progressively dropped, crosstalk penalty is lower for the last channels than for the first ones.

We also note that for 100 GHz channel spacing, the crosstalk distribution over the entire array is almost flat and is maintained below 2 dB (max 1.6 dB). Such flat distribution is desirable since it guarantees a uniform BER performance for all the channels. In contrast, when a much narrower channel spacing of 50 GHz is used, the central channels experience a significantly higher level of crosstalk compared to the side channels, which leads to a non-flat distribution of crosstalk. This is illustrated in Fig. 7.7b, where each curve is normalized to its maximum value (dB/dB). It is worth noting that for this particular example, the model indicates a minimum spacing of 120 GHz ($\sim$1 nm) between channels in order to keep the maximum crosstalk penalty below 1 dB. This is a more stringent requirement compared to the case of two rings. More generally, this analysis shows how our newly proposed model can be used to investigate the effect of the various design parameters on the scalability of the MRR link in terms of the number of wavelengths that can be utilized while maintaining the power penalty below a certain level.

Crosstalk Equilibration

The observations of Fig. 7.7 led us to examine the effect of reshuffling the resonance wavelengths of the rings in the array shown in Fig. 7.4a, i.e. the ith ring is tuned to the incoming λ_j instead of λ_i. An example is illustrated in Fig. 7.8 where 6 channels are present at the receiver (channel spacing = 100 GHz). The worst-case (obtained through exhaustive enumeration) of the crosstalk happens if the first ring is tuned to the middle channel, i.e. the first ring maximally captures the signals of not-yet-filtered neighboring channels. In this case, the distribution of the crosstalk penalty has a maximum of 2 dB at
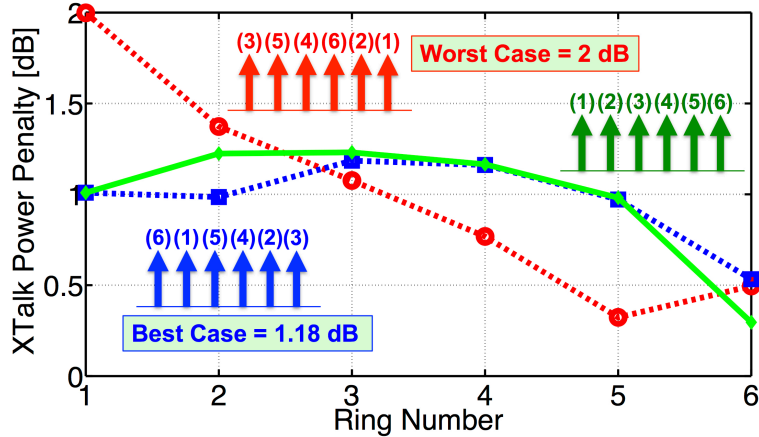
Figure 7.8: An example of how reshuffling the resonances of the ring array will help to reduce/flatten the crosstalk penalty distribution over the entire array of 6 rings. For the red curve, the first ring is tuned to λ_3, second ring is tuned to λ_5 and so forth. For the blue curve, the first ring is tuned to λ_6, the second ring is tuned to λ_1 and so forth. The solid green curve is for the default ordered case. The spacing between channels in all cases is set to 100 GHz.

the first ring and a minimum of 0.3 dB at the fifth ring. On the other hand, and logically, if the resonances are reshuffled such that channels at the extremities of the Lorentzian spectrum of the ring are filtered first as shown for the blue case in the inset of Fig. 7.8, the distribution of the crosstalk penalty shows a flatter distribution than the red curve and an improvement of 0.8 dB penalty is achieved in terms of worst-case over the entire array (without changing the spacing between channels). A slight improvement is also achieved compared to the default ordered case (i.e. R1 $\rightarrow$ λ_1, R2 $\rightarrow$ λ_2, R3 $\rightarrow$ λ_3, R4 $\rightarrow$ λ_4, R5 $\rightarrow$ λ_5, R6 $\rightarrow$ λ_6) plotted as a solid green curve. This analysis can be also extended into investigating the benefits of a *non-uniform spacing (a flex grid)* for the demux array.

7.4   Estimation of Intra-Channel Crosstalk

In this section, we compare our crosstalk model based on Eq. (7.5) with the intra-channel crosstalk data recorded by using two cascaded rings (Fig. 7.2c) that simultaneously add and drop (switch) signals at the same wavelength. The block diagram of the experimental

setup for recording eye diagram and BER measurements for the intra-channel crosstalk is presented in Fig. 7.9a. Eye diagrams obtained for the intra-channel crosstalk in a single-ring structure (Fig. 7.2b) and two cascaded rings (Fig. 7.2c) at 20 Gb/s data rate, are shown in Figs. 7.9b and 7.9c, respectively. As expected, single-ring switch exhibits far severe crosstalk than the two cascaded rings. The BER curves for the cascaded rings were measured for data rates from 10 Gb/s to 25 Gb/s as presented in Fig. 7.9d. For each data rate, we extracted the intra-channel crosstalk penalty for ADD-DROP operation by taking the difference between the extrapolated powers (with Gaussian fitting) required to maintain a BER of 10^{-12} for ADD-only (reference case) and simultaneous ADD-DROP operations. The results for this double-ring structure are plotted in Fig. 7.9e as red circles. To approximate the intra-channel crosstalk power penalty associated with a ring acting as a spatial switch, as shown in Fig. 7.2b and Fig. 7.2c, the power of the interfering signal(s) P_{XTalk}^{av} must be identified. This power is the spectral power that does not get coupled into the ring and remains in the waveguide. P_{XTalk}^{av} depends on two parameters: first, the rejection ratio of the through path of the ring resonator, which we denote by t_0 (e.g. t_0 = 0.01 for 20 dB rejection, and t_0 = 0 for critically coupled rings); second, the ratio of the 3-dB rejection bandwidth of ring resonator to the bandwidth of the NRZ-OOK modulation, which we denote by ξ = FWHM/$(2r_b)$, where r_b is the bit rate. With these two parameters, the spectral crosstalk power of the single-ring structure in Fig. 7.2b can be written as $P_{XTalk}^{av} \approx P_{NRZ}^{av}(1 - (1 - t_0) \times \gamma)$ where

$$\gamma = 1 - \frac{1 - \exp(-2\pi\xi)}{2\pi\xi}. \tag{7.9}$$

γ is a factor that reflects the filtering of the spectrum of an NRZ-OOK modulated channel by a Lorentzian filter [13]. By directly inserting Eq. (7.9) into Eq. (7.5), the crosstalk power penalty associated with BER = 10^{-12} for a single microring switch can be evaluated. Similar calculation can be performed to evaluate the crosstalk penalty of two cascaded

microrings. The result is presented in Fig. 7.9e by a solid blue curve. Although the model matches the measurements for up to 20 Gb/s (which corresponds to about 4 dB penalty), it underestimates the penalty at higher data rates. As can be seen from the measurements, the power penalty sharply goes up around 25 Gb/s data rate ($\sim$ 7 dB penalty) while the model predicts a threshold of 31 Gb/s for 7 dB penalty. One possible explanation for deviation of the model from the measurements of intra-channel crosstalk is that the model does not take the third and forth terms of Eq. (7.1) into account. These two terms describe the crosstalk in terms of the degree of coherence between the main signal and the crosstalk power. The reason is the unknown (or hard to predict) statistics of the coherence of the beating terms in Eq. (7.1). Considering full coherence between the crosstalk and the signal in Eq. (7.1) will lead to significant overestimation of the power penalty even for the low data rates.

Another way of modeling the intra-channel crosstalk penalty is based on the closure of the eye diagram (a worst-case analysis) [151]. In this case, instead of considering the crosstalk power as a noise mechanism, we estimate how the crosstalk signals distort the eye diagram of the NRZ-OOK modulation for '0' and '1' bits. The power penalty for the crosstalk is then given by

$$\text{PP}_{XTalk} \approx -10 \log_{10} \left(1 - 2 \sum_i \sqrt{\gamma_i} \right) \tag{7.10}$$

where γ_i is the crosstalk power ratio from the ith channel. This equation is calculated and plotted in Fig. 7.9e as a solid green curve. As a comparison, this model shows less consistency with the measurements than Eq. (7.5) in which the crosstalk power is treated as noise.

Both the eye-closure approach and our noise-based approach fail to correctly predict the measured crosstalk penalty at high data-rates. However, we observe that Eq. (7.5)
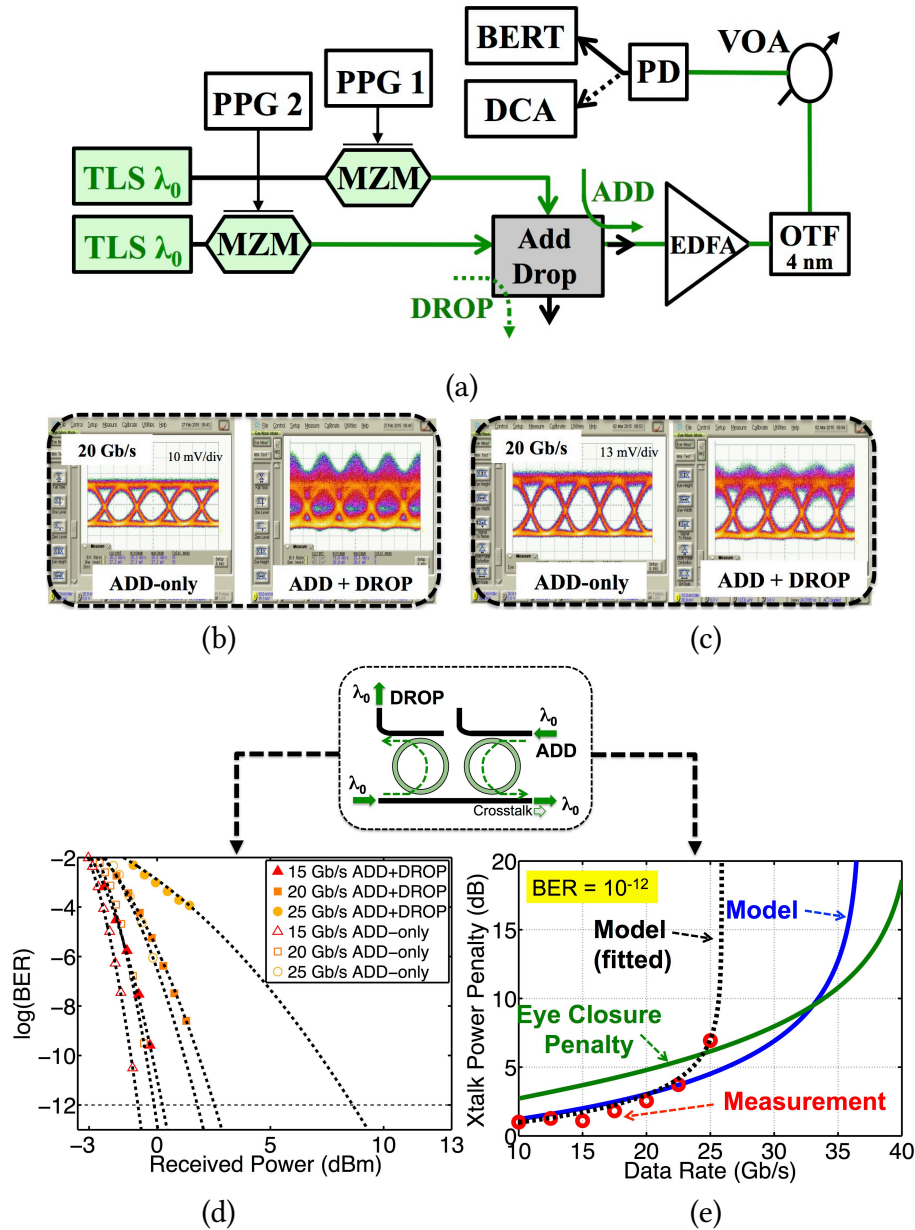
Figure 7.9: (a) Experimental setup for measuring BER curves and recording eye diagrams of the *intra-channel crosstalk* between two NRZ-OOK channels with the same wavelength. (b) Eye diagram for a single-ring structure at 20 Gb/s. (c) Eye diagram for a double-ring structure at 20 Gb/s. (d) BER curves with Gaussian extrapolation for 10^{-12} at 25 Gb/s data rate. (e) Measured intra-channel crosstalk penalty (red circles) and the estimation of penalty based on the model (solid blue curve) for two cascaded rings.

with its parameters modified as

$$\text{PP}_{XTalk} \approx -5\log_{10}\left(1 - 0.72\, Q_{\text{BER}} \frac{P_{XTalk}^{av}}{P_{NRZ}^{av}} \frac{r+1}{r-1}\right) \tag{7.11}$$

falls in good agreement with the measurement data (plotted in Fig. 7.9e as a dotted black curve). The derivation of Eq. (7.11) starts with the observation that the inter-channel crosstalk model (Eq. 7.5) accounts very well for the intra-channel crosstalk penalty when the data rate is low (see Fig. 7.9e); however, it tends to underestimate the penalty when the crosstalk becomes strong. Starting with the argument of the $log_{10}(...)$ in Eq. (5.41) and using the fact that $1 - 0.5x \approx \sqrt{1-x}$ for small x, and $1 - 0.5x > \sqrt{1-x}$ for large x, we conclude that replacing $1 - 0.5x$ in Eq. (7.5) with $\sqrt{1-x}$ will lead to almost the same penalty when the crosstalk is weak and lead to a larger penalty when the crosstalk is strong. This assumption transforms Eq. (7.5) into the following equation as an approximation for the intra-channel crosstalk

$$\text{PP}_{XTalk} \approx -5\log_{10}\left(1 - Q_{\text{BER}} \frac{P_{XTalk}^{av}}{P_{NRZ}^{av}} \frac{r+1}{r-1}\right). \tag{7.12}$$

Finally, we notice that introducing a factor of 0.72 in the argument of $log_{10}(...)$ function will lead to a better fit to the measured data presented in Fig. 7.9e.

Hence, we posit that Eq. (7.11) can be taken as a useful model for estimating the intra-channel crosstalk in microring-based Mux/DeMux structures. To make a comparison between Eq. (7.5) and Eq. (7.11) for the intra-channel crosstalk, the two equations are calculated for one and two cascaded rings (Fig. 7.2b and 7.2c). The curves are presented in Fig. 7.10. Solid curves correspond to Eq. (7.5) and dotted curves correspond to Eq. (7.11). As before, we assume that the through port rejection ratio of rings is 20 dB and the extinction ratio of channels at the drop port of the rings is about 9 dB such that $(r+1)/(r-1) \approx 1.23$. These numbers are from the measured spectra of the rings described in [81]. It can be seen that the crosstalk is very severe for single-ring structure

and in order to successfully switch the two channels to the other waveguides, the FWHM bandwidth of the ring must be far more than 4 times the data rate (i.e. ξ = FWHM/($2r_b$) > 2), which corresponds to a power penalty of 2.4 dB. If one considers ring resonators with a bandwidth of 30 GHz, as the ones we considered in the previous section, this limits the NRZ-OOK rate to 7.5 Gb/s (upper x-axis in Fig. 7.10). In contrast, if two rings are utilized (Fig. 7.2c), the filtering effect is accumulated, intra-channel crosstalk is reduced, and higher bit-rates can be allowed (or, equivalently requirements on the rings' FWHM bandwidth can be relaxed). This has been also reported by Gambini *et al.* for both bus and ring on-chip architectures [62]. Inspection of the plots presented in this figure further reveals that the model predicts a forbidden region for each structure. For example, the forbidden region of the single-ring structure based on Eq. (7.5) is marked in Fig. 7.10 with a black arrow. If data rate is too high or the rejection bandwidth of the rings is too small that the ξ factor falls within this forbidden region, the power penalty is considered infinite, meaning that the eye diagram of the optical signal is severely if not totally impaired due to the crosstalk fluctuations. From the solid curves in Fig. 7.10 (underestimation of the intra-channel crosstalk) we can see that the forbidden region of the single-ring structure is above 18 Gb/s data rate. Hence, at 20 Gb/s the eye-diagram is severely influenced by the crosstalk noise as it falls within the forbidden region. Fig. 7.9b agrees well with this prediction.

7.5   Summary of the Chapter

In this chapter, we presented a modeling approach for estimating the *inter-channel* crosstalk power penalty in silicon photonic links with integrated microring add/drop de/multiplexers. The crosstalk penalty depends on the bandwidth of each ring, spacing between channels, modulation rate of each channel, and to some extent, is independent of the characteristics of the photodiode. We used experimental measurements to validate
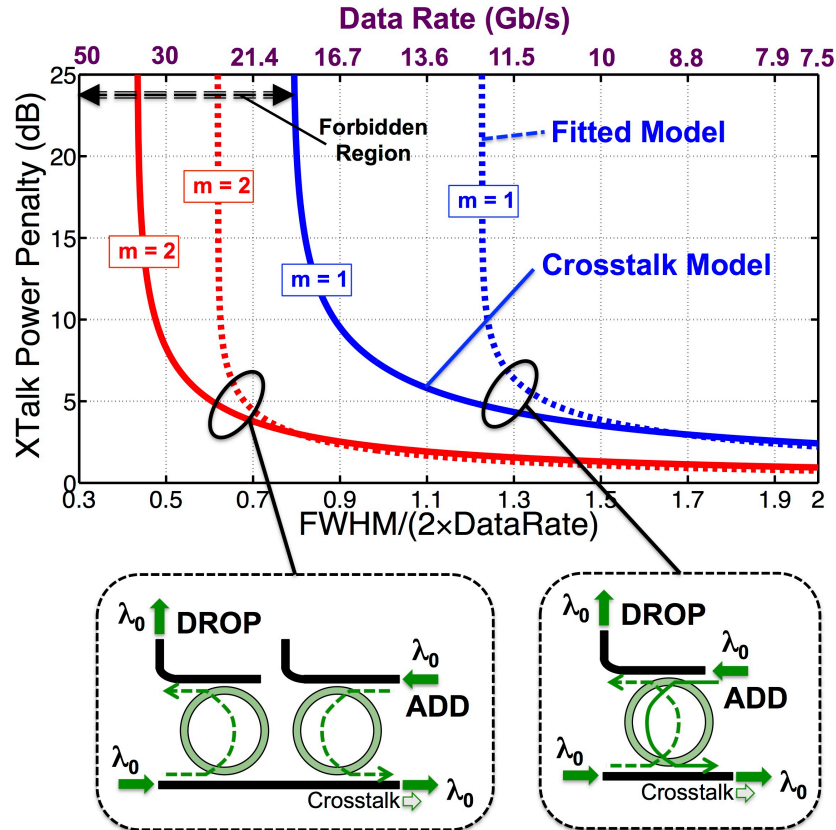
Figure 7.10: Estimated intra-channel crosstalk power penalty for the structures of one (blue) and two (red) cascaded microring resonators. It is assumed that the two channels are launched with the same power. The upper axis corresponds to the data rate when the FWHM of the rings is set to 30 GHz.

the proposed model. We also leveraged our model to analyze the crosstalk distribution across channels when employing a ring-based array of demultiplexers, and observed that channel reshuffling can be beneficial.

Our approach is less applicable to *intra-channel* crosstalk modeling, especially at high rates, due to the unknown statistical behavior of the coherence between crosstalk channels. However, for low data rates the predictions of the model are reliable and can be used to estimate the performance of spatial add/drop wavelength switching.

Chapter 8

Scalability of Microring-Based Silicon Photonic Interconnects

Abstract– *In this chapter, we present a comprehensive physical layer design and modeling platform for silicon photonic interconnects. The platform is based on explicit, closed-form expressions for optical power penalties, derived for both signal-dependent and signal-independent noise contexts. We show how the modeling approach is used for the design space exploration of silicon photonic links and can be leveraged to optimize the wavelength division multiplexed (WDM) capacity, evaluate the scalability, and study the sensitivity of the system to key device parameters. We apply the methodology to the design of microring-based silicon photonic links, including an evaluation of the impairments associated with cascaded ring modulators, as well as the spectral distortion and crosstalk effects of demultiplexer ring arrays for non-return-to-zero (NRZ) on-off keying (OOK) modulated WDM signals.*

With the vast rise in parallel multicore architectures, the scalability of computing performance is increasingly reliant on the availability of high-bandwidth, energy efficient data communications infrastructure. Silicon photonics has attracted considerable interest [3, 154, 175] as an emerging technology that can offer close integration of CMOS electronics with optical devices [189, 7]. This technology has the potential for delivering ultra high bandwidth interconnect solutions that are at the same time energy efficient and available at low cost. Silicon photonics has been proposed for designing efficient networks-on-chip [217, 45, 166, 167, 76], for supporting global interconnects for datacenters [130], and high performance computing (HPC) systems [88, 194]. It has been also presented as an enabling technology for realizing Exascale computing systems (i.e. Su-
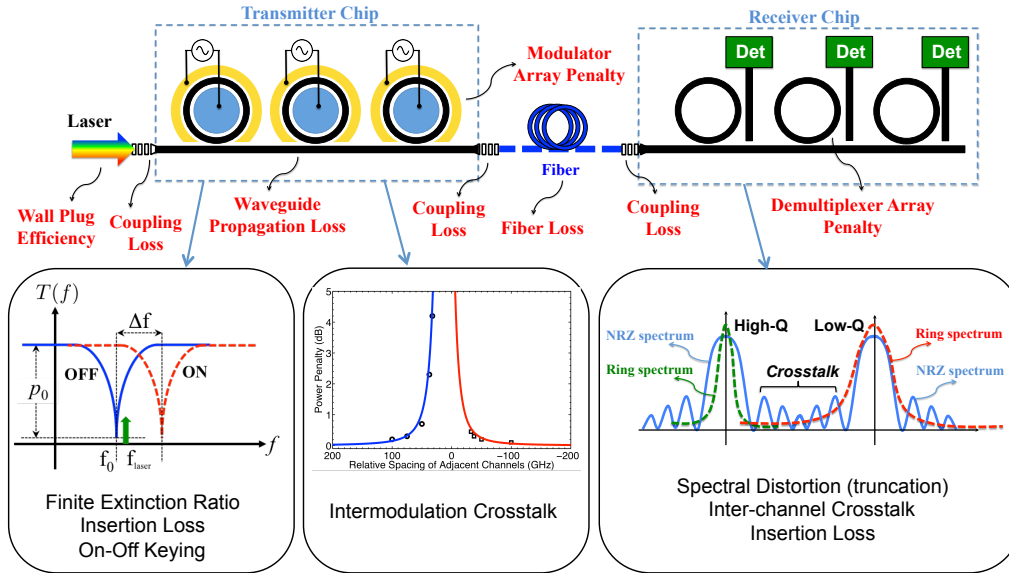
Figure 8.1: Structure of a microring-based silicon photonic link and the associated optical power penalties.

percomputers capable of realizing 10^{18} operations per second) [158, 44]. The chip-scale integration of electronics and photonics may further alleviate the limitations of electronic chips in terms of I/O pin count [121, 71].

8.1 Optical Power Budget and Power Penalties

Consider a simple chip-to-chip wavelength division multiplexed (WDM) silicon photonic link as shown in Fig. 8.1. On the transmitter side, cascaded microrings along an on-chip waveguide modulate the incoming multi-wavelength light generated either separately by external single-λ lasers or together by a comb laser source [38]. The incoming wavelengths, once imprinted with data, are then transmitted through an optical fiber to a receiver chip. The receiver chip consists of multiple passive microrings whose resonances are tuned to channel wavelengths. The total capacity of this link is obtained by multiplying the number of channels (wavelengths) with the modulation bit rate. Intuitively, it is tempting to maximize the number of wavelengths, N_λ, and/or to choose higher bit rates, r_b, for each channel. This allows for higher utilization of the available spectrum in the
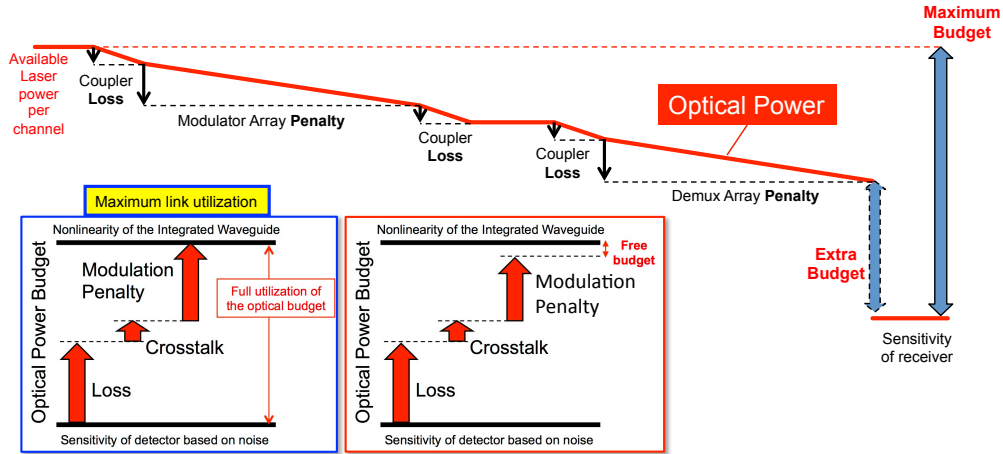
Figure 8.2: Plot diagram of the optical power budget from the input laser source to the receiver end of the link.

transmission media. However, as either the number of wavelengths and/or the bit rate grows, crosstalk and other undesired impairments emerge and eventually prevent reliable transmission. The total capacity of the link is closely tied to the optical power losses and impairments that the light experiences over the entire link. Summing up all the power penalties of the link, PP^{dB}, for a single channel, the following inequality must hold [73]:

$$\left[P_{laser}^{dBm} - 10 \log_{10} N_\lambda\right] - P_{sensitivity}^{dBm} \geq PP_{tot}^{dB}. \tag{8.1}$$

where the first term denotes the optical power per channel assuming that all the input laser lines have initially the same optical power. Figure 8.2 shows a diagram of how the optical budget is per channel shrinks due to the optical power penalties along the link.

In general, aggregated optical power P_{laser} (sum over all wavelengths) must stay below the nonlinear threshold of the silicon waveguides at any point of the link [74, 132]. In contrast, the signal power should stay above the sensitivity of the detectors $P_{sensitivity}$ (minimum number of photons or equivalently a certain amount of optical power) at the receive side to establish an error-free transmission link. A typical receiver may have a sensitivity of -12.7 dBm at 8 Gb/s operation [39], while a good receiver may exhibit a sen-

sitivity of about -21 dBm at 10 Gb/s [117]. This highly depends on the CMOS technology used (i.e. 65nm CMOS or 28nm CMOS, etc.) and the type of electronic circuit designed for the receiver. The difference between lower and higher thresholds (i.e. laser power per channel minus the sensitivity of the receiver) results in the *maximum power budget* that can be exploited. This budget accounts for the power penalty per channel (how much higher should the signal hit the detector), PP_{tot}, including the optical losses and signal impairments that cause deterioration on the bit error rate (BER) of the link.

We will show that the power impairments induced by the rings depend on the spacing between channels, N_λ, which is inversely proportional to the number of wavelengths, and on the modulation rate r_b. Eq. (8.1) is therefore a nonlinear function of N_λ and r_b, and determination of the $N_\lambda \times r_b$ product that maximizes the bandwidth directly from it is intractable. It is however possible to optimize the link, i.e. optimize its constituting devices, such that PP_{tot} is minimized for different (N_λ , r_b) combinations. We proceed to do this optimization for NRZ-OOK (non-return-to-zero On-Off keying) modulation format.

8.2   Properties of NRZ-OOK Modulation Format

In an NRZ-OOK modulation, the electrical signal is a sequence of logical 0's and 1's that are imprinted on the envelope of the light. Therefore, the envelope electric field (or simply the baseband signal) of the modulated light has a form of

$$E(t) = \sum_{k=-\infty}^{\infty} a_k \, p(t - kT_b) \tag{8.2}$$

where T_b is the bit interval and its inverse ($r_b = 1/T_b$) is the bit rate of the modulation, a_k is the amplitude for 0's and 1's bits, and $p(t)$ is the ideal square pulse of unit amplitude and duration T_b. We denote the high level of $E(t)$ by A and the low level by B as shown in Fig. 8.3a. Ideally, if the modulation has an infinite extinction ratio (i.e. $A^2/B^2 \to \infty$),

then the low level must be zero ($B = 0$); however, in practice optical modulators produce a limited extinction ratio. Typical numbers for 10G devices are 8–10 dB of extinction [52]. It is a practical assumption that all bits are equiprobable and the source of information that generates these bits is a stationary ergodic source. Therefore $Prob(a_k = A) = Prob(a_k = B) = 0.5$.

The average behavior, i.e. statistics, of such modulation is given by the auto-correlation of the envelope electric field over time. The average power spectral density of the envelope of modulation is then the Fourier transform of its autocorrelation, which simplifies to the following equation [147, 79]:

$$S_E(f) = \mu^2\, \delta(f - f_c) + \sigma^2 \frac{1}{r_b} \text{sinc}^2((f - f_c)/r_b), \tag{8.3}$$

where

$$\mu = \frac{A + B}{2}, \tag{8.4a}$$

$$\sigma = \frac{A - B}{2}, \tag{8.4b}$$

f_c is the optical frequency of the light, and $\delta(f - f_c)$ is the Dirac delta function centered at carrier frequency f_c. A plot of this spectrum is presented in Fig. 8.3b. To prove Eq. (8.4), we start by calculating the auto-correlation function of the electric field

$$R_{EE}(\tau) = \frac{1}{T_b} \int_{-\infty}^{\infty} E(t)E(t + \tau)\, dt =$$

$$= \left(\sum_m \sum_n a_m a_n \right) \frac{1}{T_b} \int_{-\infty}^{\infty} p(t - mT_b)p(t - nT_b + \tau)dt. \tag{8.5}$$

that corresponds to the optical power. The average auto-correlation function, $\overline{R_{EE}}(\tau)$ is

then the statistical average of this function:

$$\overline{R}_{EE}(\tau) = \overline{a_m a_n} \times \overline{R}_{pp}(\tau + (m - n)T_b) \tag{8.6}$$

where the first term is the auto-correlation of the discrete sequence:

$$R_{aa}[k] = \overline{a_m a_{m+k}} = \begin{cases} \overline{a_m^2} & \text{if} \quad k = 0 \\ \overline{a_m} \times \overline{a_m} & \text{if} \quad k \neq 0 \end{cases} = \begin{cases} \sigma^2 + \mu^2 \\ \mu^2 \end{cases} \tag{8.7}$$

and σ^2 and μ are the variance and average of the two A and B levels as given by Eq. (8.4). $\overline{R}_{pp}(\tau)$ is the average auto-correlation of the unit pulse defined by

$$\overline{R}_{pp}(\tau) = \frac{1}{T_b} \int_{-\infty}^{\infty} p(t)p(t + \tau)dt = \frac{1}{T_b} (p(\tau) \circledast p(-\tau)) = \Lambda\left(\frac{\tau}{T_b}\right) \tag{8.8}$$

where $\circledast$ denotes the convolution, and $\Lambda(\tau)$ is the standard triangle function of amplitude 1 and duration 2. Figure 8.3d shows a plot of $\overline{R}_{pp}(\tau)$ function. Let us now define the Fourier transform of this function

$$S_p(f) = \mathcal{F}\{\overline{R}_{pp}(\tau)\} = \frac{1}{r_b} \operatorname{sinc}^2\left(\frac{f}{r_b}\right) \tag{8.9}$$

and note that $S_p(f)$ is basically the *power spectral density* of $p(t)$ due to the Parseval's theorem:

$$\begin{aligned} \text{Average Power} \quad &= \frac{1}{T_b} \int_{-\infty}^{\infty} |p(t)|^2 dt \\ &= \int_{-\infty}^{\infty} \frac{1}{T_b} |\mathcal{F}\{p(t)\}|^2 df = \int_{-\infty}^{\infty} S_p(f)df. \end{aligned} \tag{8.10}$$

In the same way, we can define the power spectral density of the discrete sequence a_m
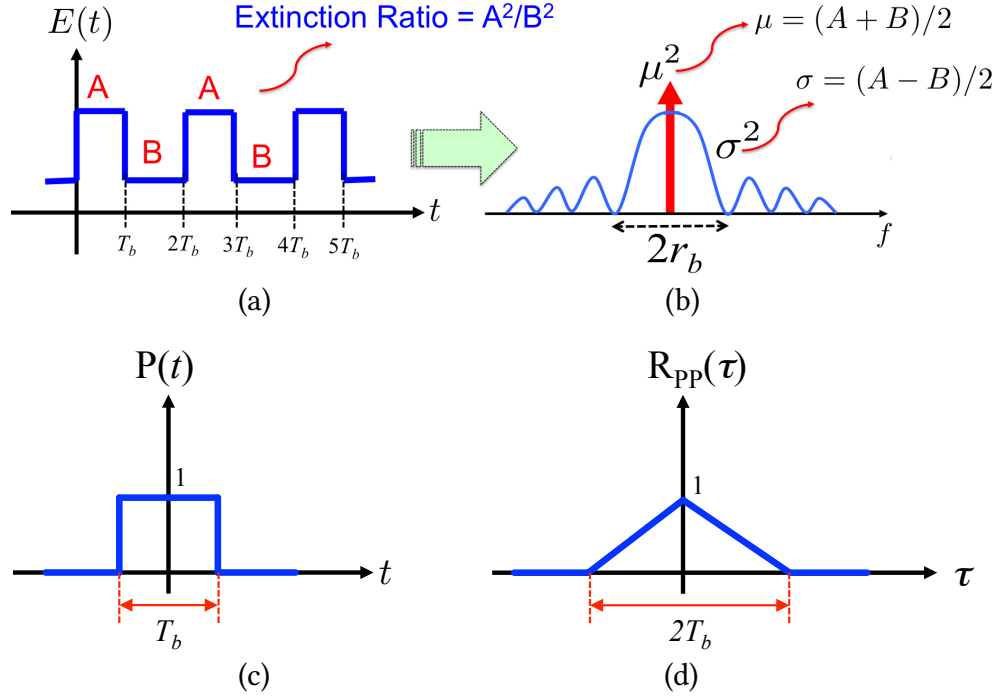
Figure 8.3: (a) Envelope electric field of the NRZ-OOK modulation, and (b) its power spectral density. Extinction ratio is the ratio of the optical power of the high level to the optical power of the low level of the electric field.

based on the discrete Fourier transform of its auto-correlation function:

$$S_a(f) = \mathcal{F}\{R_{aa}[k]\} = \mathcal{F}\{\mu^2 + \sigma^2\,\delta[k]\} = \mu^2 \sum_{m=-\infty}^{\infty} \delta(f - m) + \sigma^2. \qquad (8.11)$$

Finally, the power spectral density of the electric field is calculated by taking the Fourier transform of Eq. (8.6):

$$
\begin{aligned}
S_E(f) \quad &= \mathcal{F}\{\sum_{k=-\infty}^{\infty} R_{aa}[k]\,\overline{R}_{pp}(\tau - kT_b)\} = \sum_{k=-\infty}^{\infty} R_{aa}[k]e^{-j2\pi fkT_b} S_p(f) \\
&= S_p(f) \times S_a(T_b f)
\end{aligned}
\qquad (8.12)
$$

and by noting that

$$S_a(T_b f) = \left[r_b\,\mu^2 \sum_{m=-\infty}^{\infty} \delta(f - m\,r_b) \right] + \sigma^2 \qquad (8.13)$$

237

we reach to the following equation

$$S_E(f) = \left[\mu^2 \sum_{m=-\infty}^{\infty} \text{sinc}^2(m)\, \delta(f - m\, r_b) \right] + \sigma^2\, S_p(f). \qquad (8.14)$$

Since $\text{sinc}(m)$ is only non-zero at $m = 0$, it is seen that Eq. (8.3) is reached.

The power of the modulated signal is given by integrating Eq. (8.3) over all frequencies, yielding

$$\text{P}_{\text{av}} = \mu^2 + \sigma^2 = \frac{A^2 + B^2}{2}. \qquad (8.15)$$

We note that the power of the NRZ-OOK modulation can be decomposed based on the mean and variance of the envelope electric field. The μ^2 power is at the center frequency (represented by the delta function) and the power of the variance (σ^2) is spread over a spectrum of frequencies (represented by the *sinc* function). We will leverage Eq. (8.3) to estimate the power penalty of each optical component along the link.

Bit Error Rate of OOK Modulation

The quality of a digital communication system is characterized by the bit error rate (BER) [also known as bit error ratio]. Considering the two-level OOK signaling, the receiver converts the optical power into electrical current and then based on a threshold level (I_{th}), a decision is made for the value of each bit in each bit interval. If we assume that signal noise at the receiver has a Gaussian distribution for both the high and low levels (based on the central limit theorem[1]) as shown in Fig. 8.4a, we can mathematically characterize the probability of error as

$$\text{BER} = \text{Prob}(i(t) < I_{\text{th}}|\text{`1'})\,\text{Prob}(\text{`1'}) + \text{Prob}(i(t) > I_{\text{th}}|\text{`0'})\,\text{Prob}(\text{`0'}). \qquad (8.16)$$

---

[1]In probability theory, the central limit theorem (CLT) establishes that, in most situations, when independent random variables are added, their properly normalized sum tends toward a normal distribution (informally a "bell curve") even if the original variables themselves are not normally distributed. The theorem is a key concept in probability theory because it implies that probabilistic and statistical methods that work for normal distributions can be applicable to many problems involving other types of distributions.
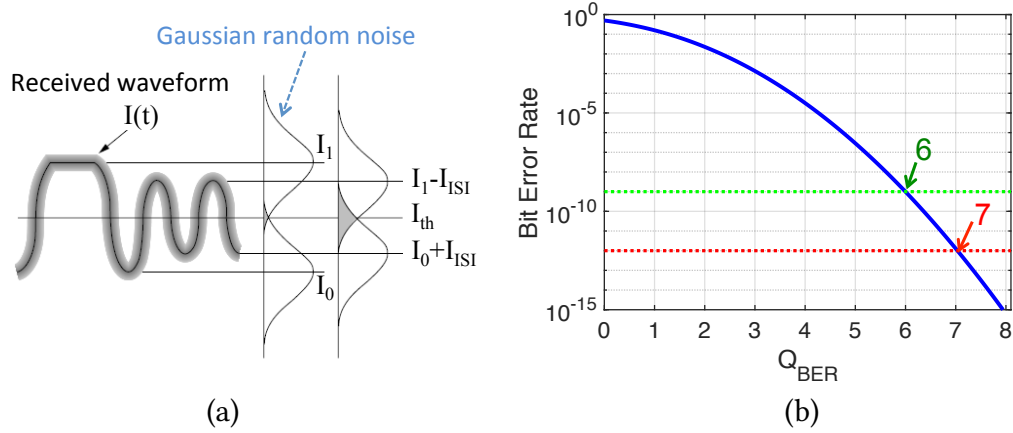
Figure 8.4: (a) An example of received signal and the threshold level and Gaussian noise distribution. (b) Plot of BER as a function of Q_{BER}.

Since we assume that all bits have the same probability $\mathrm{Prob}(`0`) = \mathrm{Prob}(`1`) = 0.5$. Therefore, the probability of error on the '1' bit is given by

$$\mathrm{Prob}(i(t) < I_{\mathrm{th}}|`1`) = \int_{-\infty}^{I_{\mathrm{th}}} \frac{1}{\sigma_{n,1}\sqrt{2\pi}} \exp\left(-\frac{(i-I_1)^2}{2\sigma_{n,1}^2}\right) di \qquad (8.17)$$

which can also be written in terms of complementary error function:

$$P_e(`1`) = \frac{1}{2} \mathrm{erfc}\left(\frac{I_1 - I_{\mathrm{th}}}{\sigma_{n,1}\sqrt{2}}\right). \qquad (8.18)$$

The probability of error for '0' bit is similarly given by

$$\mathrm{Prob}(i(t) > I_{\mathrm{th}}|`0`) = \int_{I_{\mathrm{th}}}^{\infty} \frac{1}{\sigma_{n,0}\sqrt{2\pi}} \exp\left(-\frac{(i-I_0)^2}{2\sigma_{n,0}^2}\right) di. \qquad (8.19)$$

which can also be written in terms of complementary error function:

$$P_e(`0`) = \frac{1}{2} \mathrm{erfc}\left(\frac{I_{\mathrm{th}} - I_0}{\sigma_{n,0}\sqrt{2}}\right). \qquad (8.20)$$

239

As can be seen, BER is a function of the decision threshold, I_{th}. The optimal decision threshold can be derived for minimizing the BER by taking the derivative of BER

$$\frac{\partial \text{BER}}{\partial I_{\text{th}}} = 0 \Rightarrow \frac{\sigma_{n,1}}{\sigma_{n,0}} = \exp\left[\frac{(I_{\text{th}} - I_0)^2}{2\sigma_{n,0}^2} - \frac{(I_{\text{th}} - I_1)^2}{2\sigma_{n,1}^2}\right] \tag{8.21}$$

and solving the quadratic equation. If we take the log of both sides and assume that $\ln(\sigma_{n,1}/\sigma_{n,0}) \approx 0$, then the minimum BER is approximately given by

$$\text{BER} \approx \frac{1}{2}\text{erfc}\left(\frac{I_1 - I_0}{(\sigma_{n,1} + \sigma_{n,0})\sqrt{2}}\right) = \frac{1}{2}\text{erfc}\left(\frac{Q_{\text{BER}}}{\sqrt{2}}\right) \tag{8.22}$$

where Q_{BER} defined as

$$Q_{\text{BER}} = \frac{I_1 - I_0}{\sigma_{n,1} + \sigma_{n,0}} = \frac{A^2 - B^2}{\sigma_{n,A} + \sigma_{n,B}} \tag{8.23}$$

is called the quality factor of the eye diagram of the OOK signal, which is closely related to signal to noise ratio at the receiver. Figure 8.4b shows a plot of BER as a function of Q_{BER}. We can see that a BER of 10^{-9} corresponds to $Q_{\text{BER}} \approx 6$ and a BER of 10^{-12} corresponds to $Q_{\text{BER}} \approx 7$.

Power Penalty Evaluation Based on Signal Statistics

As the modulated light goes through the waveguides, transmission medium, and other optical devices, the quality of the signal degrades. To evaluate the breadth of the degradation, one approach consists of evaluating the new high and low levels (assumed to be A' and B') when the light reaches the receiver. Alternatively, one can analyze how the power spectral density experiences distortion due to the non-ideal frequency response of the link. In other terms, one can evaluate how μ and σ evolve into μ' and σ'. Once these latter values are known, the power penalty (or Q penalty) associated with the distortion can be derived using one of the two following models widely used in literature. If

the receiver is limited by a signal-independent noise (SIN) mechanism (such as thermal noise) that does not beat against the optical signal, and considering a Gaussian distribution for the uncertainty of high and low levels, the power penalty to get the same BER performance and compensate for the signal degradations is estimated as [151]

$$\text{PP}^{\text{SIN}} \approx -10 \log_{10} \left(\frac{\sigma'}{\sigma} \frac{\mu'}{\mu} \right). \tag{8.24}$$

In contrast, in amplified systems the noise is mainly due to the amplified spontaneous emission (ASE) and that depends on the intensity of the input light to the amplifier (signal dependent noise, SDN). Using the model proposed by Downie [54] that relates the variance of the noise to the high and low levels of the electric field, the power penalty is estimated as

$$\text{PP}^{\text{SDN}} \approx \begin{cases} -10 \log_{10} \left(\mu'/\mu \right) & \text{if } \sigma'/\mu' > 1 \\ \\ -10 \log_{10} \left(\sigma'/\sigma \right) & \text{if } \sigma'/\mu' \leq 1 \end{cases}. \tag{8.25}$$

This equation implies that the power penalty in this case is related to the worst-case decrease of μ or σ. If the center frequency loses more power than the spectral frequencies, then $\sigma'/\mu' > 1$ and the power penalty is due to the change in μ. If the spectrum loses more power than the center frequency, then $\sigma'/\mu' < 1$, and the penalty is due to the change in σ. In accordance with Eq. (8.24) and (8.25), two approaches can be taken to obtain power penalties:

1) Estimate the high and low levels (A' and B') first, and then find μ' and σ' using equations (8.4a) and (8.4b). This is the approach we take for characterizing penalties of ring modulators.

2) Directly estimate μ' and σ' based on μ, σ, and the spectral response of the optical device. This is the approach we take for the demultiplexing ring filters.

Also note that in this chapter, we mainly consider the SIN model of the power penalty and assume no amplification in the link. If amplification is to be considered in the receiver,
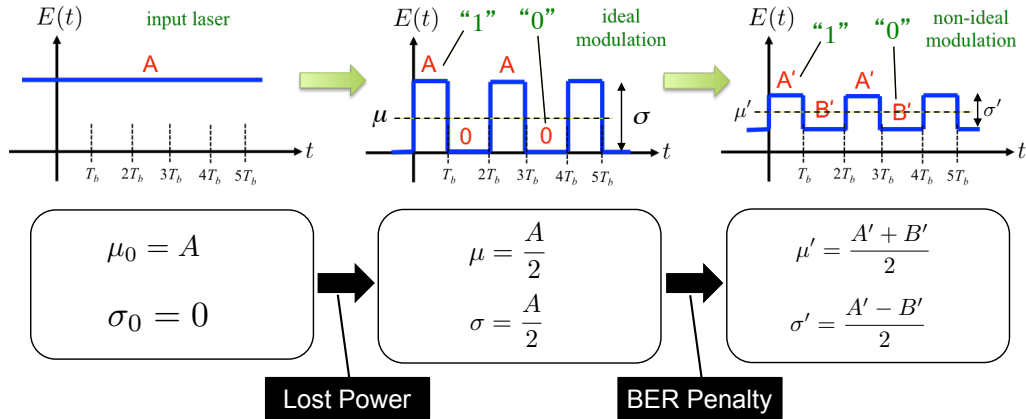
Figure 8.5: Characterization of the power penalty of optical OOK modulation.

as it is the case in the experiments against which this work has been compared, Eq. (8.25) should be employed instead of Eq. (8.24).

8.3 Modulation Power Penalty

Microring resonators are capable of modulating light at very high speeds while operating at low powers. This property combined with the small footprint of these optical modulators turn them into very promising candidates for integrated photonics. The physics behind the operation of such modulators has been well established [96]. Just like all other practical modulators, microrings are not capable of providing ideal modulation and thus some impairments and penalties on the optical signal are inevitable. In this section, we provide a general description of the power penalties that are of the most importance.

As depicted in Fig. 8.5, the electric field envelope of the unmodulated input light has constant amplitude that we denote by A, and constant power $P_{in} = A^2$. The modulation action occurs by shifting of the resonance of the ring as shown in Fig. 8.6. An ideal NRZ-OOK ring modulator traps the input light depending on whether the data bit is '1' or '0' (shown in Fig. 8.6). The mean and variance of this ideal modulation is $\mu = A/2$, and $\sigma = A/2$, which gives the average power $P_{mod} = \mu^2 + \sigma^2 = A^2/2 = P_{in}/2$ (see Fig. 8.5). We note that $-10\log_{10}(P_{mod}/P_{in}) = 3$ dB. As data bits are assumed equiprobable, this means

obfuscating half of the power. Therefore:

$$\text{Ideal Modulator:} \quad \text{PP} = 3\,\text{dB}\,. \tag{8.26}$$

A non-ideal modulation differs from the ideal modulation in the way that the high level is A' instead of A (where $A' < A$) and the low level is B' instead of zero, so that $\text{P}'_{\text{mod}} = \mu'^2 + \sigma'^2 = (A'^2 + B'^2)/2$ as shown in Fig. 8.5. This causes the power penalty to grow beyond the 3 dB of an ideal modulation. If the transmission of the ring for bit '0' and '1' at the laser wavelength is denoted by T_0 and T_1, then $A' = \sqrt{T_1}$ and $B' = \sqrt{T_0}$. The values of T_0 and T_1 can be estimated by considering a Lorentzian shape for the resonance spectrum of the ring [67]. For example, if the Lorentzian spectrum of the modulator is given by

$$T(\lambda, \lambda_{\text{res}}, \Delta\lambda_{\text{res}}, t_0) = \frac{(2/\Delta\lambda_{3\text{dB}} \times (\lambda - \lambda_{\text{res}}))^2 + t_0}{(2/\Delta\lambda_{3\text{dB}} \times (\lambda - \lambda_{\text{res}}))^2 + 1} \tag{8.27}$$

then $T_0 = T(\lambda_{\text{laser}}, \lambda_{\text{res}}, \Delta\lambda_{\text{res}}, t_0)$ and $T_1 = T(\lambda_{\text{laser}}, \lambda'_{\text{res}}, \Delta\lambda'_{\text{res}}, t'_0)$, where the prime terms indicate the change in the resonance and optical bandwidth of the ring due to the free carrier dispersion (FCD) and free carrier absorption (FCA).

The extinction ratio (ER) of modulation (also known os the optimal modulation amplitude or OMA) is the ratio of power of bit '1' to the power of bit '0', i.e. $r = A'^2/B'^2$. Using equations (8.4a) and (8.4b) again, the mean and variance of the non-ideal modulation levels can be written as $\mu' = B'(\sqrt{r} + 1)/2$, and $\sigma' = B'(\sqrt{r} - 1)/2$. After some further algebraic manipulations[2] the following relations are obtained:

$$\frac{\mu'}{\mu} = \frac{\sqrt{r} + 1}{\sqrt{r} + 1}\sqrt{\frac{P'_{\text{mod}}}{P_{\text{mod}}}} \quad , \quad \frac{\sigma'}{\sigma} = \frac{\sqrt{r} - 1}{\sqrt{r} + 1}\sqrt{\frac{P'_{\text{mod}}}{P_{\text{mod}}}}. \tag{8.28}$$

These can be inserted in Eq. (8.24) or (8.24), depending on the noise conditions. Consid-

---

[2]in particular, by rewriting levels in terms of average modulation powers $A = \sqrt{2\text{P}_{\text{mod}}}$ and $B' = \sqrt{2\text{P}'_{\text{mod}}/(r+1)}$.

ering a dominant signal-independent noise (SIN) in the system, so Eq. (8.24), the total modulation power penalty is given by

$$\text{PP}_{\text{Mod}}^{\text{SIN}} = -10 \log_{10}\left(\frac{r-1}{r+1}\frac{P'_{\text{mod}}}{P_{\text{mod}}}\right) - 10\log_{10}\left(\frac{P_{\text{mod}}}{P_{\text{in}}}\right) \tag{8.29}$$

where the second term is the -3 dB corresponding to the ideal modulation. By eliminating P_{mod} and noting that $P'_{\text{mod}} = (r+1)A'^2/2r$, the power penalty can be rewritten as

$$\boxed{\text{PP}_{\text{Mod}}^{\text{SIN}} = -10\log_{10}\left(\frac{r-1}{r+1}\right) - 10\log_{10}\left(\frac{A'^2}{P_{\text{in}}}\right) - 10\log_{10}\left(\frac{r+1}{2r}\right)} \tag{8.30}$$

This formulation is helpful as it permits to distinguish three contributing power penalty factors. The first term is the power penalty due to the finite extinction ratio (ER) of modulation (or OMA) as widely used in the literature [96]. The second term is the modulator insertion loss (IL), defined as the difference between the input power and the power of bit '1'. Finally, the third term is the penalty due to the ON-OFF keying nature of the imperfect modulation. For an ideal modulator (infinite extinction ratio, $r \to \infty$, and no loss), this last term is 3 dB while the first two terms are reduced to zero. Figure 8.6 provides more details of the power penalty of microring modulators. Figure 8.7a provides an example of the spectral response of a microring modulators with an active shift of 0.2 nm. Figure 8.7b shows the calculation of the modulation extinction ratio (i.e. OMA) and the power penalty of modulator based on Eq. (8.30). Clearly, the best OMA and minimum penalty requires a good alignment between the laser and the resonance of the modulator.

If a silicon photonic link is dominated by the amplified spontaneous emission (ASE) noise that beats against the optical signal, the power penalty of modulator is estimated from Eq. (8.25) by noticing that $(\mu'/\mu)>(\sigma'/\sigma)$. The result is

$$\text{PP}_{\text{Mod}}^{\text{SDN}} \approx -10\log_{10}\left(\frac{\sqrt{r}-1}{\sqrt{r}+1}\right) - 5\log_{10}\left(\frac{P'_{\text{mod}}}{2P_{\text{in}}}\right). \tag{8.31}$$
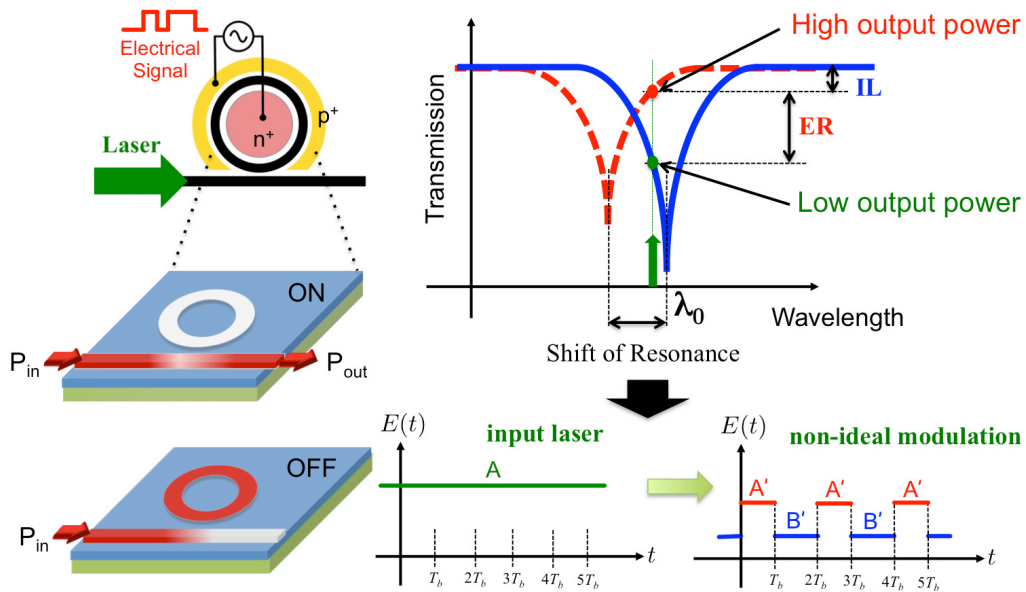
Figure 8.6: (left) Schematic view of a microring modulator. The high-speed electrical signal is applied to the pn-junction embedded inside the silicon ring. (right) Modulation of the input laser by shifting the resonance of the ring to create high and low levels of optical power at the output.
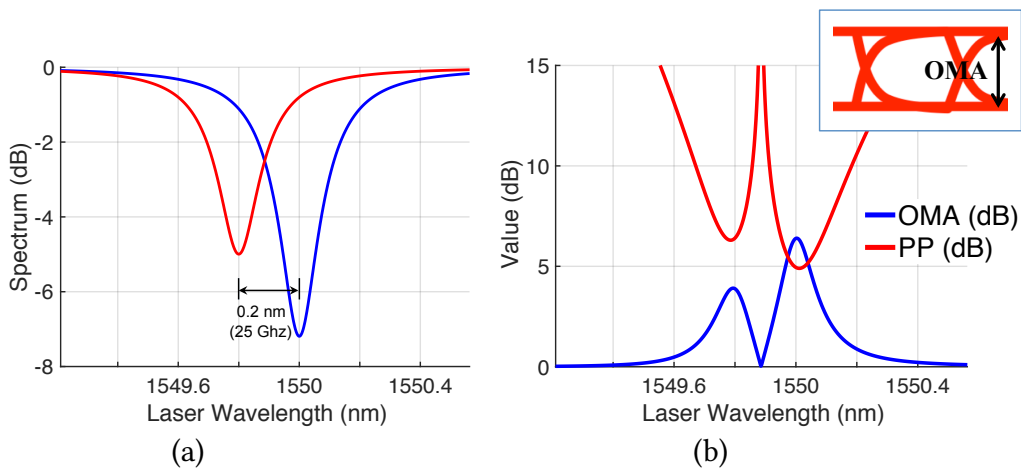
Figure 8.7: (a) Spectral response of a microring modulator for 25 GHz of spectral shift. (b) Estimated OMA and power penalty of the modulator in part a.

The first term in this equation is the power penalty due to the finite extinction ratio of modulation (which is different from the one in Eq. (8.30)) and the second term represents the average power loss due to the combination of ON-OFF keying and insertion loss of modulator.

Now, we consider several ring modulators mounted on a silicon waveguide, each one modulating a specific wavelength (as in Fig. 8.1). Cascading all the ring modulators on a single bus waveguide induces two additional power penalties. The first one to be accounted for is the passive cumulative insertion loss of the laser power of each channel as it passes by the other rings. If the spacing between wavelengths is small and the number of modulators is large, this could become a significant loss. The second one is the array-induced multiplexing crosstalk [135] (also known as *intermodulation crosstalk*). Multiple wavelengths of the comb laser pass by each ring, but only one of them gets modulated. Due to the carrier injection or depletion of the p-n junction of the ring, the spectrum of the ring is switched between two resonance frequencies as shown in the inset of Fig. 8.8: f_0 (OFF state) and f_1 (ON state), where $f_1 > f_0$ (blue shift). If this shift ($\Delta f = f_1 - f_0$) is too large, then the shifted spectrum may capture some of the power of the neighboring channel that passes by the ring (which then acts as a noise), hence inducing an intermodulation crosstalk. If f_Δ represents the channel spacing and $f_\Delta > 0$, then a neighboring channel at $f_0 + f_\Delta$ would suffer higher crosstalk than a neighboring channel at $f_0 - f_\Delta$. Using the crosstalk model presented in the previous chapter, we reproduce the results of intermodulation crosstalk reported by Padmaraju *et al.* [135] for a microring modulator with $Q \approx 6000$ at the resonance wavelength 1546.7 nm and $q_0 \approx 0.04$. The reproduced results are plotted in Fig. 8.8 where circles and squares are penalties from experimental data for $f_\Delta > 0$ and $f_\Delta < 0$, respectively.

At this point, we have introduced the main ingredients to estimate the total power penalty induced by a ring-resonator based modulator array. Knowing the parameters Q, coupling coefficients of ring-waveguides, and loss of each ring, its associated Lorentzian transform can be obtained, from which stems A' and B'. These permit to evaluate Eq. (8.30) or (8.31) as well as the intermodulation crosstalk penalty.
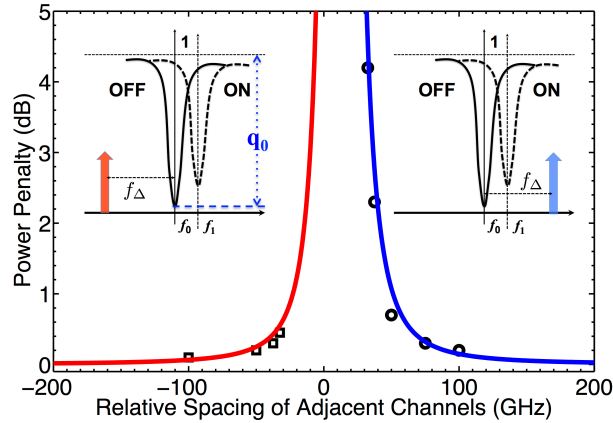
Figure 8.8: $PP_{\text{Mod-Xtalk}}$ for two cases: (blue) when the adjacent channel is at a higher frequency ($f_\Delta > 0$), (red) when the adjacent channel is at a lower frequency ($f_\Delta < 0$). Due to the unilateral shift of the ring modulator, the first case exhibits higher penalty than the second case. Measurement data [135] are denoted by circles and squares.

8.4   Microring Demultiplexers

Consider an NRZ-OOK modulated signal passing through microring resonator based filter as shown in Fig. 8.9a. The filter has two outputs: the *drop port* and the *through port*. Considering this filter as a linear time independent (LTI) system, it can be characterized by two spectral responses $H_{\text{drop}}(f)$ and $H_{\text{through}}(f)$ that relate the output electric field present at each port to the input electric field. Knowing that the input power spectral density is $S_E(f)$, we can calculate the output power spectral density as $S_{\text{out}}(f) = S_E(f)\,|H(f)|^2$ at each port [147]. Unless $H(f)$ is constant for all f, which is not realizable with ring resonators, some degree of *spectral distortion* will appear. This translates into eye closure in the time domain as shown in Fig. 8.9b. At the receiver, the received signal is sampled in each bit interval. We consider this as a reconstruction of the NRZ-OOK signal with new average high and low levels denoted by A' and B', and new statistical mean μ' and variance σ'^2. Using Eq. (8.3), the change in the mean and variance by going through the
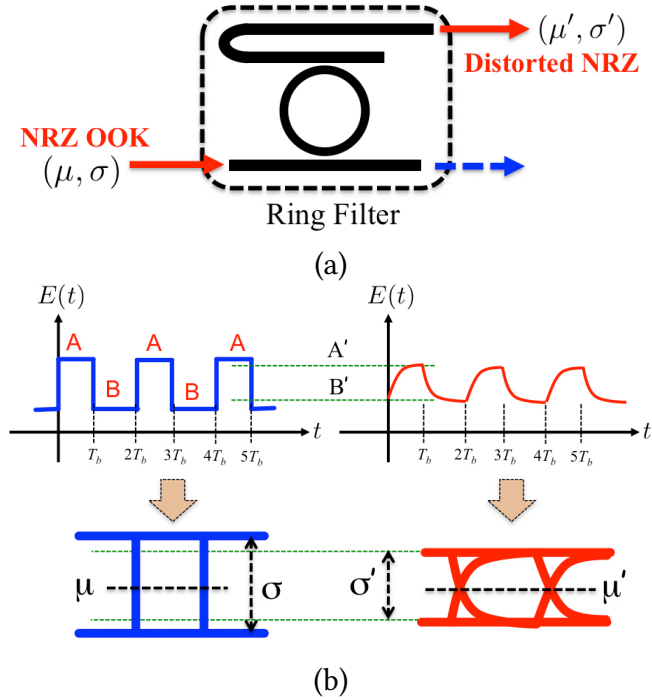
Figure 8.9: (a) Schematic representation of a demux ring filter. The input NRZ modulation is characterized by the mean and variance of its high and low levels. The dropped signal is a distorted NRZ modulation that can be characterized by new mean and variance parameters. (b) Comparison between the input NRZ modulation to a filter and its output response. The output is a signal that effectively has new high and low levels A' and B' with different mean and variance than the original signal. This can be seen by inspecting the eye-diagrams of the input and output optical signals.

$H(f)$ system can be written as

$$\left(\frac{\mu'}{\mu}\right)^2 \approx |H(f_c)|^2 \tag{8.32a}$$

$$\left(\frac{\sigma'}{\sigma}\right)^2 \approx \int_{-\infty}^{\infty} |H(f)|^2 \frac{1}{r_b} \mathrm{sinc}^2\left(\frac{f - f_c}{r_b}\right)\, df. \tag{8.32b}$$

The relation between μ and μ' is approximately independent of the shape of the filter since the power associated with the mean of the two levels is centered at a single frequency. On the other hand, the relation between σ and σ' depends on the shape of the filter, indicating *spectral distortion* to the power spectrum of the signal.

Fig. 8.10a shows demux ring resonators at the receive side of a WDM silicon photonics
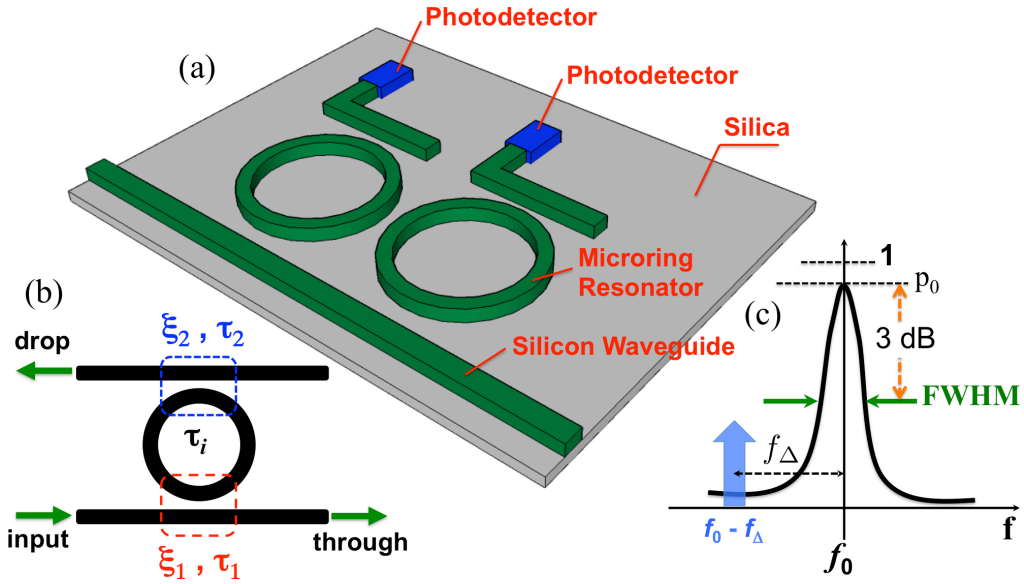
Figure 8.10: (a) Graphical view of microring demultiplexers at the receiver. (b) Intrinsic and coupling decay rates of a ring add-drop filter. (c) Transmission spectrum of the drop path of a demux ring with the 3-dB bandwidth denoted as FWHM. f_Δ is the possible detuning between the resonance and the channel.

link, where each one is tuned to drop one of the NRZ-OOK modulated channels. Each ring has an intrinsic loss that corresponds to a decay time-constant denoted by τ_i. The coupling rates and the coupling decay time constants for the ring-waveguide interactions are denoted by (ξ_1, τ_1) and (ξ_2, τ_2) for the through and drop paths as depicted in Fig. 8.10b. If the quality factor of the rings is high enough and the coupling of ring-waveguide is weak, one can employ the equations of the temporal coupled mode theory (tCMT) for microring resonators [105]. By virtue of the conservation of energy in the ring-waveguide system, the relation between (ξ_1, τ_1) and (ξ_2, τ_2) is then given as $\xi_{1,2} \approx \sqrt{2/\tau_{1,2}}$ and the drop and through field transfer functions are

$$H_{\text{drop}}(\Omega) \approx \xi_1 \xi_2 \frac{1}{j\Omega + 1/\tau} \tag{8.33a}$$

$$H_{\text{through}}(\Omega) \approx \frac{j\Omega + 1/\tau'}{j\Omega + 1/\tau} \tag{8.33b}$$

where $\Omega = 2\pi(f - f_0)$ is the frequency deviation from the resonance of the ring (f_0), and the time constants τ and τ' are given by

$$1/\tau = 1/\tau_i + 1/\tau_2 + 1/\tau_1 \tag{8.34a}$$

$$1/\tau' = 1/\tau_i + 1/\tau_2 - 1/\tau_1. \tag{8.34b}$$

The full width at half maximum of the resonance is calculated as FWHM = $1/(\pi\tau)$, and the quality factor of the resonance is given by $Q = f_0/\text{FWHM} = \omega_0\tau/2$. The power transfer function from the input to the drop port is $\left|H_{\text{drop}}(\Omega)\right|^2$, which has a Lorentzian lineshape as depicted in Fig. 8.10c. The main parameters of this function are the resonance frequency of the ring (f_0), 3-dB bandwidth (FWHM), and the peak drop power $p_0 = (\xi_1\xi_2\tau)^2$. Considering the Lorentzian function presented in Eq. (8.33a), our first goal is to minimize the insertion loss (make it as close as possible to 0 dB) of the drop path of the demux ring. For convenience, we introduce the normalized rate parameters as $x = (1/\tau_1)/(1/\tau_i)$ and $y = (1/\tau_2)/(1/\tau_i)$. For a demux ring, the drop insertion loss is

$$\left|H_{\text{drop}}(0)\right|^2 = \frac{4xy}{(1 + x + y)^2}. \tag{8.35}$$

This expression is completely symmetric for both x and y. Considering a given value for y, equating the derivative of the $\left|H_{\text{drop}}(0)\right|^2$ with respect to x to zero yields $x_{\text{opt}} = 1 + y$ which is exactly the *critical coupling condition* (input coupling rate is equal to the intrinsic loss rate plus the output coupling rate) for a ring resonator. Substituting x_{opt} then gives $\left|H_{\text{drop}}(0)\right|^2 = y/(1 + y)$ and thus y must be chosen large enough to get as close to 0 dB as possible. A large y can correspond to a big τ_i which in turn corresponds to a low loss ring. Substituting for $\tau_i \approx 2n_g/(c\,\alpha_{\text{loss}})$ where n_g is the group index of the mode, α_{loss} is the loss factor (1/m) of the ring (including waveguide loss and bending loss) and $c = 3 \times 10^8$

m/s is the speed of light, the minimum insertion loss of drop path (in dB) is calculated as

$$\text{IL}_{\text{demux}} = -10 \log_{10} \left(1 - \frac{2Q}{Q_i} \right) \tag{8.36}$$

where the intrinsic quality factor is $Q_i = 2\pi n_g/(\alpha_{\text{loss}} \lambda_0)$, and λ_0 is the resonance wavelength of the ring. This equation suggests that the optimum design for a demux filter is to have it as low-loss as possible at the critical coupling point. The coupling coefficients of the electric field between the waveguide and the ring at the input and drop ports of the demux ring are then estimated as

$$\kappa_1 \approx \sqrt{(2\pi R\, \alpha_{\text{loss}})x} \;\;, \;\; \kappa_2 \approx \sqrt{(2\pi R\, \alpha_{\text{loss}})y} \tag{8.37}$$

where $2\pi R$ is the circumference of the ring, α_{loss} is the loss factor of the ring expressed in 1/m, and $x = 1 + y$ at the critical coupling point. These coupling coefficients can be used to estimate the gap distances between the ring and the waveguides.

Once the drop and through shapes of the filter are known, the next step consists of looking at the impact that the drop path has on the spectrum of an NRZ-OOK modulation. This means evaluating the integral of Eq. 8.32b for a Lorentzian function. The result is given by

$$\gamma = \frac{1}{1 + \xi^2} - \frac{1}{2\pi\nu}\text{Re}\left(\frac{1 - \exp(-2\pi\nu(1 - j\xi))}{(1 - j\xi)^2} \right) \tag{8.38}$$

in which $\nu = f_0/(2Qr_b)$, $\xi = 2Qf_\Delta/f_0$, and $\text{Re}(\dots)$ indicates the real part of the complex number. Here, f_Δ denotes any possible detuning between the channel and the resonance of the demux ring as shown in Fig. 8.10c. To prove this relation, we first note that the variance of the NRZ-OOK modulation changes by getting dropped by the ring as $(\sigma'/\sigma)^2 = p_0 \times \gamma$ where

$$\gamma = \int_{-\infty}^{\infty} \frac{\text{sinc}^2(F)}{1 + ((F - F_\Delta)/\nu)^2} dF, \tag{8.39}$$

$\nu = \text{FWHM}/(2r_b)$, $F_\Delta = f_\Delta/r_b$, and F is the frequency normalized to bit rate. In order to

calculate this integral, one has to employ the residue theorem. Rewriting the integral in the form of

$$\gamma = \frac{\nu^2}{2\pi^2}\text{Re}\int_{-\infty}^{\infty}\frac{1-\exp(j2\pi F)}{F^2((F-F_\Delta)^2+\nu^2)}dF \tag{8.40}$$

and choosing the appropriate contour in the upper half plane shown in Fig. 8.11 indicates that there is only one pole at $Z_0 = F_\Delta + j\nu$ inside this contour. The closed-loop integral is then calculated as

$$\oint \frac{1-\exp(j2\pi Z)}{Z^2((Z-F_\Delta)^2+\nu^2)}dZ = \frac{\pi}{\nu Z_0^2}(1-\exp(j2\pi Z_0)). \tag{8.41}$$

Obviously, due to the higher-order denominator, the integral over $C(\infty)$ is zero. The integral over $C(\varepsilon)$ where $\varepsilon \to 0$ on the other hand is not zero and must be calculated. Substituting $Z = \varepsilon\exp(j\theta)$ into the integrand of Eq. (8.41), will then yield the following result

$$\int_{C(\epsilon)} \frac{1-\exp(j2\pi Z)}{Z^2((Z-F_\Delta)^2+\nu^2)}dZ = \frac{-2\pi^2}{|Z_0|^2}. \tag{8.42}$$

Combining the above equations finally leads to

$$\gamma = \frac{\nu^2}{|Z_0|^2} + \frac{\nu}{2\pi}\text{Re}\left(\frac{1-\exp(j2\pi Z_0)}{Z_0^2}\right) \tag{8.43}$$

which can also be cast as Eq. (8.38).

We note that the first term of Eq. (8.38) is not a function of bit rate, reflecting only the filtering nature of the device at a distance f_Δ from its resonance. If there is no detuning between the channel and the filter, then $f_\Delta = 0$, $\xi = 0$, and γ is reduced to $\gamma = 1 - (1 - \exp(-2\pi\nu))/(2\pi\nu)$. The second term, on the other hand, depends on the bit rate, highlighting the fact that the bandwidth of the NRZ-OOK signal is a contributor to the spectral distortion imposed by the ring filter. If modulation speed is low enough, the second term is negligible compared to the first term.

We proceed to employ Eq. (8.38) to evaluate the power penalty of the filtering action.
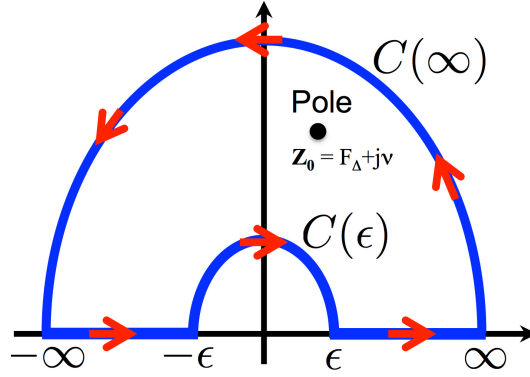
Figure 8.11: Selected contour in the upper half-plane for evaluating the integral of the spectral distortion penalty. There is a pole at the origin and a pole at $Z_0 = F_\Delta + j\nu$.

By introducing Eq. (8.33) into equations (8.32a) and (8.32b), we obtain $(\mu'/\mu)^2 \approx p_0/(1 + \xi^2)$ and $(\sigma'/\sigma)^2 \approx p_0 \times \gamma$. From there, using Eq. (8.24), we have

$$\boxed{\text{PP}^{\text{SIN}}_{\text{Demux}} \approx -10\log_{10}(p_0) - 5\log_{10}\left(\frac{1}{1+\xi^2}\right) - 5\log_{10}(\gamma)} \qquad (8.44)$$

for a link with signal-independent noise (SIN). The first term in PP is simply the drop insertion loss of the resonance spectrum of the ring. The second term reflects the insertion loss at the channel wavelength, and the third term reflects the penalty of spectral distortion due to the filtering. If the link is dominated by a signal-dependent noise (SDN), power penalty is estimated from Eq. (8.25) as

$$\text{PP}^{\text{SDN}}_{\text{Demux}} \approx -5\log_{10}(p_0) - 5\log_{10}(\gamma) \qquad (8.45)$$

Figure 8.12a shows the measured power penalties (circles) reported by Biberman *et al.* [28] at BER = 10^{-9} for a ring of a bandwidth about 10 GHz and four different bit rates: 5 Gb/s (PP $\approx$ 0.3 dB), 7.5 Gb/s (PP $\approx$ 0.5 dB), 10 Gb/s (PP $\approx$ 0.7 dB), and 12.5 Gb/s (PP $\approx$ 1.15 dB). The solid curve on this figure is the estimated power penalty up to 60 Gb/s (PP $\approx$ 3.3 dB) with f_Δ = 0 (exact tuning between the channel and the ring filter). Although the model provides a good estimation of the power penalties based on these measured data
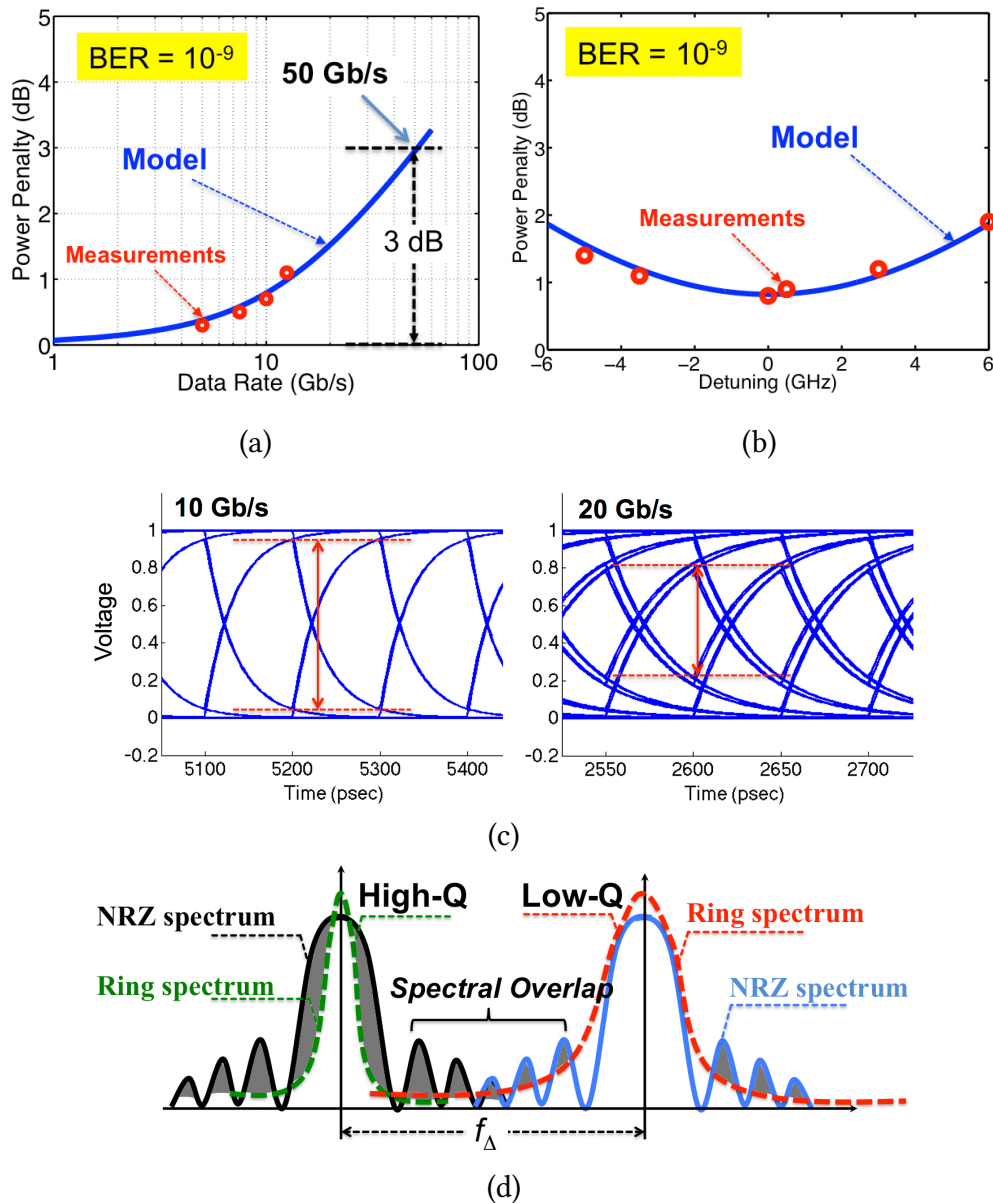
Figure 8.12: (a) Comparison of the estimated spectral distortion (solid curve) and measured penalties [28] (circles) for various bit rates. (b) Estimation of the penalty due to small detunings between the resonance of the ring and the channel wavelength. Measurements are taken from [91]. (c) Simulated eye-diagrams of the received signal for 10 Gb/s and 20 Gb/s data rates. (d) Schematic view of two adjacent channels with a fair amount of spectral overlap.

points, it will likely provide an underestimation of the power penalties for very high data rates.

Fig. 8.12b presents the measured penalties (circles) of small detunings (up to 6 GHz) reported by Lee *et al.* [91] at BER = 10^{-9} for a ring resonator with FWHM = 9.6 GHz ($Q \approx 20000$) and a bit rate of 10 Gb/s (PP $\approx$ 0.8 dB at zero detuning). The solid curve represents the estimated BER penalties for a link with signal-dependent noise mechanism (SDN–Eq. 8.45). The SDN is selected based on the fact that the experiment used optical amplification to compensate for power loss in the system. Here again, we observe good agreement between the measurements and calculations. These two comparisons against measurements let us posit that Eq. (8.38) can serve as a basis for quantifying the rate-dependent penalties due to *sideband truncation* of an NRZ-OOK channel.

Figure 8.12c illustrates the effect of the spectral distortion on the eye-diagram of the received signal for 10 Gb/s and 20 Gb/s data rates. As it can be observed, the eye-diagram of the 20 Gb/s signal experiences more closure in accordance with higher power penalty.

Besides the spectral distortion of each channel, power penalties also arise in the de-multiplexing process in the form of crosstalk. Due to the infinite tail of the Lorentzian resonance and spectral overlap of the channels, capturing and dropping some of the optical power from adjacent channels is unavoidable, as shown in Fig. 8.12d. We assume that the worst-case crosstalk happens at the first demux ring in Fig. 8.1, since all the channels are at their maximum power and none has been dropped yet (considering the discussion on the crosstalk in the previous chapter). The power that is captured by the Lorentzian tail of the ring from the spectra of the adjacent channels has a negative impact on the statistics of the high and low levels at the detector and increases the BER. In order to compensate for this, the crosstalk effect on the BER is turned into a power penalty.

At this point, we can estimate the worst-case power penalties inflicted by the demux array. Worst-case of crosstalk happens at the first ring, while the worst-case of waveguide loss happens for the last ring. Equation (8.44) estimate the impact of truncation while the

effect of crosstalk is modeled based on the previous chapter. Using these equations and the ones presented in the previous section for the modulator side, we can proceed to the next steps, which consist of optimizing the ring parameters for minimum power penalty and evaluating the ideal link format ($N_\lambda \times r_b$) for maximum bandwidth.

8.5 Scalability of a Microring-Based Silicon Photonic Link Under Ideal Drive Conditions

In this section, we analyze the scalability of the silicon photonic link with the assumption that the electronic circuitry (i.e. drivers for modulators and the trans-impedance amplifiers at the receiver) impose no limitations. In other words, we only consider the limitations that the photonics part of the link impose to fully close the optical power budget. The following assumptions are then applied:

1. an extreme FSR of 50 nm is possible for the rings (which corresponds to approximately 4 μm diameter of the rings) [208].

2. the FSR and loss of a ring resonator are independent of each other for the purpose of this optimization. Loss of a ring is similar to loss of a waveguide.

3. modulators can provide any desired shift of resonance and there is no limitation in terms of driver circuitry.

4. Q-factor of the ring resonators can be set to any desired value.

5. there is no amplifier in the link, thus the link is dominated by a signal-independent noise mechanism (SIN).

Optimization of Modulator Power Penalty

For the modulators, the parameters subject to optimization are Q-factor of the rings, shift of the resonance due to the free carrier dispersion, the change of the spectral parameters

due to the free carrier absorption, and the spacing between channels. First, we calculate
how the shift of the ring resonator, $\Delta\lambda_{\mathrm{res}}$ and the change in the refractive index of the
silicon waveguide, Δn^{si} are related to each other:

$$n_{\mathrm{eff}}(\lambda_{\mathrm{res}}) \times 2\pi R = m\,\lambda_{\mathrm{res}}\,,$$

$$[n_{\mathrm{eff}}(\lambda_{\mathrm{res}} + \Delta\lambda_{\mathrm{res}}) + \Delta n_{\mathrm{eff}}] \times 2\pi R = m\,(\lambda_{\mathrm{res}} + \Delta\lambda_{\mathrm{res}}) \tag{8.46}$$

where Δn_{eff} is the extra change in the effective index of the optical mode due to the free
carrier dispersion. Since m is the same in both of these equations (we are tracking the
same resonance mode), these two equations can be combined to result in

$$\frac{\lambda_{\mathrm{res}}}{n_{\mathrm{eff}}(\lambda_{\mathrm{res}})} = \frac{\lambda_{\mathrm{res}} + \Delta\lambda_{\mathrm{res}}}{n_{\mathrm{eff}}(\lambda_{\mathrm{res}} + \Delta\lambda_{\mathrm{res}}) + \Delta n_{\mathrm{eff}}}\,. \tag{8.47}$$

If we assume $\Delta\lambda_{\mathrm{res}} \ll \lambda_{\mathrm{res}}$, we can write

$$n_{\mathrm{eff}}(\lambda_{\mathrm{res}} + \Delta\lambda_{\mathrm{res}}) \approx n_{\mathrm{eff}}(\lambda_{\mathrm{res}}) + \frac{\partial n_{\mathrm{eff}}}{\partial\lambda} \times \Delta\lambda_{\mathrm{res}} \tag{8.48}$$

and then use the definition of group index $n_g = n_{\mathrm{eff}} - \lambda\,\partial n_{\mathrm{eff}}/\partial\lambda$ to simplify Eq. (8.47)
and reach the following relation:

$$\boxed{\Delta n_{\mathrm{eff}} \approx n_g(\lambda_{\mathrm{res}}) \times \frac{\Delta\lambda_{\mathrm{res}}}{\lambda_{\mathrm{res}}}} \tag{8.49}$$

and finally using the sensitivity coefficients of the optical mode, $\Delta n_{\mathrm{eff}} \approx (\Gamma_{\mathrm{core}} \times n_g/n_g^{si})\,\Delta n_{si}$, we find the change in the index of silicon core as a function shift of resonance as

$$\Delta n_{si} \approx \frac{n_g^{si}(\lambda_{\mathrm{res}})}{\Gamma_{\mathrm{core}}} \times \frac{\Delta\lambda_{\mathrm{res}}}{\lambda_{\mathrm{res}}} \approx 3.99 \times \frac{\Delta\lambda_{\mathrm{res}}}{\lambda_{\mathrm{res}}} \tag{8.50}$$

for a 450×220 nm ridge waveguide with 50 nm of slab height. Finally, the change in the loss of the ring is estimated by solving

$$\Delta n_{si} = -10^{-4} \left[0.6069 \, \Delta \alpha_{si} + 2.5138 \, (\Delta \alpha_{si})^{0.8} \right] . \qquad (8.51)$$

for lateral PIN type modulators. If a lateral PN modulator (i.e. carrier depletion) is considered, a mode overlap factor must be included as well [103]. We then proceed to calculate the effective loss factor of the optical mode as $\Delta \alpha_{\text{eff}} \approx 1.1 \, \Delta \alpha_{si}$ and then find the new quality factor and extinction of resonance:

$$T = \frac{\left(2Q \frac{\lambda - \lambda_{\text{res}}}{\lambda_{\text{res}}} \right)^2 + t_0}{\left(2Q \frac{\lambda - \lambda_{\text{res}}}{\lambda_{\text{res}}} \right)^2 + 1} \Rightarrow T = \frac{\left(2Q' \frac{\lambda - \lambda'_{\text{res}}}{\lambda'_{\text{res}}} \right)^2 + t'_0}{\left(2Q' \frac{\lambda - \lambda'_{\text{res}}}{\lambda'_{\text{res}}} \right)^2 + 1} . \qquad (8.52)$$

where $\lambda'_{\text{res}} = \lambda_{\text{res}} + \Delta \lambda_{\text{res}}$. The change in the extinction of resonance is calculated as

$$t_0 = \left(\frac{\tau_i - \tau_c}{\tau_i + \tau_c} \right)^2 \quad , \quad t'_0 = \left(\frac{\tau'_i - \tau_c}{\tau'_i + \tau_c} \right)^2 \qquad (8.53)$$

where τ_i and τ_c are the original intrinsic loss and coupling time constants of the modulator and τ'_i is the intrinsic loss time constant of the ring after the spectral shift, given by

$$\tau_i = \frac{2 \, n_g}{c \, \alpha_{\text{eff}}} \Rightarrow \tau'_i = \frac{2 \, n_g}{c \, (\alpha_{\text{eff}} + \Delta \alpha_{\text{eff}})} \qquad (8.54)$$

Similarly, the quality factor of the modulator is calculated as

$$Q = \frac{\pi c}{\lambda_{\text{res}}} \left(\frac{1}{\tau_i} + \frac{1}{\tau_c} \right) \quad , \quad Q' \approx \frac{\pi c}{\lambda_{\text{res}}} \left(\frac{1}{\tau'_i} + \frac{1}{\tau_c} \right) . \qquad (8.55)$$

Therefore, by starting with Q and t_0 and knowing $\Delta \lambda_{\text{res}}$, we can find Q' and t'_0 which results in calculating the shifted spectrum (see for example Fig. 8.7).

Considering these parameters, trade-offs then should be made between modulator
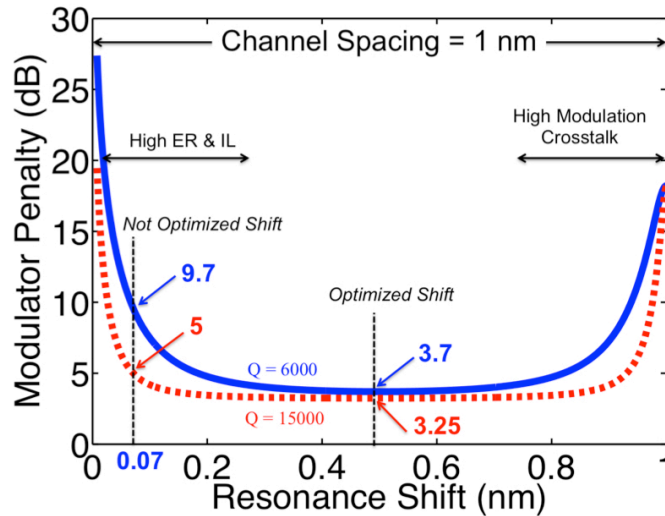
Figure 8.13: Estimated modulator penalty for a channel spacing of 1 nm and different resonance shifts. If the shift of the resonance is small, then the insertion loss and extinction ratio penalty are high. On the other hand, if the shift is too big, then the crosstalk penalty is very high. The optimum shift is about half of the channel spacing.

insertion loss, extinction ratio penalty, ON-OFF keying penalty, cumulative arrayed-induced insertion loss, and the multiplexing crosstalk. Figure 8.13 shows an example of calculating the total penalty of the ring modulator for a case where channel spacing is 1 nm (hence 50 rings) and the quality factor is either 6000 (low-Q, solid curve) or 15000 (high-Q, dotted curve). It is assumed that the rings are originally critically coupled to the waveguide. As this plot indicates, power penalty of a modulator is always higher than 3 dB (which is the penalty of an ideal NRZ-OOK modulator) and the *minimum value of the penalty occurs when the shift of the resonance of the ring is approximately half of the channel spacing*. Unfortunately, such a big shift of the resonance usually requires high peak-to-peak voltages for the electrical signal that drives the pn-junction of the ring (according to the formulation presented in [67] and [206]). Consequently, for CMOS-compatible voltages ($V_{pp} \sim 1$–2 V [51, 132]), the resonance shift is away from its optimum value and this leads to higher penalties on the modulators. For example in Fig. 8.13, a shift of about 0.5 nm leads to a penalty of 3.7 dB for Q = 6000 and a shift of 0.07 nm ($\sim$ 8.5 GHz) leads to about 9.7 dB penalty. It is important to note that here we make the assumption that the

ring modulators are driven with pre-emphasized NRZ signals to effectively compensate for the limited bandwidth of the modulator [209]. Hence, we always consider that the static ER can be achieved under dynamic drive. With this assumption, higher Q factor for the ring modulators is more desirable as it reduces the penalty of the modulator for a fixed amount of resonance shift.

Optimization of Demux Power Penalty

For the ring demultiplexers, the parameters subject to optimization are the Q-factor of the rings, bit rate of channels, and channel spacing. Trade-offs are then established among drop insertion loss, spectral distortion, and crosstalk penalties. As an example, Fig. 8.14a shows the optimized overall penalty of demux rings for 50 channels with 1 nm spacing between them, various bit rates, and various Q-factors. This penalty is the sum of the spectral distortion (Eq. 8.44), the cross-talk penalty and ring drop insertion loss. The latter is function of the intrinsic loss of the ring α_{loss} and Q-factor. Note that the intrinsic loss is not subject to optimization: the minimal value that fabrication can provide should simply be employed to limit insertion loss. In contrast, the Q-factor can clearly be optimized if the objective is to minimize the overall power penalty. Looking at Fig 8.14a, one notes that if the Q of the ring is small, crosstalk has a big impact on the penalty. On the other hand, if the Q is too high, then the spectral distortion penalty and insertion loss have higher impacts. Therefore, there is a minimum point for the power penalty at an optimum Q-factor. The dependence of the optimum penalty and optimum Q on the bit rate is emphasized on Fig. 8.14a with the dotted black curve showing how the optimum point behaves as the bit rate increases. This indicates that higher bit rates lead to wider filters in order to compensate for higher spectral distortion.

In Fig. 8.14b, same optimization is done for the demux rings but with 8 dB/cm of ring loss [80]. It is clear that in this case, high-Q rings will lead to a high insertion loss. In Fig. 8.14c we plot the penalty of different bit rates and different number of channels for
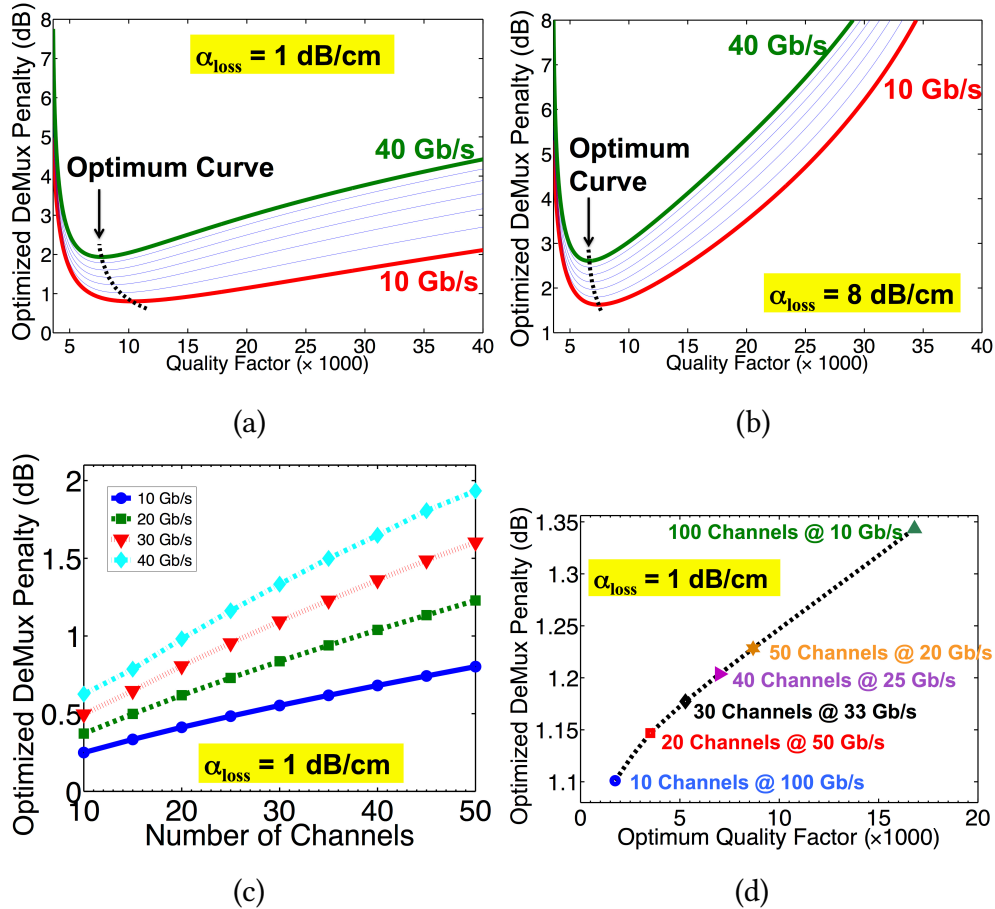
Figure 8.14: Optimizing the penalty of demux ring filters by including the insertion loss as a function of Q-factor of the ring. (a) Loss factor of the ring is 1 dB/cm and 50 channels. (b) Loss factor of the ring is 8 dB/cm and 50 channels. (c) Power penalty for different number of channels and bit rates (loss = 1 dB/cm) when using demux rings with optimized Q. (d) Optimum penalty and its corresponding Q-factor for an aggregate rate of 1 Tb/s and different channel rates (loss = 1 dB/cm).

optimally set Q-factors. It is seen that penalty tends to increase as the spacing between channels decreases. Finally, we perform the optimization on microring demux filter with the assumption of having a fixed amount of *total aggregate rate* (1 Tb/s is assumed). In this case, increasing the number of channels, e.g. from 10 to 100, is accompanied by a decrease in the modulation speed. The results are shown in Fig. 8.14d where six cases are clearly marked on the plot. With optimum design of the rings, the penalties result in a narrow range 1.1 dB–1.35 dB. Meanwhile, the corresponding quality factors can vary over a wide range from 1800 to 17000.
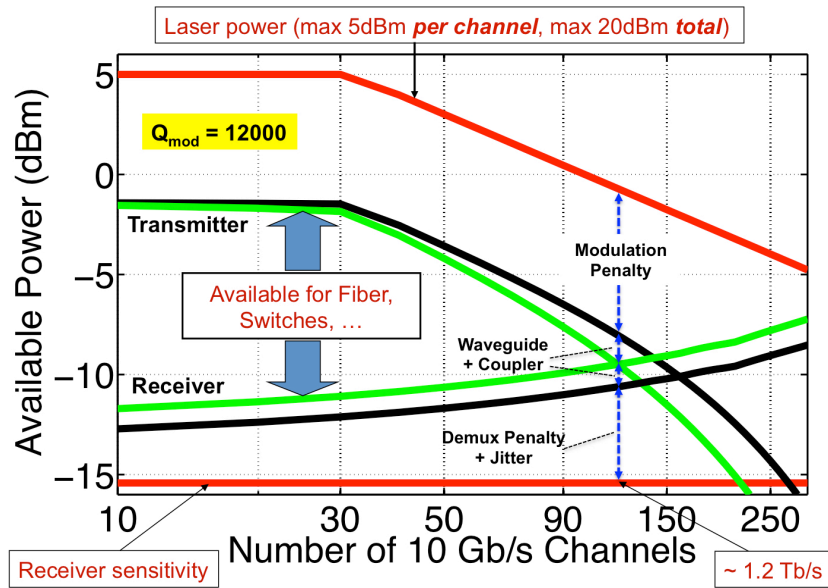
Figure 8.15: Calculation of the available optical power at the output of the transmitter and minimum required power at the input of the receiver. Both calculations reflect the worst-case. The difference is the available power budget for losses and penalties associated with fibers, or switches that are placed between transmitter and receiver. The capacity of the link is estimated to be 1.2 Tb/s. Quality factor of ring modulators is taken to be 12000 and intrinsic loss factor α_{loss} to be 1 dB/cm.

Scalability Analysis

The scalability of a silicon photonic link in terms of the maximum total aggregate rate depends on the available optical power budget. Based on the characterization of power penalties for ring modulators at the transmitter and ring demultiplexers at the receiver, it is possible to maximize the achievable aggregation. For each particular bit-rate, the analysis is carried out by finding the maximum value of N_λ such that Eq. (8.1) holds. Some of the power penalties of the link do not generally depend on the bit-rate of channels or on the number of channels/channel spacing. Examples include the waveguide propagation loss (~ 1 dB/cm [92]) and chip-to-fiber/fiber-to-chip coupling insertion loss (~ 1 dB) [132]. The penalties of the modulators are also approximately independent of bit-rate, but depend on the number of channels due to array-induced insertion loss and inter-modulation crosstalk. This can be taken as an encouragement to use higher bit-rates at the modulators. However, higher bit-rates limit the sensitivity of the receiver circuitry which, as we
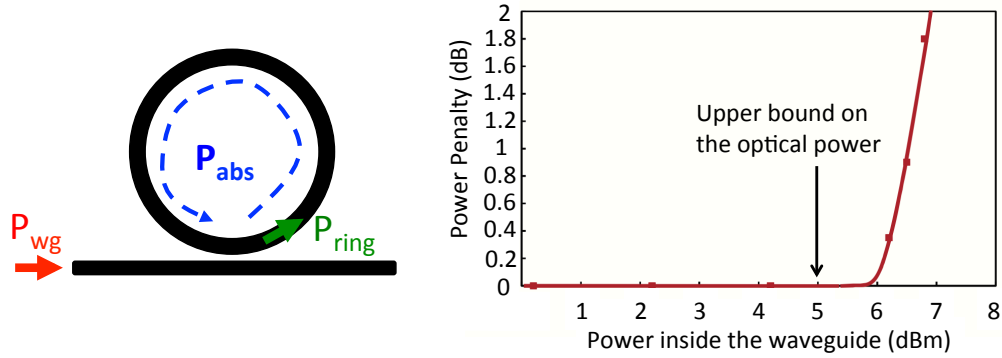
Figure 8.16: Characterization of the upper bound on the optical power inside the waveguide. The plot shows additional power penalty of the ring modulator due to the absorption of optical power (taken from [98]) and its self heating effect.

will see, eventually puts a cap on the maximum achievable bandwidth. Overall, this gives rise to an interesting design space to explore.

We initiate our explorations by analyzing the scalability of a silicon photonic link as a function of number of channels for 10 Gb/s NRZ-OOK modulation, as shown in Fig. 8.15. It is assumed that the laser power cannot exceed 5 dBm for each wavelength (after coupling into the waveguide), but the total amount of the optical power that is injected into the silicon waveguide is kept below 20 dBm to avoid the nonlinearities of the silicon waveguide. The upper bound of the optical power per channel is due to the absorption of the optical power inside the modulators *caused mainly by the free carrier absorption (FCA), especially in depletion-mode modulators*, which leads to the self heating [168, 169] of the rings, drift of the resonance, and a bi-stability cycle in the spectral response of the resonator [179, 214]. Consequently, the generated eye diagram of the modulator starts to close and an additional power penalty is imposed on the modulator. As shown in the plot of Fig. 8.16, the upper bound of the optical power has been experimentally measured by Li *et al.* to be about 5 dBm inside a ring modulator [98]. Beyond this value, the power penalty sharply increases.

In order to mathematically examine the self heating effect of microring resonators, we first note that if λ_{res} is the original resonance of the modulator and λ'_{res} is the drifted

resonance due to the absorption of optical power, then

$$\lambda'_{\text{res}} = \lambda_{\text{res}} + \Delta\lambda_{\text{res}} \Rightarrow \Delta\lambda_{\text{res}} = \frac{\lambda_{res}}{n_g(\lambda_{\text{res}})} \times \Delta n_{\text{eff}} \tag{8.56}$$

and the effective index changes with the temperature of the ring as

$$\Delta n_{\text{eff}} = \frac{dn_{\text{eff}}}{dT} \times \Delta T \approx 2.1 \times 10^{-4} \times \Delta T_{\text{ring}} \tag{8.57}$$

for a 450×220 nm waveguide. As shown in Fig. 8.16, if P_{ring} is the optical power inside the ring (right after the light is coupled into the ring) at any wavelength, the total absorbed power is $P_{\text{abs}} = (1 - L) P_{\text{ring}}$ where $L = \exp(-2\pi R\alpha_{\text{eff}})$ is the round trip power attenuation of the ring. Using the SFG method, the relation between the power inside the ring and inside the waveguide is given by

$$P_{\text{ring}} = P_{\text{wg}} \frac{\kappa^2}{1 + t^2 L - 2t \sqrt{L} \cos\left(2\pi(\lambda_{\text{laser}} - \lambda'_{\text{res}})/\text{FSR}\right)} \tag{8.58}$$

where κ and t are coupling coefficients of the electric field. In order to find the relation between ΔT_{ring} and P_{abs}, we assume that the ring is sitting on a buried oxide (BOX) layer of thickness d_{box} and temperature profile between the ring and the substrate has a linear gradient. Therefore at the location of the ring we have

$$\left|\frac{\partial T}{\partial n}\right| = \Delta T_{\text{ring}}/d_{\text{box}} \Rightarrow P_{\text{abs}} \approx k_{sio_2} \times 2\pi(w^2 + 2Rw + Rh)\frac{\Delta T_{\text{ring}}}{d_{\text{box}}} \tag{8.59}$$

where $w = 0.45\mu$m and $h = 0.22\mu$m are the width and height of the ring and $k_{sio_2} = 1.38$ W/m.K is the thermal conductivity of oxide. Therefore, the heating efficiency coefficient H_T defined by $\Delta T_{\text{ring}} = H_T \times P_{\text{ring}}$ is given as

$$H_T = \frac{[1 - \exp(-2\pi R\alpha_{\text{eff}})]\, d_{\text{box}}}{2\pi(w^2 + 2Rw + Rh)} \times \frac{1}{k_{sio_2}} \approx 0.5703 \quad \text{K/mW} \tag{8.60}$$
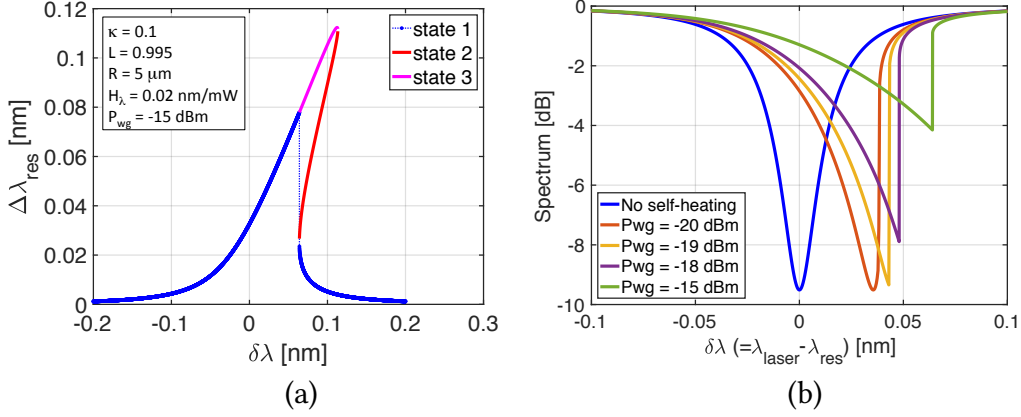
Figure 8.17: (a) Bistability of a microring modulator due to the self heating. (b) Deformation of the optical spectrum due to the self heating of the resonator.

for $R = 5\mu m$, $\alpha_{\text{eff}} = 20$ dB/cm, and $d_{\text{box}} = 2\mu m$. Finally, the self-heating coefficient of resonance, H_λ, defined as $\Delta\lambda_{\text{res}} = H_\lambda \times P_{\text{ring}}$ is given by

$$H_\lambda = \frac{\lambda_{\text{res}}}{n_g} \times \frac{dn_{\text{eff}}}{dT} \times H_T \approx 0.0423 \quad \text{nm/mW} \tag{8.61}$$

for a resonance at 1550 nm. Note that Eq. (8.56) and Eq. (8.58) are coupled to each other for a given optical power inside the waveguide. It is easy to see that for $\Delta\lambda_{\text{res}} \ll \text{FSR}$, the following equation is obtained:

$$\Delta\lambda_{\text{res}} = H_\lambda P_{\text{wg}} \times \frac{A_1}{1 + A_2 \left(2\pi(\delta\lambda - \Delta\lambda_{\text{res}})/\text{FSR}\right)^2} \tag{8.62}$$

where

$$\delta\lambda = \lambda_{\text{laser}} - \lambda_{\text{res}} \quad , \quad A_1 = \frac{\kappa^2}{(1 - t\sqrt{L})^2} \quad , \quad A_2 = \frac{t\sqrt{L}}{(1 - t\sqrt{L})^2} . \tag{8.63}$$

By solving this equation and finding $\Delta\lambda_{res}$ for a given P_{wg} at any laser wavelength, the transmission spectrum of the modulator can be derived and from that the modulation power penalty under the self heating can be obtained. Figure 8.17a shows an example of solutions of Eq. (8.62) for P_{wg} = -15 dBm. In this example, $Q \approx 40000$ which is a

265

rather high value. Figure 8.17b plots the calculated spectrum of the modulator under different input optical powers. Clearly, the self-heating of the modulators results in a deformation of the Lorentzian spectrum, and the lower the Q factor is the less sensitive the modulator becomes to self-heating effect. The upper bound of 5 dBm in Fig. 8.16 has been achieved with the assumption that *an active thermal tuning control loop will keep the Lorentzian spectrum in alignment with the input laser at all time* to correct for thermal drift. Otherwise, the upper bound of the optical power will be much lower than 5 dBm (it was measured to be -4.5d Bm in the original experiment [98]).

The sensitivity of the receiver for each channel is set to a value of -15.5 dBm (given by a rate-dependent model). These assumptions set the limit of the maximum power budget to 20.5 dB for each 10 Gb/s channel. From the transmitter side, we add up the modulator penalties (loss, extinction ratio, crosstalk) and the waveguide loss, assuming 1 dB/cm loss for the silicon waveguide. Since the number of required modulators and demultiplexers is equal to the number of channels, the length of the waveguides linearly scales up with the number of channels (the waveguide padding between modulators is set to 100 μm and set to 20 μm for demultiplexing filters). Then, we add up the worst-case penalty associated with the demultiplexers (including worst-case crosstalk and worst-case waveguide loss). We also account for a total of 2 dB extra margin to compensate possible jitter in clocking and polarization-dependent losses. In contrast, we do not account for power penalties related to *chromatic dispersion*. We assume this last to be negligible for short distance links with modulation rates < 100 Gb/s. Figure 8.15 shows that at about 120 channels, the power budget completely closes and thus the scalability of this link for 10 Gb/s channels is limited to about 1.2 Tb/s. For lower number of channels, there is some budget remaining. It can be exploited to introduce additional elements in the link (more couplers, switches, etc.). Alternatively, launching powers per channel can be reduced to limit the power consumption of the laser.

Next, the dependence of the scalability of the link on the modulation speed is investi-
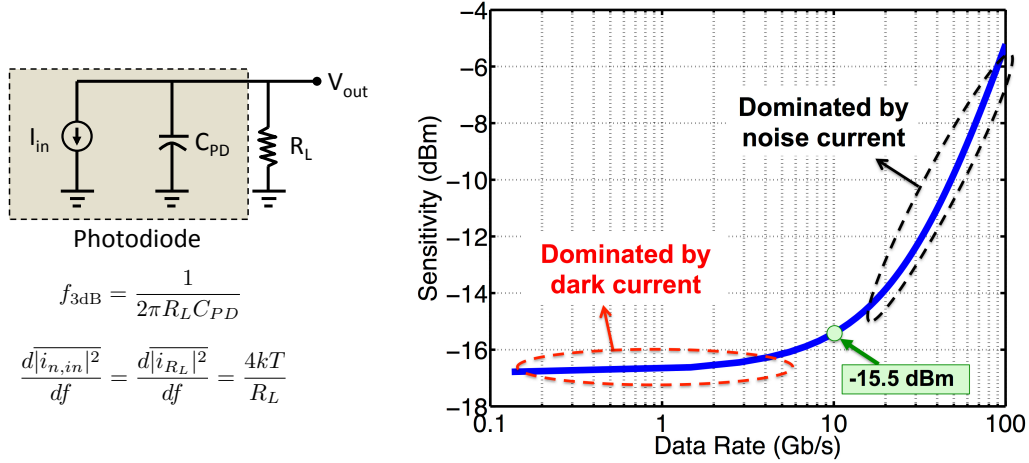
Figure 8.18: Estimated sensitivity of the receiver as a function of the modulation data rate based on a simple RC model. At low data rates, dark current dominates over the noise current and thus the sensitivity is almost constant, whereas at high data rates the noise current dominates over the dark current and sensitivity is reduced.

gated. In this case, we still assume the same condition on laser power (maximum 5 dBm per channel, maximum 20 dBm total), but a bit-rate dependent model for the sensitivity of the receiver (assuming a Germanium-based PIN detector) is employed. The model for the sensitivity is simply based on the thermal noise of an RC equivalent circuit for the receiver as shown in Fig. 8.18. The 3dB bandwidth and the input-referred power spectral density of the noise are given by

$$f_{\text{3dB}} = \frac{1}{2\pi R_L C_{PD}} \quad , \quad \frac{d\overline{|i_{n,in}|^2}}{df} = \frac{d\overline{|i_{R_L}|^2}}{df} = \frac{4kT}{R_L} \tag{8.64}$$

where C_{PD} is the capacitance of the photodiode and R_L is the input load resistance of the receiver. We consider a worst-case of 200 fF for C_{PD} [33], and then calculate the RMS noise current as a function of the receiver bandwidth:

$$I_{\text{noise}}^{\text{RMS}} = f_{\text{3dB}} \sqrt{8\pi kT\, C_{PD}}\,. \tag{8.65}$$

Using the definition of the BER for OOK modulation with Gaussian noise at the receiver,

the sensitivity is defined by the following equation:

$$P_{\text{sensitivity}} \approx 10 \log_{10}\left(Q_{\text{BER}} \times \frac{I_{\text{dark}} + I_{\text{noise}}^{\text{RMS}}}{R} \times \frac{10^{ER/10} + 1}{10^{ER/10} - 1} \right) \qquad (8.66)$$

where Q_{BER} is the argument of the complementary error function (Q_{BER} = 7 for BER = 10^{-12}), R is a responsivity [A/W] of the photodiode (assumed $\sim$0.7 A/W [34]), ER is the extinction ratio of input signal (assumed $\sim$10 dB as a reference), I_{dark} is the dark current [mA] (assumed $\sim$1 μA) and I_{noise} is the total RMS input-referred noise current [mA] of the receiver [34, 10, 33, 50, 148, 125]. For simplicity, we also assume that the receiver has enough bandwidth to accommodate the designated bit rate (BW $\geq$ DataRate). The noise current is estimated by setting the 3 dB bandwidth of the RC circuit equal to the data rate [34]. Figure 8.18 shows a plot of the obtained sensitivity for BER = 10^{-12}. The sensitivity is dominated by dark current at low data rates, whereas at high data rates is dominated by the noise current. A sensitivity of -15.5 dBm is then obtained for 10 Gb/s NRZ OOK modulation.

We identify the highest number of channels supported for each modulation rate up to 120 Gb/s (shown in Fig. 8.19). By evaluating the link format ($N_{\lambda} \times r_b$ product), it is found that a chip-to-chip silicon photonic link, under the considered assumptions, can support a peak aggregated bandwidth of $\sim$ 2.1 Tb/s, composed of 47 channels at 45 Gb/s. This corresponds to a channel spacing of 1.06 nm (i.e. $\sim$ 125 GHz) and a modulator spectral shift of 0.53 nm.

The breakdown of the power penalties in this case includes 5.56 dB for the modulator array, 1.6 dB for waveguide loss and coupling at the transmitter side, negligible fiber loss and dispersion effects (considering a short-reach application), 1.2 dB for waveguide loss and coupling at the receiver, 3.5 dB for the optimized demultiplexing array, and 2 dB for extra jitter and polarization dependent penalties. The overall power penalty at 45 Gb/s for 47 channels then sums up to 13.8 dB which is equal to the total power budget per
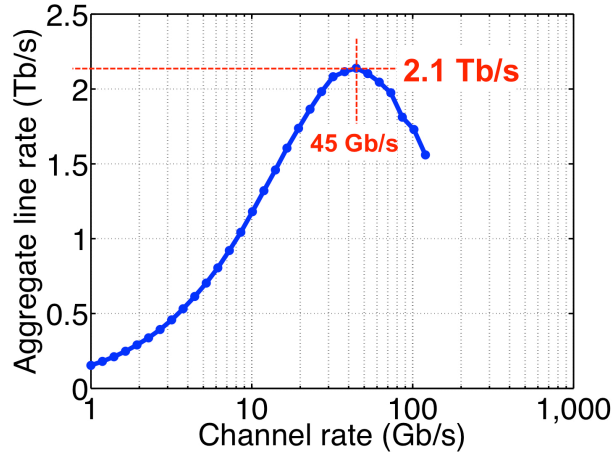
Figure 8.19: Plot of estimated maximum capacity of the link based on the power penalties of the link and the modulation rate of each channel. It is seen that the link can support up to 2.1 Tb/s at a modulation rate of 45 Gb/s.

wavelength (3.4 dBm laser power for each channel, and -10.4 dBm receiver sensitivity).

Overall, this result is to be considered with care, as the generation and detection of 45 Gb/s electric signal might be problematic and energy consuming. Nevertheless, it is a good indicator of the general capabilities of microring-based silicon photonic links from an optical point-of-view.

8.6 Scalability Under Non-Ideal Drive Conditions in 65 nm CMOS

In this section, we improve upon the scalability analysis of the previous section and also add the energy/bit calculations by considering more realistic assumptions on the photonic and electronic elements. The updated list of assumptions are as follows:

1. a power-law relation is now established between the loss of the rings (in dB/cm unit) and the radius (in μm unit) [16]:

$$\alpha_{\mathrm{dB/cm}} \approx a \times R^{-b}.$$ \hfill (8.67)
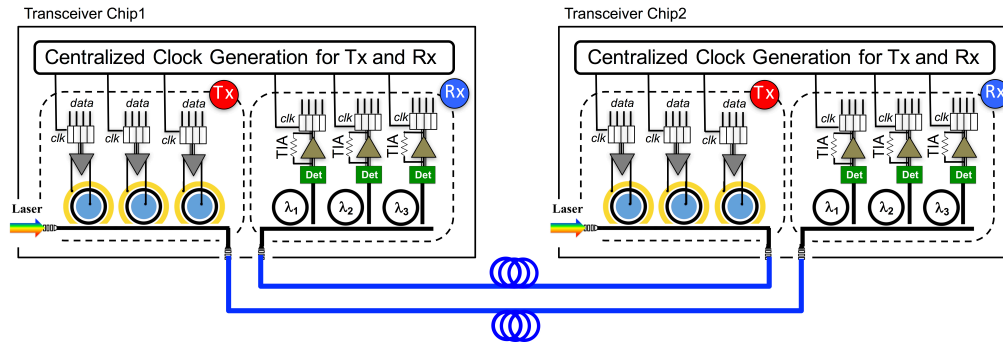
Figure 8.20: Schematic of a full-duplex chip-to-chip optical interconnect based on silicon photonics. Each is equipped with an optical transceiver interface that can support wavelength-division-multiplexing (WDM) optical signaling. The receiver includes wavelength selective filters for demultiplexing.

2. the FSR and loss of a ring resonator are now related to each other through the radius. Therefore, the optical bandwidth of the link (i.e. FSR) is also subject to optimization.

3. modulators *cannot* provide any desired shift of resonance. We assume both PIN and PN modulators are limited to a maximum of 2 nm of spectral shift [206]. PN modulators are chosen in our analysis.

4. Optimization of demultiplexing filters is now carried out based on radius (more precisely FSR) and coupling gap. The gap can only vary from 100 nm to 450 nm.

5. a maximum thermal tuning with 2 nm/mW efficiency is possible for both ring modulators and demultiplexers. The tuning range of resonance is less than 2 nm for both modulators and filters.

6. the driver and receiver circuits are optimized for 65 nm CMOS technology based on recent literature [141, 143, 142] for each given data rate. A database of the optimal power consumption (in mW) is created for the receiver based on not only the data rate but also the input photoelectric current.

Figure 8.20 shows a general scheme of a uni-directional point-to-point link established between two transceivers. Two such links are combined simultaneously to obtain full-duplex capability. No additional clock-data recovery (CDR) mechanism is considered
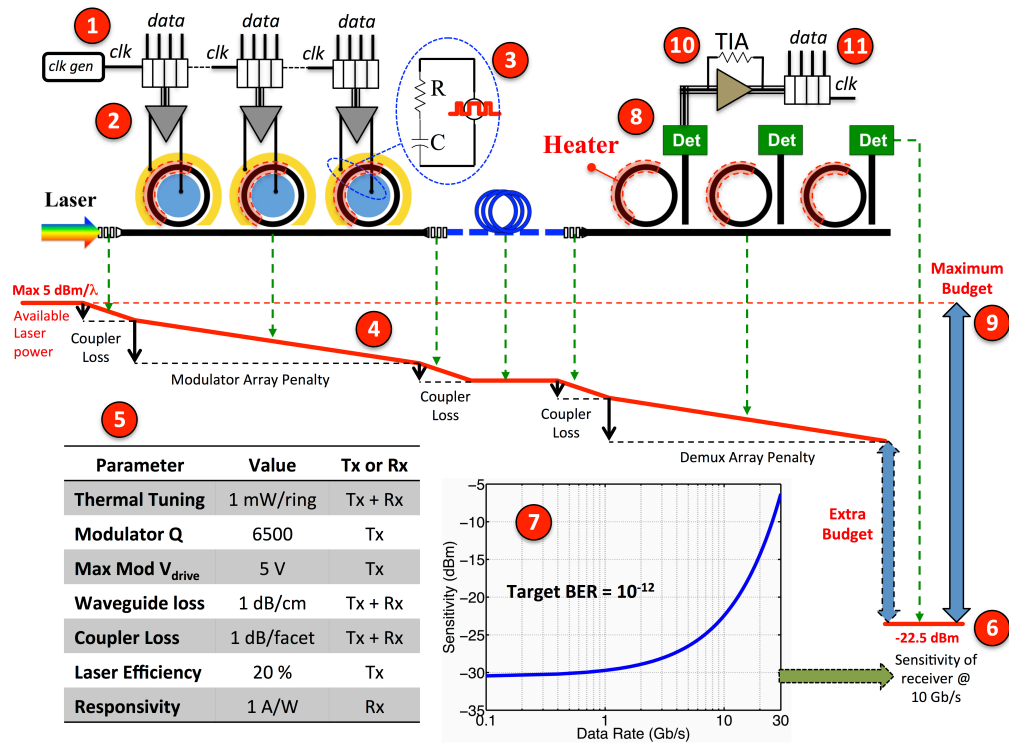
Figure 8.21: Schematic of a unidirectional silicon photonic link. (1) Clock generation and serialization. (2) High-speed drivers. (3) RC representation of optical modulator and thermal tuning. (4) Evolution of optical power budget along the link. (5) Key assumptions for optical devices. (6) Optical sensitivity level of the receiver. (7) Our sensitivity model as a function of data rate for 65 nm CMOS node. (8) Microring demultiplexing array with thermal tuning followed by photodetectors. (9) Maximum available optical power budget. (10) Transimpedance amplifier frontend of the receiver. (11) Deserialization of the electrical data.

inside the Rx unit in our analysis. It is assumed that clock generators located in the two transceivers are fed by a common, closely located centralized clock (another possibility is to synchronize the Rx via a clock forwarding mechanism, with one of the WDM channels dedicated to clock [67]).

Figure 8.21 shows a summary of all the components that go into our modeling of the link. The updated model for the sensitivity of the receiver with an optimal design in 65 nm CMOS shows a sensitivity of -22.5 dBm which has a good improvement (∼7 dB) over the simplistic RC model. In order to estimate the power consumption of the laser source, we assume 20% wall-plug efficiency (15% has been mentioned for on-chip hybrid lasers in

[86, 72], 30% has been considered by Georgas *et al.* [67]).

Carrier depletion PN modulators

Carrier depletion modulators are formed by embedded a PN junction inside the silicon waveguide as shown in 8.22a. The depletion region with a lateral with of L_j is formed due to the balance of positive and negative charges, which also leads to a built-in potential barrier, V_{bi} between the p and n regions. The lateral junction length is a function of the applied reversed voltage. The junction capacitance can be written as

$$C_j(V) = \frac{\epsilon_0 \epsilon_{si}}{L_j(V)} 2\pi R\, h\, s = \frac{C_{j,0}}{\sqrt{1 - V/V_{bi}}} \tag{8.68}$$

where $0 < s < 1$ indicates what fraction of the ring is covered by the PN junction. By changing the applied reverse voltage from V_1 to V_2, the total amount of change in the charge carriers in the waveguide is given by

$$\Delta Q = \int_{V_1}^{V_2} \frac{C_{j,0}}{\sqrt{1 - V/V_{bi}}}\, dv = -2\, C_{j,0}\, V_{bi} \left(\sqrt{1 - \frac{V_2}{V_{bi}}} - \sqrt{1 - \frac{V_1}{V_{bi}}} \right). \tag{8.69}$$

The change of charge per unit volume of the waveguide is the change in the charge densities:

$$\Delta N \approx \Delta P \approx -2\, V_{bi}\, \frac{\epsilon_0 \epsilon_{si}}{L_{j,0}\, w}\, s\, f(V_2, V_1) = -A_0\, f(V_2, V_1) \tag{8.70}$$

where $f(V_2, V_1)$ is the term in parenthesis in Eq. (8.69). Note that if $|V_2| > |V_1|$ then $\Delta Q < 0$ which means that the ring modulator will experience a *red shift* (due to the extraction of carriers). The change in the index of silicon can then be written as

$$\Delta n_{si} \approx A_1\, f(V_2, V_1) \tag{8.71}$$

(a)   (b)
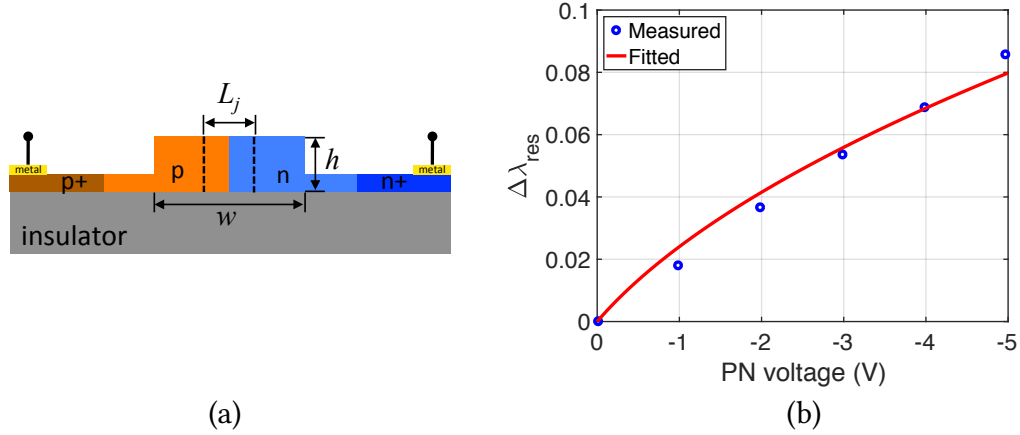
Figure 8.22: (a) Cross section of a carrier depletion PN modulator. (b) DC measurements of the electro-optic efficiency of a PN modulator (data taken from [206]).

and then the relation between the change in the effective index of the optical mode and the index of silicon is given by

$$\Delta n_{\text{eff}} \approx \left(\Gamma_{\text{core}} \frac{n_g}{n_g^{si}} \right) \Delta n_{si} \times \frac{\iint_{\Delta L_j} \mathbf{E} \cdot \mathbf{E}^* dS}{\iint_{\text{core}} \mathbf{E} \cdot \mathbf{E}^* dS} \tag{8.72}$$

according to the perturbation theory. The last term is due to the fact that the index profile of silicon is only changing over the regional change in the junction length. By combining this equation with Eq. (8.49), the DC characteristic of the modulator (for $V_1 = 0$ and $V_2 = -V_{dc}$) is estimated as

$$\Delta \lambda_{\text{res}} \approx a \times \left(-1 + \sqrt{1 + \frac{V_{dc}}{V_{bi}}} \right) . \tag{8.73}$$

By fitting this equation to the measured data points in [206] for $V_{bi} \approx 0.7$ V, we estimate $a = 0.043055$ nm, which constitutes our model for the PN modulators in our analysis. Note that the radius of the modulator in general does not appear in a coefficient, hence this model has a weak dependence on the radius.
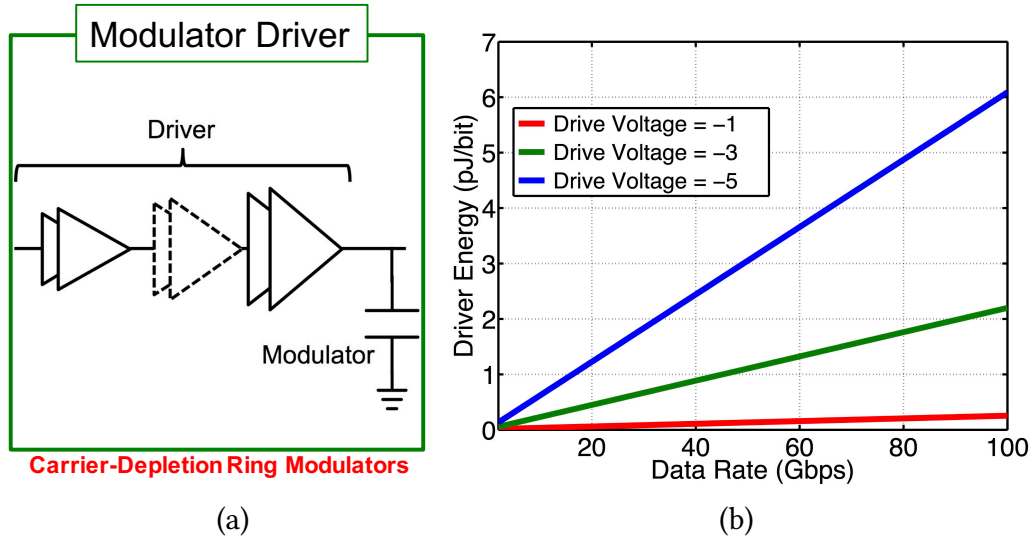
Figure 8.23: (a) Schematic of the high-speed driver with the capacitive load of the optical modulator. (b) Plots of our model for the energy consumption of the high-speed driver as a function of data rate and driving voltage in 65 nm CMOS node. C_{mod} = 50 fF.

High-speed Drivers

The high-speed driver provides fast and large output voltage swing to charge and discharge the optical modulator, assumed here to be a ring modulator operated through carrier depletion mechanism. This is marked by (2) on Fig. 8.21. The driver architecture is based on cascaded amplifier and buffer stages as shown in Fig. 8.23a. The model for the energy consumption of the driver is from the optimized design presented in [143] and is adapted to various capacitive loads and drive voltages based on the following linear equation:

$$E \approx \text{slope} \times DR + \text{constant} \tag{8.74}$$

where E is the energy consumption in pJ/bit unit and DR is data rate in Gbit/sec unit. The slope and constant factors are calculated as

$$\text{slope} = 1.4 \times 10^{-2} \times \left(\frac{V_{mod}}{2V_{DD}}\right)^2 \times \left(\frac{C_{mod}}{C_{\text{ref}}}\right) \tag{8.75a}$$

$$\text{constant} = 8.4 \times 10^{-2} + \frac{1}{4}\left(C_{mod}V_{mod}^2 - C_{\text{ref}}(2V_{DD})^2\right) \times 10^{-3} \tag{8.75b}$$
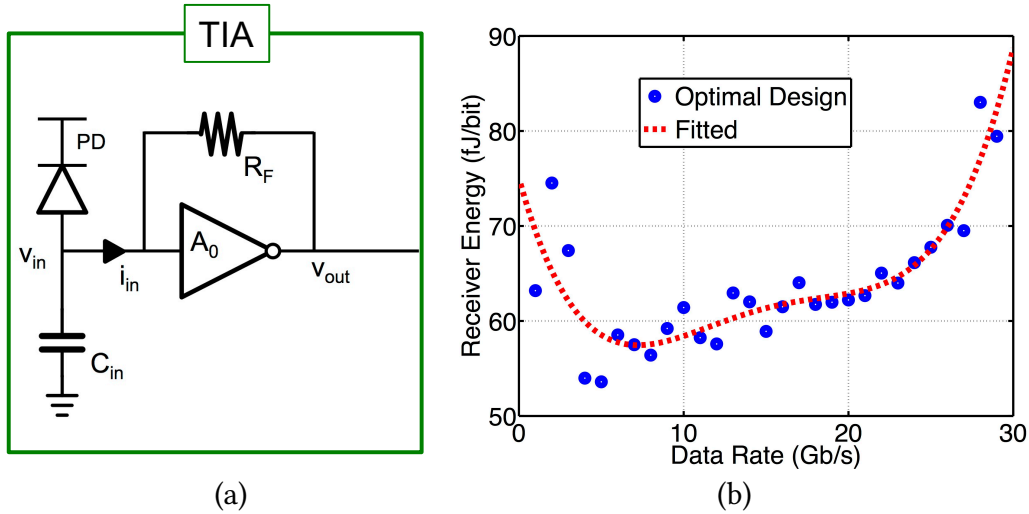
Figure 8.24: (a) Schematic of a photodiode followed by a transimpedance amplifier (TIA). (b) Optimal energy design points of the TIA as a function of data rate at the optical sensitivity level of the receiver.

where V_{DD} = 1.2 V is the supply voltage in 65 nm CMOS, V_{mod} is the peak-to-peak driving voltage applied to the optical modulator, C_{mod} is the equivalent capacitive load of the modulator in fF unit, and C_{ref} = 50 fF is a reference capacitor that has been used in the simulations of the driver. A plot of Eq. (8.75) is presented in Fig. 8.23b for C_{mod} = 50 fF and three drive swings. We set the modulator capacitance to 65 fF in our analysis, which is the approximate junction capacitance of the modulator in [206]. Note that the capacitance of the modulator does not impact the power penalty of the modulator in our analysis.

Receiver

In this analysis, we assume a simple PIN photodiode with a responsivity of 1 A/W and a speed in excess of 45 GHz [50]. The transimpedance amplifier (TIA) is the electrical frontend of the receiver. As shown in Fig. 8.24a, the feedback resistance, R_f, converts the generated photocurrent into a voltage that is fed into the decision circuitry to lift the signal from a small analog voltage to a full digital signal. We model the consumption and capability of the TIA using a database of various designs in 65 nm CMOS [144]. The design that for a given bit-rate and photocurrent (received from the diode) leads to the
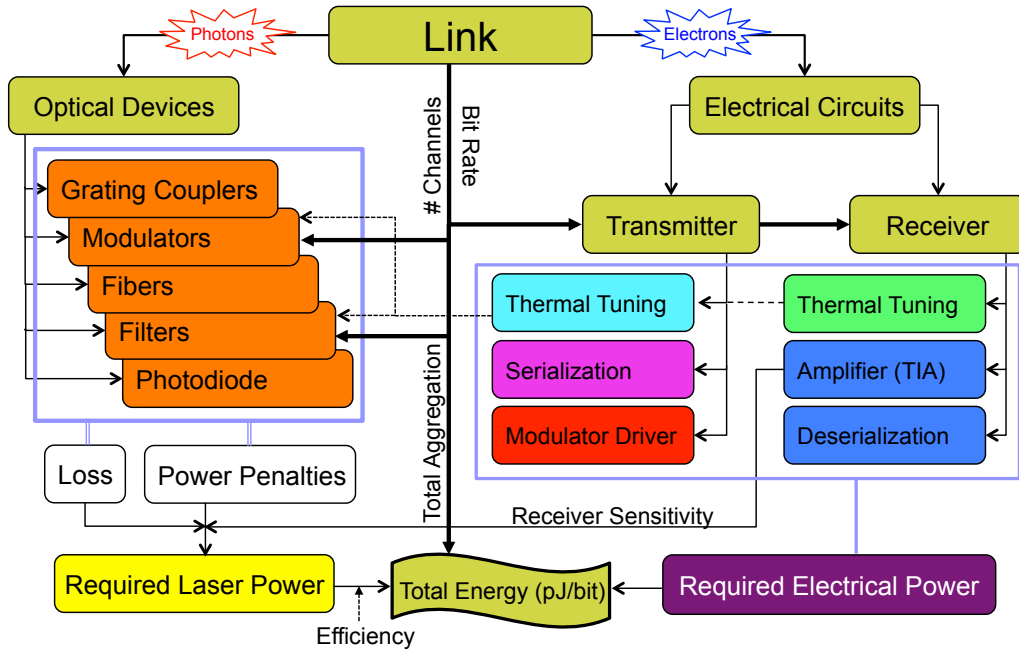
Figure 8.25: Diagram of inter-relations between photonic and electronic elements to estimate the aggregated bandwidth and energy consumption of the link.

lowest power consumption is selected. For example, the data points in Fig. 8.24b indicate the energy consumption of the TIA at the sensitivity level of the receiver for each given data rate. A plot of the sensitivity values (in dBm) as a function of data rate is marked by (7) in Fig. 8.21. Compared to the simple RC model of the receiver, the transimpedance gain and the bandwidth of the receiver are given by

$$|Z_{\text{TIA}}|(f=0) = R_f \times \left| \frac{A_0 - 1}{A_0} \right| \approx R_f$$

$$f_{\text{3dB}} = \frac{|A_0 - 1|}{2\pi R_f C_{in}} = f_{\text{3dB,RC}} \times |A_0 - 1| \,. \tag{8.76}$$

The output of the TIA stages goes to the decision circuits and deserialization stage if electrical data are serialized at the Tx.
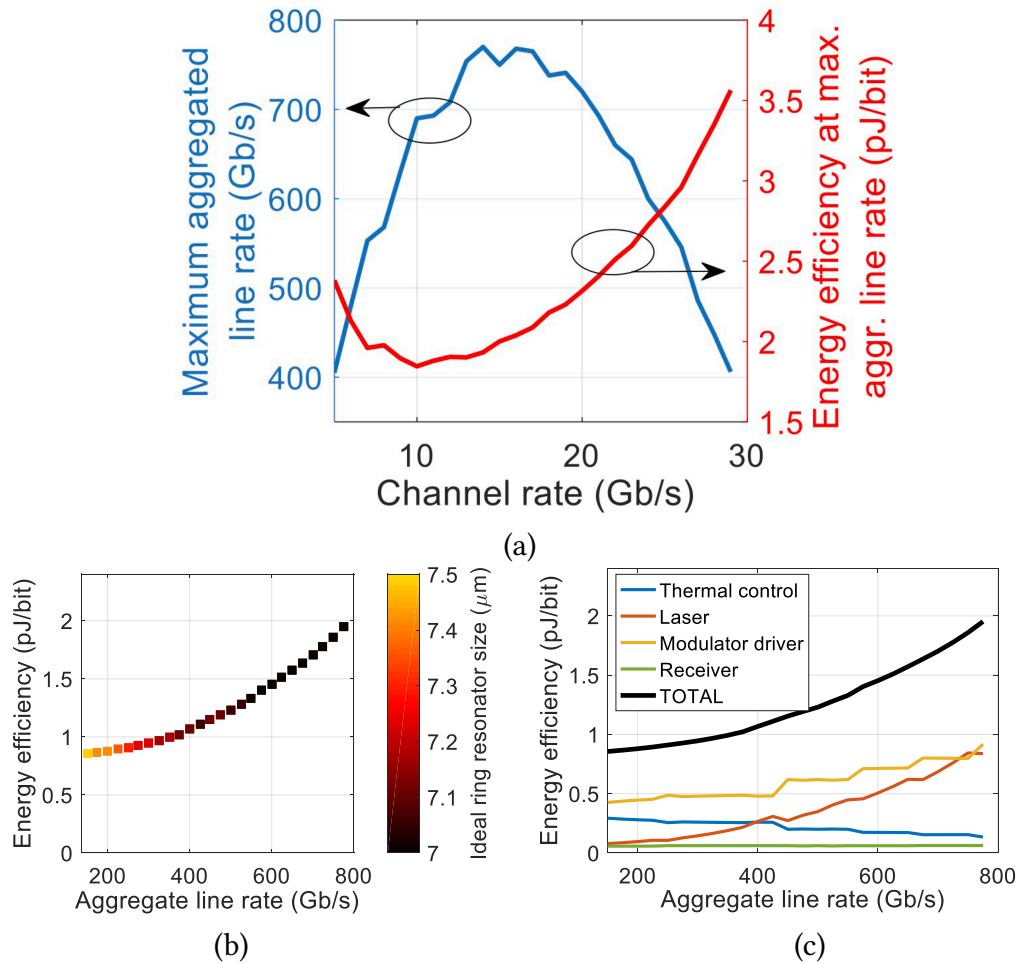
Figure 8.26: (a) Evaluation of the maximum supported aggregation and the associated energy consumption for various channel rates. (b) Minimum energy consumption of the link for given aggregations based on optimum value for the ring radius. (d) Breakdown of energy consumption.

Scalability Analysis

Based on all the electrical and optical models provided in the previous sections, we perform an analysis of the performance of the silicon photonic link. The objective of this analysis is to find the right combination of N_λ (number of channels) and DR (optical rate of each channel). Figure 8.25 shows the connection between photonics elements and electrical elements in the simulations in order to calculate the optical power penalties and the energy/bit metric of the link.

We start by looking for the $N_\lambda \times DR$ product leading to the maximal aggregation

for any given data rate. Here, we sweep the optical data rate of each channel and the FSR of the rings from 10 nm to 50 nm, and increase the number of channels until the available optical power budget is fully utilized. Figure 8.26a plots the results *indicating a maximum possible aggregation of about 765 Gbps at about 17 Gbps channel rate.* The energy consumption associated with each case is also plotted. Note that the minimum point of the energy curve is not at the maximum point of the aggregation curve. This further reiterates the fact that designing a silicon photonic link requires certain electrical and optical tradeoffs.

We then continue by looking for the $N_\lambda \times DR$ product leading to the minimal energy consumption for a given optical aggregation rate for the link. Figure 8.26b shows the minimal energy for 200 Gbps to 800 Gbps aggregation based on the optimization of the radius and gaps of the rings. The rings are always set to their critical coupling operation point. This chart shows that ring radius of about 7μm (hence FSR $\approx$ 12.45 nm) leads to the best energy performance of the link. The breakdown of the energy numbers is plotted in Fig. 8.26c indicating that the laser consumption becomes critical at higher data rates while the static thermal tuning has a declining trend.

As expected, the results shown here differ substantially from the ones reported in the previous section (maximal aggregation rate of 2.1 Tb/s). The main two reasons are the dependence of the loss of the rings on the radius (hence FSR) and the low electro-optic efficiency of depletion-based modulators (although depletion type modulators can operate at higher speeds than their injection type counterparts).

8.7   Discussion

This chapter only focused on the simplest optical modulation format, that is OOK. However, currently the optical interconnects (especially for datacenters) are moving towards PAM-4 and higher-order modulation formats. Compared to OOK, PAM-4 signaling will

save a factor of 2 on the spectral efficiency (in terms bits/hz), hence reducing the inter-channel spectral crosstalk for WDM signaling. However, PAM-4 suffers from a worse optical SNR compared to OOK because the four levels are closer to each other and open-ness of the eyes is smaller than OOK. Overall, we believe that PAM-4 is a viable choice to increase the throughput limit of a WDM link with OOK signaling, but at the cost of higher energy/bit.
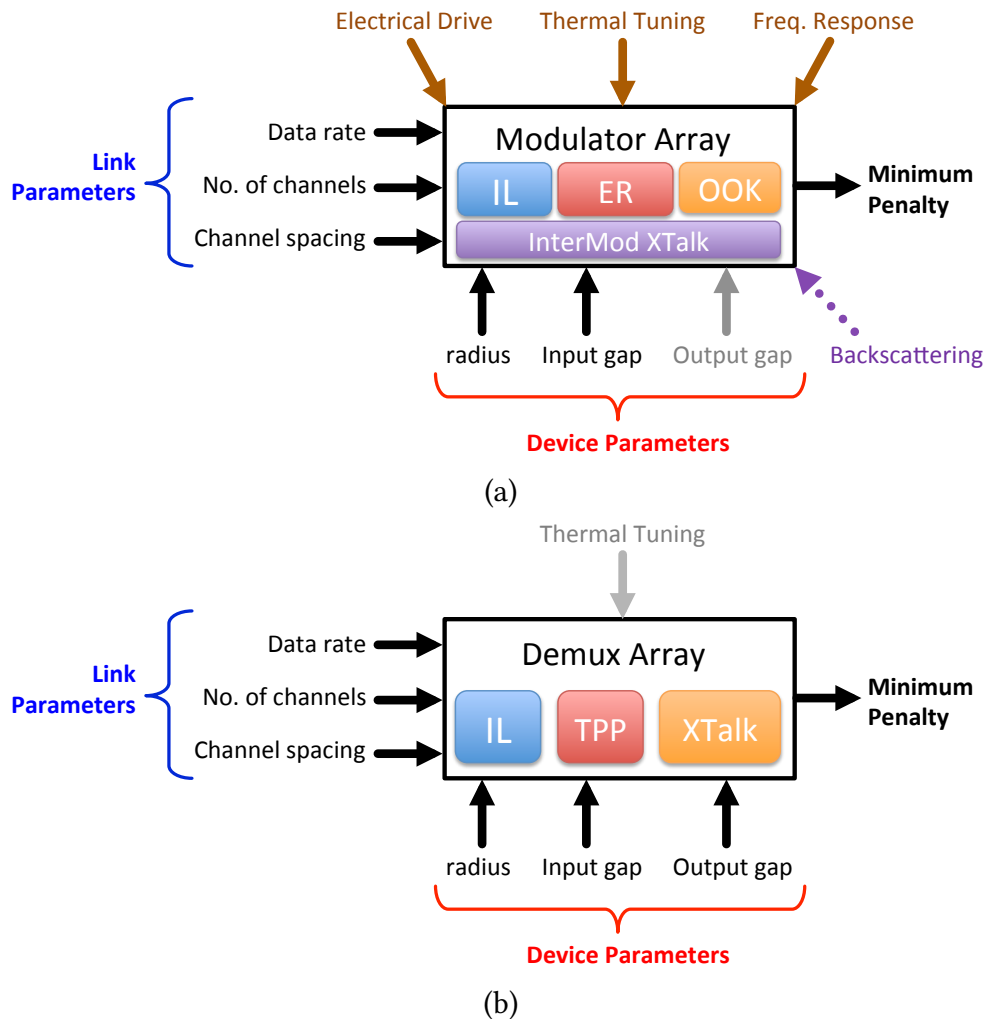
Figure 8.27: (a) Block diagram of power penalty evaluation for modulators. (b) block diagram of power penalty evaluation for demultiplexing filters.

8.8 Summary of the Chapter

In this chapter, we first presented a rigorous mathematical frameworks for characterizing the optical power penalties of optical components (modulators, filters, etc.) of a digital communication link. We then specifically characterized the power penalties of microring-based modulators and filters, and used the results to show that the scalability of a silicon photonic link is maximally limited to 2.1 Tb/s under ideal drive conditions (i.e. photonics is assumed to be the bottleneck). We then went further on to include more realistic assumptions by coupling the loss of the rings to their radius (and hence the FSR) based on the design space exploration of the SOI microrings presented in the previous chapters. We also updated the receiver sensitivity model with a database of simulation results for 65 nm CMOS technology, and analyzed the link both in terms of energy efficiency and maximum throughput.

Figure 8.27a and 8.27b show a summary of the implemented procedure for calculating the power penalty of a ring-based modulator array and demux array as a function of link parameters and device parameters. Future work will need to add the impact of the backscattering and the electrical frequency response (in case the modulators are not limited by their photon life time) into the power penalty calculations. Although the model for self-heating and bistability of the modulators was already developed, it has not been directly included in the power penalty calculations, rather it is used as a limiting factor on the optical power per channel.

Conclusions

In conclusion, the studies carried out in this thesis have two major implications:

1. silicon photonic switch fabrics based on the electro-optic phase shifters, although faster than their thermo-optic counterparts, are limited in their scalability due the inherent excess loss and crosstalk of the electro-optic operation. However, by taking advantage of the repetitions in the number of switch configurations, optimal routing tables can be constructed for moderately low-radix switches. This leads to various benefits such as avoiding routing paths with high power penalties and also equalized performance on all the input-output routing paths.

2. silicon photonic links based on microring resonators can provide data aggregations in the range of 0.5–1 Tb/s at energy efficiencies in the range of 1.5–3 pJ/bit. While the design space of MRRs is not hard to explore for the optimum designs, significant amount of engineering effort still needs to accompany that to overcome challenges such as thermal sensitivity of MRRs, optical absorption and self heating, fabrication variations, and backscattering effects. The potential of silicon photonics for monolithic integration with CMOS electronics is driving engineers to find the best solutions for all the problems that MRRs are facing.

Bibliography

[1] Erik Agrell et al. "Roadmap of optical communications." In: *Journal of Optics* 18.6 (2016), p. 063002.

[2] Ryan Aguinaldo et al. "Wideband silicon-photonic thermo-optic switch in a wavelength-division multiplexed ring network." In: *Optics express* 22.7 (2014), pp. 8205–8218.

[3] J Ahn et al. "Devices and architectures for photonic chip-scale integration." In: *Applied Physics A* 95.4 (2009), pp. 989–997.

[4] T Alasaarela et al. "Reduced propagation loss in silicon strip and slot waveguides coated by atomic layer deposition." In: *Optics express* 19.12 (2011), pp. 11529–11538.

[5] Luca Alloatti et al. "Photonics design tool for advanced CMOS nodes." In: *IET optoelectronics* 9.4 (2015), pp. 163–167.

[6] Reja Amatya et al. "Low power thermal tuning of second-order microring resonators." In: *Lasers and Electro-Optics, 2007. CLEO 2007. Conference on.* IEEE. 2007, pp. 1–2.

[7] Yasuhiko Arakawa et al. "Silicon photonics for next generation system integration platform." In: *Communications Magazine, IEEE* 51.3 (2013), pp. 72–77.

[8] Mehdi Asghari. "Silicon photonics: a low cost integration platform for datacom and telecom applications." In: *National Fiber Optic Engineers Conference.* Optical Society of America. 2008, NThA4.

[9] Solomon Assefa et al. "A 90nm CMOS integrated nano-photonics technology for 25Gbps WDM optical communications applications." In: *Electron Devices Meeting (IEDM), 2012 IEEE International.* IEEE. 2012, pp. 33–8.

[10] Solomon Assefa et al. "CMOS-integrated optical receivers for on-chip interconnects." In: *Selected Topics in Quantum Electronics, IEEE Journal of* 16.5 (2010), pp. 1376–1385.

[11]   AH Atabaki et al. "Optimization of metallic microheaters for high-speed reconfigurable silicon photonics." In: *Optics express* 18.17 (2010), pp. 18312–18323.

[12]   R Baets and PE Lagasse. "Calculation of radiation loss in integrated-optic tapers and Y-junctions." In: *Applied optics* 21.11 (1982), pp. 1972–1978.

[13]   M Bahadori et al. "Comprehensive design space exploration of silicon photonic interconnects." In: *Journal of Lightwave Technology* (2015).

[14]   Meisam Bahadori et al. "Comprehensive design space exploration of silicon photonic interconnects." In: *Journal of Lightwave Technology* 34.12 (2016), pp. 2975–2987.

[15]   Meisam Bahadori et al. "Crosstalk penalty in microring-based silicon photonic interconnect systems." In: *Journal of Lightwave Technology* 34.17 (2016), pp. 4043–4052.

[16]   Meisam Bahadori et al. "Design Space Exploration of Microring Resonators in Silicon Photonic Interconnects: Impact of the Ring Curvature." In: *Journal of Lightwave Technology* (2018).

[17]   Meisam Bahadori et al. "Energy-bandwidth design exploration of silicon photonic interconnects in 65nm CMOS." In: *IEEE Optical Interconnects Conference (OI), 2016.* IEEE. 2016, pp. 2–3.

[18]   Meisam Bahadori et al. "Energy-performance optimized design of silicon photonic interconnection networks for high-performance computing." In: *Proceedings of the Conference on Design, Automation & Test in Europe.* European Design and Automation Association. 2017, pp. 326–331.

[19]   Meisam Bahadori et al. "Loss and crosstalk of scalable MZI-based switch topologies in silicon photonic platform." In: *Photonics Conference (IPC), 2016 IEEE.* IEEE. 2016, pp. 615–616.

[20]   Meisam Bahadori et al. "Optimization of Microring-based Filters for Dense WDM Silicon Photonic Interconnects." In: *Measurement* 4 (), p. 5.

[21]   Meisam Bahadori et al. "Optimization of microring-based filters for dense WDM silicon photonic interconnects." In: *Optical Interconnects Conference (OI), 2015 IEEE.* IEEE. 2015, pp. 84–85.

[22]   Meisam Bahadori et al. "Thermal Rectification of Integrated Microheaters for Microring Resonators in Silicon Photonics Platform." In: *Journal of Lightwave Technology* (2017).

[23]   Meisam Bahadori et al. "Thermal Rectification of Integrated Microheaters for Microring Resonators in Silicon Photonics Platform." In: *Journal of Lightwave Technology* 36.3 (2018), pp. 773–788.

[24]   Christopher Batten et al. "Building many-core processor-to-DRAM networks with monolithic CMOS silicon photonics." In: *IEEE Micro* 29.4 (2009).

[25]   Glenn Baxter et al. "Highly programmable wavelength selective switch based on liquid crystal on silicon switching elements." In: *Optical Fiber Communication Conference, 2006 and the 2006 National Fiber Optic Engineers Conference. OFC 2006.* IEEE. 2006, 3–pp.

[26]   Hamed Pishvai Bazargani et al. "Microring resonator design with improved quality factors using quarter Bezier curves." In: *Lasers and Electro-Optics (CLEO), 2015 Conference on.* IEEE. 2015, pp. 1–2.

[27]   Aleksandr Biberman et al. "Broadband CMOS-compatible silicon photonic electro-optic switch for photonic networks-on-chip." In: *Conference on Lasers and Electro-Optics.* Optical Society of America. 2010, CPDA11.

[28]   Aleksandr Biberman et al. "Silicon microring resonator-based broadband comb switch for wavelength-parallel message routing." In: *LEOS 2007* (2007).

[29]   Alberto A Bilotti. "Static temperature distribution in IC chips with isothermal heat sources." In: *IEEE Transactions on Electron Devices* 21.3 (1974), pp. 217–226.

[30]   Wim Bogaerts and Shankar Kumar Selvaraja. "Compact single-mode silicon hybrid rib/strip waveguide with adiabatic bends." In: *IEEE Photonics Journal* 3.3 (2011), pp. 422–432.

[31]   Wim Bogaerts et al. "Silicon microring resonators." In: *Laser & Photonics Reviews* 6.1 (2012), pp. 47–73.

[32]   Wim Bogaerts et al. "Silicon-on-insulator spectral filters fabricated with CMOS technology." In: *IEEE journal of selected topics in quantum electronics* 16.1 (2010), pp. 33–44.

[33]   John E Bowers et al. "High-gain high-sensitivity resonant Ge/Si APD photodetectors." In: *SPIE Defense, Security, and Sensing.* International Society for Optics and Photonics. 2010, 76603H–76603H.

[34]   James F Buckwalter et al. "A monolithic 25-Gb/s transceiver with photonic ring modulators and Ge detectors in a 130-nm CMOS SOI process." In: *Solid-State Circuits, IEEE Journal of* 47.6 (2012), pp. 1309–1322.

[35] David M Calhoun et al. "Hardware–Software Integrated Silicon Photonics for Computing Systems." In: *Silicon Photonics III*. Springer, 2016, pp. 157–189.

[36] Giovanna Calò, Antonella D'Orazio, and Vincenzo Petruzzelli. "Broadband Mach–Zehnder switch for photonic networks on chip." In: *Journal of Lightwave Technology* 30.7 (2012), pp. 944–952.

[37] D Celo et al. "Optical proximity correction in geometry sensitive silicon photonics waveguide crossings." In: *Group IV Photonics (GFP), 2017 IEEE 14th International Conference on*. IEEE. 2017, pp. 45–46.

[38] Chin-Hui Chen et al. "A comb laser-driven DWDM silicon photonic transmitter based on microring modulators." In: *Optics express* 23.16 (2015), pp. 21541–21548.

[39] Chin-Hui Chen et al. "DWDM Silicon Photonic Transceivers for Optical Interconnect." In: *Journal of Solid-State Circuits (JSSC)* 49.6 (2014).

[40] Qixiang Cheng, Meisam Bahadori, and Keren Bergman. "Advanced path mapping for silicon photonic switch fabrics." In: *CLEO: Science and Innovations*. Optical Society of America. 2017, SW1O–5.

[41] Qixiang Cheng et al. "Highly-scalable, low-crosstalk architecture for ring-based optical space switch fabrics." In: *IEEE Optical Interconnects Conference (OI), 2017*. IEEE. 2017, pp. 41–42.

[42] Qixiang Cheng et al. "Smart routing tables for integrated photonic switch fabrics." In: *Proc. 43rd Eur. Conf. Exhib. Opt. Commun.* 2017.

[43] Gary Cook and Jodie Van Horn. "How dirty is your data? A look at the energy choices that power cloud computing." In: *Greenpeace (April 2011)* (2011).

[44] Paul W Coteus et al. "Technologies for exascale systems." In: *IBM Journal of Research and Development* 55.5 (2011), pp. 14–1.

[45] Dharanidhar Dang et al. "Mode-Division-Multiplexed Photonic Router for High Performance Network-on-Chip." In: *VLSI Design (VLSID), 2015 28th International Conference on*. IEEE. 2015, pp. 111–116.

[46] Peter De Dobbelaere et al. "Digital MEMS for optical switching." In: *IEEE Communications magazine* 40.3 (2002), pp. 88–95.

[47] Casimer DeCusatis. "Optical interconnect networks for data communications." In: *Journal of Lightwave Technology* 32.4 (2014), pp. 544–552.

[48] Christopher T DeRose et al. "A CMOS compatible external heater-modulator." In: *Optical Interconnects Conference, 2014 IEEE*. IEEE. 2014, pp. 17–18.

[49] Christopher T DeRose et al. "Low power and broadband 2×2 silicon thermo-optic switch." In: *Optical Fiber Communication Conference*. Optical Society of America. 2011, OThM3.

[50] Christopher T DeRose et al. "Ultra compact 45 GHz CMOS compatible Germanium waveguide photodiode with low dark current." In: *Optics express* 19.25 (2011), pp. 24897–24904.

[51] Po Dong et al. "High-speed and compact silicon modulator based on a racetrack resonator with a 1 V drive voltage." In: *Optics letters* 35.19 (2010), pp. 3246–3248.

[52] Po Dong et al. "Low Vpp, ultralow-energy, compact, high-speed silicon electro-optic modulator." In: *Optics Express* 17.25 (2009), pp. 22484–22490.

[53] Po Dong et al. "Low power and compact reconfigurable multiplexing devices based on silicon microring resonators." In: *Optics express* 18.10 (2010), pp. 9852–9858.

[54] John D Downie. "Relationship of Q penalty to eye-closure penalty for NRZ and RZ signals with signal-dependent noise." In: *Journal of Lightwave Technology* 23.6 (2005), pp. 2031–2038.

[55] John D Downie and A Boh Ruffin. "Analysis of signal distortion and crosstalk penalties induced by optical filters in optical networks." In: *Journal of lightwave technology* 21.9 (2003), p. 1876.

[56] Luan HK Duong et al. "A case study of signal-to-noise ratio in ring-based optical networks-on-chip." In: *Design & Test, IEEE* 31.5 (2014), pp. 55–65.

[57] Luan HK Duong et al. "Coherent crosstalk noise analyses in ring-based optical interconnects." In: *Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition*. EDA Consortium. 2015, pp. 501–506.

[58] Nicolas Dupuis. "Technologies for fast, scalable silicon photonic switches." In: *Photonics in Switching (PS), 2015 International Conference on*. IEEE. 2015, pp. 100–102.

[59] Nicolas Dupuis et al. "Design and Fabrication of Low-Insertion-Loss and Low-Crosstalk Broadband 2 × 2 Mach–Zehnder Silicon Photonic Switches." In: *Journal of Lightwave Technology* 33.17 (2015), pp. 3597–3606.

[60] Nicolas Dupuis et al. "Ultralow crosstalk nanosecond-scale nested 2× 2 Mach–Zehnder silicon photonic switch." In: *Optics letters* 41.13 (2016), pp. 3002–3005.

[61]   Louay A Eldada. "Advances in telecom and datacom optical components." In: *Optical Engineering* 40.7 (2001), pp. 1165–1179.

[62]   Fabrizio Gambini et al. "BER evaluation of a low-crosstalk silicon integrated multi-microring network-on-chip." In: *Optics express* 23.13 (2015), pp. 17169–17178.

[63]   Elsa Garmire et al. *Integrated optics*. Vol. 7. Springer Science & Business Media, 2013.

[64]   Rajdeep Gautam et al. "Single silicon microring resonator Mach-Zehnder modulator with low-power consumption using thermo-optic effect." In: *Photonics Conference (IPC), 2012 IEEE*. IEEE. 2012, pp. 834–835.

[65]   Alexander Gazman et al. "Automated Thermal Stabilization of Cascaded Silicon Photonic Ring Resonators for Reconfigurable WDM Applications." In: *ECOC* (2017).

[66]   Alexander Gazman et al. "Software-defined control-plane for wavelength selective unicast and multicast of optical data in a silicon photonic platform." In: *Optics express* 25.1 (2017), pp. 232–242.

[67]   Michael Georgas et al. "Addressing link-level design tradeoffs for integrated photonic interconnects." In: *Custom Integrated Circuits Conference (CICC), 2011 IEEE*. IEEE. 2011, pp. 1–8.

[68]   Albert Greenberg et al. "VL2: a scalable and flexible data center network." In: *ACM SIGCOMM computer communication review*. Vol. 39. 4. ACM. 2009, pp. 51–62.

[69]   Biswajeet Guha, Jaime Cardenas, and Michal Lipson. "Athermal silicon microring resonators with titanium oxide cladding." In: *Optics express* 21.22 (2013), pp. 26557–26563.

[70]   Nicholas C Harris et al. "Efficient, compact and low loss thermo-optic phase shifter in silicon." In: *Optics express* 22.9 (2014), pp. 10487–10493.

[71]   Mikhail Haurylau et al. "On-chip optical interconnect roadmap: challenges and critical directions." In: *Selected Topics in Quantum Electronics, IEEE Journal of* 12.6 (2006), pp. 1699–1705.

[72]   Martijn JR Heck and John E Bowers. "Energy efficient and energy proportional optical interconnects for multi-core processors: Driving the need for on-chip sources." In: *IEEE Journal of Selected Topics in Quantum Electronics* 20.4 (2014), pp. 332–343.

[73]   Robert Hendry et al. "Modeling and evaluation of chip-to-chip scale silicon pho-
        tonic networks." In: *High-Performance Interconnects (HOTI), 2014 IEEE 22nd Annual
        Symposium on*. IEEE. 2014, pp. 1–8.

[74]   Robert Hendry et al. "Physical layer analysis and modeling of silicon photonic
        WDM bus architectures." In: *HiPEAC 2014 workshop, January*. 2014.

[75]   Yasuhiro Hida, Hidekatsu Onose, and Saburo Imamura. "Polymer waveguide ther-
        mooptic switch with low electric power consumption at 1.3 μm." In: *IEEE Photonics
        Technology Letters* 5.7 (1993), pp. 782–784.

[76]   Ting Hu et al. "Silicon photonic network-on-chip and enabling components." In:
        Science China Technological Sciences 56.3 (2013), pp. 543–553.

[77]   Wei-Ping Huang, Lin Han, and Jianwei Mu. "A rigorous circuit model for simu-
        lation of large-scale photonic integrated circuits." In: *IEEE Photonics Journal* 4.5
        (2012), pp. 1622–1638.

[78]   Daryl Inniss and Roy Rubenstein. *Silicon Photonics: Fueling the Next Information
        Revolution*. Morgan Kaufmann, 2016.

[79]   Ezra Ip and Joseph M Kahn. "Power spectra of return-to-zero optical signals." In:
        Journal of lightwave technology 24.3 (2006), p. 1610.

[80]   Hasitha Jayatilleka et al. "Crosstalk Limitations of Microring-Resonator Based
        WDM Demultiplexers on SOI." In: *Optical Interconnects conference* (2015).

[81]   Hasitha Jayatilleka et al. "Crosstalk in SOI Microring Resonator-Based Filters." In:
        Journal of Lightwave Technology (2015 [in press]).

[82]   Hasitha Jayatilleka et al. "Crosstalk limitations of microring-resonator based
        WDM demultiplexers on SOI." In: *Optical Interconnects Conference (OI), 2015 IEEE*.
        IEEE. 2015, pp. 48–49.

[83]   Hasitha Jayatilleka et al. "Intraband crosstalk of SOI microring resonator-based
        optical add-drop multiplexers." In: *Group IV Photonics (GFP), 2015 IEEE 12th Inter-
        national Conference on*. IEEE. 2015, pp. 3–4.

[84]   Kuniaki Kakihara et al. "Full-vectorial finite element method in a cylindrical co-
        ordinate system for loss analysis of photonic wire bends." In: *Optics express* 14.23
        (2006), pp. 11128–11141.

[85]   FP Kapron, DB Keck, and RD Maurer. "Radiation losses in glass optical waveg-
        uides." In: *Applied Physics Letters* 17.10 (1970), pp. 423–425.

[86] Brian R Koch et al. "Integrated silicon photonic laser sources for telecom and datacom." In: *National Fiber Optic Engineers Conference*. Optical Society of America. 2013, PDP5C–8.

[87] J Komma et al. "Thermo-optic coefficient of silicon at 1550 nm and cryogenic temperatures." In: *Applied Physics Letters* 101.4 (2012), p. 041905.

[88] Ashok V Krishnamoorthy et al. "Computer systems based on silicon photonic interconnects." In: *Proceedings of the IEEE* 97.7 (2009), pp. 1337–1361.

[89] Cedric F Lam. "Optical network technologies for datacenter networks." In: *Optical Fiber Communication (OFC), collocated National Fiber Optic Engineers Conference, 2010 Conference on (OFC/NFOEC)*. IEEE. 2010, pp. 1–3.

[90] Patrick Le Maître et al. "Impact of process variability of active ring resonators in a 300mm silicon photonic platform." In: *Optical Communication (ECOC), 2015 European Conference on*. IEEE. 2015, pp. 1–3.

[91] Benjamin G Lee et al. "Transmission of high-data-rate optical signals through a micrometer-scale silicon ring resonator." In: *Optics letters* 31.18 (2006), pp. 2701–2703.

[92] Kevin K Lee et al. "Fabrication of ultralow-loss Si/SiO 2 waveguides by roughness reduction." In: *Optics letters* 26.23 (2001), pp. 1888–1890.

[93] Ang Li and Wim Bogaerts. "Fundamental suppression of backscattering in silicon microrings." In: *Optics Express* 25.3 (2017), pp. 2092–2099.

[94] Ang Li et al. "Backscattering in silicon microring resonators: a quantitative analysis." In: *Laser & Photonics Reviews* 10.3 (2016), pp. 420–431.

[95] Cuiting Li et al. "Silicon-microring-based thermo-optic non-blocking four-port optical router for optical networks-on-chip." In: *Optical and Quantum Electronics* 48.12 (2016), p. 552.

[96] Guoliang Li et al. "Ring resonator modulators in silicon for interchip photonic links." In: *IEEE Journal of Selected Topics in Quantum Electronics* 19.6 (2013), pp. 95–113.

[97] Jun Li et al. "Scaling trends for picojoule-per-bit WDM photonic interconnects in CMOS SOI and FinFET processes." In: *Journal of Lightwave Technology* 34.11 (2016), pp. 2730–2742.

[98]   Qi Li et al. "Experimental characterization of the optical-power upper bound in a silicon microring modulator." In: *IEEE Optical Interconnects Conf.* Citeseer. 2012, pp. 38–39.

[99]   Qi Li et al. "Single Microring-Based Silicon Photonic Crossbar Switches." In: *Photonics Technology Letters, IEEE* 27.18 (2015), pp. 1981–1984.

[100]   Xianyao Li et al. "Fast and efficient silicon thermo-optic switching based on reverse breakdown of pn junction." In: *Optics letters* 39.4 (2014), pp. 751–753.

[101]   Peicheng Liao et al. "Ultradense silicon photonic interface for optical interconnection." In: *IEEE Photonics Technology Letters* 27.7 (2015), pp. 725–728.

[102]   Andy Eu-Jin Lim et al. "Review of silicon photonics foundry efforts." In: *IEEE Journal of Selected Topics in Quantum Electronics* 20.4 (2014), pp. 405–416.

[103]   Sen Lin et al. "Electronic–Photonic Co-Optimization of High-Speed Silicon Photonic Transmitters." In: *Journal of Lightwave Technology* 35.21 (2017), pp. 4766–4780.

[104]   Brent E Little, Juha-Pekka Laine, and Sai T Chu. "Surface-roughness-induced contradirectional coupling in ring and disk resonators." In: *Optics letters* 22.1 (1997), pp. 4–6.

[105]   Brent E Little et al. "Microring resonator channel dropping filters." In: *Journal of lightwave technology* 15.6 (1997), pp. 998–1005.

[106]   Frankie Y Liu et al. "10-Gbps, 5.3-mW optical transmitter and receiver circuits in 40-nm CMOS." In: *IEEE journal of solid-state circuits* 47.9 (2012), pp. 2049–2067.

[107]   Ke Liu et al. "Two-dimensional design and analysis of trench-coupler based Silicon Mach-Zehnder thermo-optic switch." In: *Optics express* 24.14 (2016), pp. 15845–15853.

[108]   Liangjun Lu et al. "16×16 non-blocking silicon optical switch based on electro-optic Mach-Zehnder interferometers." In: *Optics express* 24.9 (2016), pp. 9295–9307.

[109]   Zeqin Lu et al. "Performance prediction for silicon photonics integrated circuits with layout-dependent correlated manufacturing variability." In: *Optics Express* 25.9 (2017), pp. 9712–9733.

[110]   Yangjin Ma et al. "Optimized silicon photonic components for high-performance interconnect systems." In: *Photonics Conference (IPC), 2015.* IEEE. 2015, pp. 353–354.

[111] Yangjin Ma et al. "Ultralow loss single layer submicron silicon waveguide crossing for SOI optical interconnect." In: *Optics express* 21.24 (2013), pp. 29374–29382.

[112] Andri Mahendra et al. "Multiwavelength stabilization control of a thermo-optic system with adaptive reconfiguration." In: *Applied Optics* 56.4 (2017), pp. 1113–1118.

[113] Jason CC Mak et al. "Automatic resonance alignment of high-order microring filters." In: *IEEE Journal of Quantum Electronics* 51.11 (2015), pp. 1–11.

[114] CL Manganelli et al. "Design of coupled micro-ring resonators for silicon photonic switching matrices." In: *IEEE Optical Interconnects Conference (OI), 2016*. IEEE. 2016, pp. 84–85.

[115] CL Manganelli et al. "Large-FSR thermally tunable double-ring filters for WDM applications in silicon photonics." In: *IEEE Photonics Journal* 9.1 (2017), pp. 1–10.

[116] Riccardo Marchetti et al. "Low-loss micro-resonator filters fabricated in silicon by CMOS-compatible lithographic techniques: design and characterization." In: *Applied Sciences* 7.2 (2017), p. 174.

[117] Gianlorenzo Masini et al. "CMOS photonics for optical engines and interconnects." In: *Optical Fiber Communication Conference*. Optical Society of America. 2012, OTu2I–1.

[118] Adil Masood et al. "Comparison of heater architectures for thermal control of silicon photonic circuits." In: *Group IV Photonics (GFP), 2013 IEEE 10th International Conference on*. IEEE. 2013, pp. 83–84.

[119] Hiroyuki Matsuura et al. "Accelerating Switching Speed of Thermo-optic MZI Silicon-Photonic Switches with "Turbo Pulse" in PWM Control." In: *Optical Fiber Communication Conference*. Optical Society of America. 2017, W4E–3.

[120] WR McKinnon et al. "Extracting coupling and loss coefficients from a ring resonator." In: *Optics express* 17.21 (2009), pp. 18971–18982.

[121] David AB Miller. "Device requirements for optical interconnects to silicon chips." In: *Proceedings of the IEEE* 97.7 (2009), pp. 1166–1185.

[122] Giovanni Monegato. "The numerical evaluation of one-dimensional Cauchy principal value integrals." In: *Computing* 29.4 (1982), pp. 337–354.

[123] F Morichetti et al. "Coherent backscattering in optical microring resonators." In: *Applied Physics Letters* 96.8 (2010), p. 081112.

[124] Francesco Morichetti et al. "Roughness induced backscattering in optical silicon waveguides." In: *Physical review letters* 104.3 (2010), p. 033902.

[125] Masoud Honarvar Nazari and Azita Emami-Neyestanak. "A 24-Gb/s double-sampling receiver for ultra-low-power optical communication." In: *Solid-State Circuits, IEEE Journal of* 48.2 (2013), pp. 344–357.

[126] Milos Nedeljkovic et al. "Mid-infrared thermo-optic modulators in SoI." In: *IEEE Photon. Technol. Lett.* 26.13 (2014), pp. 1352–1355.

[127] Lynn Nelson and Herwig Kogelnik. "Coherent crosstalk impairments in polarization multiplexed transmission due to polarization mode dispersion." In: *Optics express* 7.10 (2000), pp. 350–361.

[128] Mahdi Nikdast et al. "Chip-scale silicon photonic interconnects: A formal study on fabrication non-uniformity." In: *Journal of Lightwave Technology* 34.16 (2016), pp. 3682–3695.

[129] Dessislava Nikolova et al. "A highly scalable fully non-blocking silicon photonic switch fabric." In: *arXiv preprint arXiv:1512.09323* (2015).

[130] Dessislava Nikolova et al. "Scaling silicon photonic switch fabrics for data center interconnection networks." In: *Optics express* 23.2 (2015), pp. 1159–1175.

[131] Donghai Niu et al. "Optimized design and fabrication of polymer/silica thermo-optic switch with low power consumption." In: *Applied Optics* 56.21 (2017), pp. 5799–5803.

[132] Noam Ophir et al. "Silicon photonic microring links for high-bandwidth-density, low-power chip I/O." In: *IEEE Micro* 33.1 (2013), pp. 54–67.

[133] Kishore Padmaraju and Keren Bergman. "Resolving the thermal challenges for silicon microring resonator devices." In: *Nanophotonics* 3.4-5 (2014), pp. 269–281.

[134] Kishore Padmaraju et al. "Integrated thermal stabilization of a microring modulator." In: *Optics express* 21.12 (2013), pp. 14342–14350.

[135] Kishore Padmaraju et al. "Intermodulation crosstalk characteristics of WDM silicon microring modulators." In: *IEEE Photonics Technology Letters* 26.14 (2014), pp. 1478–1481.

[136] Kishore Padmaraju et al. "Thermal stabilization of a microring modulator using feedback control." In: *Optics express* 20.27 (2012), pp. 27999–28008.

[137]   Kishore Padmaraju et al. "Wavelength locking and thermally stabilizing microring resonators using dithering signals." In: *Journal of Lightwave Technology* 32.3 (2014), pp. 505–512.

[138]   P Pintus, Ming-Chun Tien, and John E Bowers. "Design of magneto-optical ring isolator on SOI based on the finite-element method." In: *IEEE Photonics Technology Letters* 23.22 (2011), pp. 1670–1672.

[139]   Paolo Pintus et al. "Optimization of Integrated Silicon Doped Heaters for Optical Microring Resonators." In: *ECOC 2016; 42nd European Conference on Optical Communication; Proceedings of.* VDE. 2016, pp. 1–3.

[140]   Clifford R Pollock and Michal Lipson. *Integrated photonics.* Vol. 20. 25. Springer, 2003.

[141]   Robert Polster. "Architecture of silicon photonic links." PhD thesis. Paris 11, 2015.

[142]   Robert Polster et al. "A TIA for optical networks-on-chip in 65nm CMOS." In: *Optical Interconnects Conference (OI), 2015 IEEE.* IEEE. 2015, pp. 109–110.

[143]   Robert Polster et al. "Efficiency optimization of silicon photonic links in 65-nm CMOS and 28-nm FDSOI technology nodes." In: *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 24.12 (2016), pp. 3450–3459.

[144]   Robert Polster et al. "Optimization of TIA topologies in a 65nm CMOS process." In: *Optical Interconnects Conference, 2014 IEEE.* IEEE. 2014, pp. 117–118.

[145]   Andrew W Poon et al. "Cascaded microresonator-based matrix switch for silicon on-chip optical interconnection." In: *Proceedings of the IEEE* 97.7 (2009), pp. 1216–1238.

[146]   Miloš A Popovíc et al. "Multistage high-order microring-resonator add-drop filters." In: *Optics letters* 31.17 (2006), pp. 2571–2573.

[147]   John G Proakis et al. *Communication systems engineering.* Vol. 1. Prentice-hall Englewood Cliffs, 1994.

[148]   Jonathan Proesel, Alexander Rylyakov, and Clint Schow. "Optical receivers using DFE–IIR equalization." In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International.* IEEE. 2013, pp. 130–131.

[149]   Lei Qiao, Weijie Tang, and Tao Chu. "32× 32 silicon electro-optic switch with built-in monitors and balanced-status units." In: *Scientific Reports* 7 (2017), p. 42306.

[150] Ciyuan Qiu et al. "Wavelength tracking with thermally controlled silicon resonators." In: *Optics Express* 19.6 (2011), pp. 5143–5148.

[151] Rajiv Ramaswami, Kumar Sivarajan, and Galen Sasaki. *Optical networks: a practical perspective.* Morgan Kaufmann, 2009.

[152] Christian Jørgen Rasmussen et al. "Theoretical and experimental studies of the influence of the number of crosstalk signals on the penalty caused by incoherent optical crosstalk." In: *1999 Optical Fiber Communication Conference and the International Conference on Integrated Optics and Optical Fiber Communication.* 1999, pp. 258–260.

[153] Robert H Rediker and FJ Leonbercer. "Analysis of integrated-optics near 3 dB coupler and Mach-Zehnder interferometric modulator using four-port scattering matrix." In: *IEEE Transactions on Microwave Theory and Techniques* 30.10 (1982), pp. 1801–1804.

[154] Graham T Reed, William R Headley, and CE Jason Png. "Silicon photonics: The early years." In: *Integrated Optoelectronic Devices 2005.* International Society for Optics and Photonics. 2005, pp. 1–18.

[155] Graham T Reed et al. "Silicon optical modulators." In: *Nature photonics* 4.8 (2010), p. 518.

[156] Sven-Arne Reinemo, Tor Skeie, and Manoj K Wadekar. "Ethernet for high-performance data centers: On the new IEEE datacenter bridging standards." In: *IEEE micro* 30.4 (2010), pp. 42–51.

[157] Hagen Renner. "Bending losses of coated single-mode fibers: a simple approach." In: *Journal of Lightwave Technology* 10.5 (1992), pp. 544–551.

[158] S Rumley et al. "Silicon photonics for exascale systems." In: *Journal of Lightwave Technology* 33.3 (2015), pp. 547–562.

[159] Sébastien Rumley et al. "Optical interconnects for extreme scale computing systems." In: *Parallel Computing* 64 (2017), pp. 65–80.

[160] Sébastien Rumley et al. "Silicon photonics for exascale systems." In: *Journal of Lightwave Technology* 33.3 (2015), pp. 547–562.

[161] Bahaa EA Saleh, Malvin Carl Teich, and Bahaa E Saleh. *Fundamentals of photonics.* Vol. 22. Wiley New York, 1991.

[162] Harry A Schafft. "Thermal analysis of electromigration test structures." In: *IEEE Transactions on electron devices* 34.3 (1987), pp. 664–672.

[163]   M Schönleber, D Klotz, and E Ivers-Tiffée. "A method for improving the robustness of linear Kramers-Kronig validity tests." In: *Electrochimica Acta* 131 (2014), pp. 20–27.

[164]   Otto Schwelb. "Crosstalk and bandwidth of lossy microring add/drop multiplexers." In: *Optics communications* 265.1 (2006), pp. 175–179.

[165]   Tae Joon Seok et al. "Large-scale broadband digital silicon photonic switches with vertical adiabatic couplers." In: *Optica* 3.1 (2016), pp. 64–70.

[166]   Assaf Shacham, Keren Bergman, and Luca P Carloni. "On the design of a photonic network-on-chip." In: *Proceedings of the First International Symposium on Networks-on-Chip.* IEEE Computer Society. 2007, pp. 53–64.

[167]   Assaf Shacham, Keren Bergman, and Luca P Carloni. "Photonic networks-on-chip for future generations of chip multiprocessors." In: *Computers, IEEE Transactions on* 57.9 (2008), pp. 1246–1260.

[168]   Myung Jin Shin et al. "Parametric characterization of self-heating in depletion-type Si micro-ring modulators." In: *IEEE Journal of Selected Topics in Quantum Electronics* 22.6 (2016), pp. 116–122.

[169]   Myung Jin Shin et al. "Parametric characterization of self-heating in depletion-type Si micro-ring modulators." In: *IEEE Journal of Selected Topics in Quantum Electronics* 22.6 (2016), pp. 116–122.

[170]   Ivan Shubin et al. "Microring-based multi-chip WDM photonic module." In: *Optics express* 23.10 (2015), pp. 13172–13184.

[171]   Arjun Singh et al. "Jupiter rising: A decade of clos topologies and centralized control in google's datacenter network." In: *ACM SIGCOMM computer communication review.* Vol. 45. 4. ACM. 2015, pp. 183–197.

[172]   Benjamin A Small et al. "Multiple-wavelength integrated photonic networks based on microring resonator devices." In: *Journal of Optical Networking* 6.2 (2007), pp. 112–120.

[173]   Tom Smy et al. "Circuit-level transient simulation of configurable ring resonators using physical models." In: *JOSA B* 28.6 (2011), pp. 1534–1543.

[174]   RICHARDA Soref and BRIANR Bennett. "Electrooptical effects in silicon." In: *IEEE journal of quantum electronics* 23.1 (1987), pp. 123–129.

[175]   Richard Soref. "The past, present, and future of silicon photonics." In: *IEEE Journal of selected topics in quantum electronics* 12.6 (2006), pp. 1678–1687.

[176]   R Stabile, A Rohit, and KA Williams. "Monolithically integrated 8×8 space and wavelength selective cross-connect." In: *Journal of Lightwave Technology* 32.2 (2014), pp. 201–207.

[177]   Zhan Su et al. "An interior-ridge silicon microring switch with integrated thermal tuner." In: *Integrated Photonics Research, Silicon and Nanophotonics.* Optical Society of America. 2015, IM2B–5.

[178]   Harish Subbaraman et al. "Recent advances in silicon-based passive and active optical interconnects." In: *Optics express* 23.3 (2015), pp. 2487–2511.

[179]   Chen Sun et al. "A 45 nm CMOS-SOI monolithic photonics platform with bit-statistics-based resonant microring thermal tuning." In: *IEEE Journal of Solid-State Circuits* 51.4 (2016), pp. 893–907.

[180]   Chen Sun et al. "A 45nm SOI monolithic photonics chip-to-chip link with bit-statistics-based resonant microring thermal tuning." In: *VLSI Circuits (VLSI Circuits), 2015 Symposium on.* IEEE. 2015, pp. C122–C123.

[181]   Chen Sun et al. "Single-chip microprocessor that communicates directly using light." In: *Nature* 528.7583 (2015), p. 534.

[182]   Keijiro Suzuki et al. "Broadband silicon photonics 8×8 switch based on double-Mach–Zehnder element switches." In: *Optics Express* 25.7 (2017), pp. 7538–7546.

[183]   Shuichi Suzuki and Yasuo Kokubun. "Design rule of wavelength filter bandwidth and pulsewidth for ultimate spectral efficiency limited by crosstalk in DWDM systems." In: *Photonics Technology Letters, IEEE* 15.11 (2003), pp. 1645–1647.

[184]   Hiroshi Takahashi, Kazuhiro Oda, and Hiromu Toba. "Impact of crosstalk in an arrayed-waveguide multiplexer on N×N optical interconnection." In: *Lightwave Technology, Journal of* 14.6 (1996), pp. 1097–1105.

[185]   David Thomson et al. "Roadmap on silicon photonics." In: *Journal of Optics* 18.7 (2016), p. 073003.

[186]   Erman Timurdogan et al. "An ultralow power athermal silicon modulator." In: *Nature communications* 5 (2014), p. 4008.

[187]   Erman Timurdogan et al. "Automated wavelength recovery for microring resonators." In: *Lasers and Electro-Optics (CLEO), 2012 Conference on.* IEEE. 2012, pp. 1–2.

[188]   Erman Timurdogan et al. "Vertical junction silicon microdisk modulator with integrated thermal tuner." In: *Lasers and Electro-Optics (CLEO), 2013 Conference on.* IEEE. 2013, pp. 1–2.

[189]   Leonid Tsybeskov, David J Lockwood, and Masakazu Ichikawa. "Silicon Photonics: CMOS Going Optical [Scanning the Issue]." In: *Proceedings of the IEEE* 97.7 (2009), pp. 1161–1165.

[190]   Yutaka Urino et al. "Demonstration of 30-Tbps/cm 2 bandwidth density by silicon optical interposers fully integrated with optical components." In: *Optical Communication (ECOC 2013), 39th European Conference and Exhibition on.* IET. 2013, pp. 1–3.

[191]   Yutaka Urino et al. "High-density and wide-bandwidth optical interconnects with silicon optical interposers." In: *Photonics Research* 2.3 (2014), A1–A7.

[192]   Ahmad Usman et al. "Interposer Technologies for High-Performance Applications." In: *IEEE Transactions on Components, Packaging and Manufacturing Technology* 7.6 (2017), pp. 819–828.

[193]   Amin Vahdat et al. "Scale-out networking in the data center." In: *Ieee Micro* 30.4 (2010), pp. 29–41.

[194]   Yurii A Vlasov. "Silicon CMOS-integrated nano-photonics for computer and data communications beyond 100G." In: *Communications Magazine, IEEE* 50.2 (2012), s67–s72.

[195]   Mark T Wade et al. "75% efficient wide bandwidth grating couplers in a 45 nm microelectronics CMOS process." In: *Optical Interconnects Conference (OI), 2015 IEEE.* IEEE. 2015, pp. 46–47.

[196]   Xiaolong Wang et al. "SOI thermo-optic modulator with fast response." In: *Chinese Optics Letters* 1.9 (2003), pp. 527–528.

[197]   Xiaoxi Wang and Shayan Mookherjea. "Fast circuit modeling of heat transfer in photonic integrated circuits." In: *CLEO: Science and Innovations.* Optical Society of America. 2017, JW2A–141.

[198]   Michael R Watts et al. "Adiabatic resonant microrings (ARMs) with directly integrated thermal microphotonics." In: *Lasers and Electro-Optics, 2009 and 2009 Conference on Quantum electronics and Laser Science Conference. CLEO/QELS 2009. Conference on.* IEEE. 2009, pp. 1–2.

[199]   Michael R Watts et al. "Adiabatic thermo-optic Mach–Zehnder switch." In: *Optics letters* 38.5 (2013), pp. 733–735.

[200]   Z Weissman, A Hardy, and E Marom. "Mode-dependent radiation loss in Y junctions and directional couplers." In: *IEEE Journal of Quantum Electronics* 25.6 (1989), pp. 1200–1208.

[201]   Ke Wen et al. "Flexfly: Enabling a reconfigurable dragonfly through silicon photonics." In: *High Performance Computing, Networking, Storage and Analysis, SC16: International Conference for*. IEEE. 2016, pp. 166–177.

[202]   Wouter J Westerveld et al. "Extension of Marcatili's analytical approach for rectangular silicon optical waveguides." In: *Journal of Lightwave Technology* 30.14 (2012), pp. 2388–2401.

[203]   Callum M Wilkes et al. "60 dB high-extinction auto-configured Mach–Zehnder interferometer." In: *Optics letters* 41.22 (2016), pp. 5318–5321.

[204]   Peter J Winzer and R-J Essiambre. "Advanced optical modulation formats." In: *Proceedings of the IEEE* 94.5 (2006), pp. 952–985.

[205]   Peter J Winzer, Martin Pfennigbauer, and René-Jean Essiambre. "Coherent crosstalk in ultradense WDM systems." In: *Journal of lightwave technology* 23.4 (2005), p. 1734.

[206]   Rui Wu et al. "Compact Modeling and System Implications of Microring Modulators in Nanophotonic Interconnects." In: *International Workshop on System-Level Interconnect Prediction (SLIP)*. 2015.

[207]   Yiyuan Xie et al. "Crosstalk noise and bit error rate analysis for optical network-on-chip." In: *Proceedings of the 47th Design Automation Conference*. ACM. 2010, pp. 657–660.

[208]   Qianfan Xu, David Fattal, and Raymond G Beausoleil. "Silicon microring resonators with 1.5-μm radius." In: *Optics express* 16.6 (2008), pp. 4309–4315.

[209]   Qianfan Xu et al. "12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators." In: *Optics express* 15.2 (2007), pp. 430–436.

[210]   Nikola Zecevic, Michael Hofbauer, and Horst Zimmermann. "Integrated pulsewidth modulation control for a scalable optical switch matrix." In: *IEEE Photonics Journal* 7.6 (2015), pp. 1–7.

[211]   Lin Zhang et al. "Analysis and engineering of chromatic dispersion in silicon waveguide bends and ring resonators." In: *Optics Express* 19.9 (2011), pp. 8102–8107.

[212]   Yi Zhang et al. "A compact and low loss Y-junction for submicron silicon waveguide." In: *Optics express* 21.1 (2013), pp. 1310–1316.

[213]   Xuezhe Zheng et al. "A high-speed, tunable silicon photonic ring modulator integrated with ultra-efficient active wavelength control." In: *Optics express* 22.10 (2014), pp. 12628–12633.

[214]   Xuezhe Zheng et al. "Enhanced optical bistability from self-heating due to free carrier absorption in substrate removed silicon ring modulators." In: *Optics express* 20.10 (2012), pp. 11478–11486.

[215]   Xuezhe Zheng et al. "Ultra-efficient 10Gb/s hybrid integrated silicon photonic transmitter and receiver." In: *Optics Express* 19.6 (2011), pp. 5172–5186.

[216]   Xiang Zhou, Hong Liu, and Ryohei Urata. "Datacenter optics: requirements, technologies, and trends." In: *Chinese Optics Letters* 15.5 (2017), p. 120008.

[217]   Amir Kavyan Kavyan Ziabari et al. "Leveraging Silicon-Photonic NoC for Designing Scalable GPUs." In: *Proceedings of the 29th ACM on International Conference on Supercomputing*. ACM. 2015, pp. 273–282.

[218]   William A Zortman et al. "Bit-error-rate monitoring for active wavelength control of resonant modulators." In: *IEEE Micro* 33.1 (2013), pp. 42–52.

[219]   William Zortman et al. "Integrated CMOS comaptible low power 10Gbps silicon photonic heater modulator." In: *Optical Fiber Communication Conference*. Optical Society of America. 2012, OW4I–5.

[220]   Maurizio Zuffada. "The industrialization of the silicon photonics: technology road map and applications." In: *Solid-State Device Research Conference (ESSDERC), 2012 Proceedings of the European*. IEEE. 2012, pp. 7–13.

[221]   *http://gdspy.readthedocs.io*. URL: http://gdspy.readthedocs.io.

[222]   *http://onboardoptics.org/*. URL: http://onboardoptics.org/.

[223]   *https://en.wikipedia.org/wiki/Dennard_scaling*. URL: https://en.wikipedia.org/wiki/Dennard_scaling.

[224]   *https://www.lumerical.com/tcad-products/fdtd/*. URL: https://www.lumerical.com/tcad-products/fdtd/.

[225]   *https://www.synopsys.com/optical-solutions/rsoft.html*. URL: https://www.synopsys.com/optical-solutions/rsoft.html.

[226]   *http://www.aimphotonics.com/*. URL: http://www.aimphotonics.com/.

[227]   *http://www.analogphotonics.com/*. URL: http://www.analogphotonics.com/.

[228]   *http://www.mellanox.com/related-docs/solutions/ECOC_Market_Focus-Arlon_Martin.pdf*. URL: http://www.mellanox.com/related-docs/solutions/ECOC_Market_Focus-Arlon_Martin.pdf.

Relevant Author's Publications

1. Q. Cheng, R. Dai, **M. Bahadori**, P. Morrissey, R. Polster, S. Rumley, P. O'Brien, and K. Bergman, "Microring-Based Si/SiN Dual-Layer Switch Fabric", *Optical Interconnects* Conference, June 2018

2. **M. Bahadori**, and K. Bergman, "Low-Power Optical Interconnects based on Resonant Silicon Photonic Devices: Recent Advances and Challenges", *Great Lakes Symposium on VLSI (GLSVLSI)* Conference, May 2018

3. A. Gazman, Z. Zhu, **M. Bahadori**, and K. Bergman, "Wavelength Locking of Multicast Signals using Photo-Conductive Effect in Silicon Photonic Platform", *Optical Interconnects* Conference, June 2018

4. **M. Bahadori** , S. Rumley, Q. Cheng, and K. Bergman, "Impact of Backscattering on Microring-based Silicon Photonic Links", *Optical Interconnects* Conference, June 2018

5. Q. Cheng, S. Rumley, **M. Bahadori**, and K. Bergman, "Scalable photonic switching in high performance datacenters", *Optics Express* [invited paper], January 2018

6. **M. Bahadori**, M. Nikdast, S. Rumley, L. Y. Dai, N. Janosik, T. Van Vaerenbergh, A. Gazman, Q. Cheng, R. Polster, and K. Bergman, "Design Space Exploration of Microring Resonators in Silicon Photonic Interconnects: Impact of the Ring Curvature", *IEEE Journal of Lightwave Technology*, March 2018

7. **M. Bahadori**, A. Gazman, N. Janosik, S. Rumley, Z. Zhu, R. Polster, Q. Cheng, K. Bergman, "Thermal Rectification of Integrated Microheaters for Microring Resonators in Silicon Photonics Platform", *IEEE Journal of Lightwave Technology*, vol.

36, issue 3, pp. 773–788, December 2017

8. A. Gazman, **M. Bahadori**, Z. Zhu, and K. Bergman, "Programmable Optical Power Distribution in Silicon Photonic Platform", presented at *Optical Interconnects Conference*, June 2017

9. Q. Cheng, **M. Bahadori**, and K. Bergman, "Advanced Path Mapping for Silicon Photonic Switch Fabrics", presented at *CLEO (Science and Innovations) Conference*, May 2017

10. S. Rumley, **M. Bahadori**, R. Polster, S. D. Hammond, D. M. Calhoun, K. Wen, A. Rodrigues, and K. Bergman, "Optical interconnects for extreme scale computing systems", *Parallel Computing Journal*, February 2017

11. A. Gazman, C. Browning, **M. Bahadori**, Z. Zhu, P. Samadi, S. Rumley, V. Vujicic, L. P. Barry, and K. Bergman, "Software-defined control-plane for wavelength selective unicast and multicast of optical data in a silicon photonic platform", *Optics Express*, vol. 25, no. 1, pp. 232–242, 2017

12. **M. Bahadori**, S. Rumley, R. Polster, A. Gazman, M. Traverso, M. Webster, K. Patel, and K. Bergman, "Energy-Performance Optimized Design of Silicon Photonic Interconnection Networks for High-Performance Computing", presented at *Design, Automation, and Test in Europe (DATE) Conference*, March 2017

13. **M. Bahadori**, S. Rumley, "Analysis of Worst-case of Loss and Crosstalk for Planar Silicon Photonic Switch Architectures", presented at *Design, Automation, and Test in Europe (DATE) Conference*, March 2017

14. K. Wen, P. Samadi, S. Rumley, C. P. Chen, Y. Shen, **M. Bahadori**, K. Bergman, J. Wilke, "Flexfly: enabling a reconfigurable dragonfly through silicon photonics", presented at *International Conference for High Performance Computing, Networking, Storage and Analysis*, November 2016

15. **M. Bahadori**, S. Rumley, R. P. Polster, and K. Bergman, "Loss and Crosstalk of Scalable MZI-based Switch Topologies in Silicon Photonic Platform", presented at *IEEE*

Photonics Conference (IPC), October 2016

16. S. Rumley, **M. Bahadori**, D. Nikolova, and K. Bergman, "Physical Layer Compact Models for Ring Resonators based Dense WDM Optical Interconnects", presented at *European Conference on Optical Communication (ECOC)*, 2016

17. **M. Bahadori**, S. Rumley, H. Jayatilleka, K. Murray, N. A. F Jaeger, L. Chrostowski, S. Shekhar, and K. Bergman, "Crosstalk Penalty in Microring-Based Silicon Photonic Interconnect Systems", *IEEE Journal of Lightwave Technology*, vol. 34, no. 17, pp. 4043–4052, September 2016

18. **M. Bahadori**, R. Polster, S. Rumley, Y. Thonnart, J.-L. Gonzalez-Jimenez, and K. Bergman, "Energy-Bandwidth Design Exploration of Silicon Photonic Interconnects in 65nm CMOS", presented at *Optical Interconnects Conference*, March 2016

19. **M. Bahadori**, S. Rumley, and K. Bergman, "Impact of High-Speed Modulation on the Scalability of Silicon Photonic Interconnects", presented at *OPTICS*, 2016

20. S. Rumley, **M. Bahadori**, K. Wen, D. Nikolova, and K. Bergman, "PhoenixSim: Crosslayer Design and Modeling of Silicon Photonic Interconnects", *1st International Workshop on Advanced Interconnect Solutions and Technologies for Emerging Computing Systems (AISTECS)@ HiPEAC*, January 2016

21. **M. Bahadori**, S. Rumley, D. Nikolova, and K. Bergman, "Comprehensive Design Space Exploration of Silicon Photonic Interconnects", *IEEE Journal of Lightwave Technology*, vol. 34, no. 12, pp. 2975–2987, June 2016

22. **M. Bahadori**, A. Gazman, S. Rumley, Q. Li, and K. Bergman, "Nonlinear Temperature-Dependent Transfer Characteristics of Silicon Photonic Microring Resonators", presented at *Integrated Photonics Research Conference*, June 2015

23. **M. Bahadori**, D. Nikolova, S. Rumley, C. P. Chen, and K. Bergman, "Optimization of Microring-based Filters for Dense WDM Silicon Photonic Interconnects", presented at *Optical Interconnects Conference*, April 2015