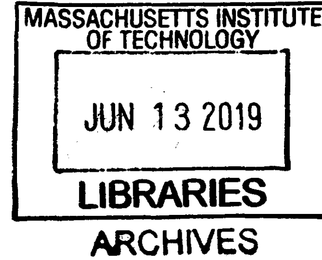
# Adjoint-Based Particle Defect Yield Modeling

by

Zhengxing Zhang

B.S. in Physics
Peking University, 2017




Submitted to the
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

© Massachusetts Institute of Technology 2019. All rights reserved.

Signature redacted

Author ............. ..................
Department of Electrical Engineering and Computer Science
May 23, 2019

Signature redacted

Certified by..........................................
Duane S. Boning
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Signature redacted

Accepted by..................................
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

# Adjoint-Based Particle Defect Yield Modeling

by

Zhengxing Zhang

Submitted to the Department of Electrical Engineering and Computer Science
on May 23, 2019, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

## Abstract

Silicon photonics, where photons instead of electrons are manipulated, shows promise for higher data rates, lower energy communication and information processing, biomedical sensing, and novel optically based applications such as wavefront engineering and beam steering of light. However, silicon photonics does not yet have mature process, device, and circuit variation models for the existing IC and photonic process steps; this lack presents a key challenge for design in this emerging industry.

This thesis addresses analysis of the process variation impact of particle defects. Such particle defects can arise in photolithography, deposition, etch, and other processes, and can perturb the intended function of photonic devices and circuits. The adjoint method previously used in optimization is modified and implemented to facilitate the simulation of the impact of defects in silicon photonic devices. More specifically, we demonstrate the methodology to build both component-level and circuit-level models based on the adjoint method. For the component-level models, we show how S-parameters of the device components are impacted by different types of particle defects. For the circuit-level models, we show the impact on circuit output spectrum and performance features based on component-level models, and perform critical area extraction for yield estimation. The model and result will be used to help generate layout design rules, predicting, and optimizing yield of complex silicon photonic devices and circuits for tomorrow's silicon photonics designers.

Thesis Supervisor: Duane S. Boning
Title: Professor of Electrical Engineering and Computer Science

# Acknowledgments

I would like to thank Prof. Duane S. Boning, my thesis supervisor. This thesis would never have be done without his support and guidance throughout the research in these two years, and I am truly happy to continue my research for a doctoral degree with him in the next couple of years.

I would like to thank Prof. Jacob White, Prof. Luca Daniel, Michael McIlrath, and all of the students in the Boning group and Photonic DFM group. It is always a good time for me to discuss recent research with them in the group meeting, and their questions and insights help me form deeper understanding for my research and of the entire field.

I would like to thank my family for always being supportive and respecting my decisions ever since I was a child. It is a really warm home that I have been missing and craving for when I am studying on the other side of the planet.

This research has been supported in part under AIM Photonics: this material is based on research sponsored by Air Force Research Laboratory under agreement number FA8650-15-2-5220. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory or the U.S. Government.

# Table of Contents


**1 Introduction** **17**

1.1 Motivation . . . . . 19

1.2 Thesis Structure . . . . . 19

**2 Background** **21**

2.1 Silicon Photonics Process . . . . . 21

2.2 S-Parameters . . . . . 22

2.2.1 Mode Expansion . . . . . 24

2.2.2 Equivalent Source Theory . . . . . 25

2.3 Adjoint Method . . . . . 26

2.3.1 Adjoint Method in Electromagnetics . . . . . 27

**3 Adjoint Method on Particle Defects** **29**

3.1 Adjoint Method for S-Parameters . . . . . 29

3.2 Polarizability . . . . . 31

3.3 Limitations . . . . . 33

**4 Implementation on Photonic Device Components** **35**

4.1 Straight Waveguide . . . . . 35

4.1.1 Silicon Pillar and Silicon Dioxide Hole . . . . . 37

4.1.2 Metal Sphere . . . . . 38

4.2 Y-Splitter . . . . . 44

4.3 Directional Coupler . . . . . 45

**5 Wavelength Dependence** **51**

5.1 Group Delay . . . . . 51

5.2 Phase Unwrapping . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
5.3 Memory Reduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55

**6 Circuit Level Variation Analysis and Yield Modeling 59**
6.1 Y-MZI: Implicit Example . . . . . . . . . . . . . . . . . . . . . . . . . . 59
6.1.1 Adjoint Method for INTERCONNECT . . . . . . . . . . . . . . 60
6.1.2 Curve Fitting . . . . . . . . . . . . . . . . . . . . . . . . . . . 63
6.2 Ring Resonator: Explicit Example . . . . . . . . . . . . . . . . . . . . . 64
6.3 Yield Modeling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 66

**7 Conclusions and Future Work 69**

# List of Figures


1-1 (a) Block diagram and (b) schematic layout of the dFT spectrometer design from [1]. . . . . . . . . . . . . . . . . . . . . . . . . . . 18

2-1 Demonstration of a silicon pillar defect in proximity to a straight silicon waveguide: (a) Top view; (b) Side view. . . . . . . . . . . . . . . . . . 22

2-2 Demonstration of a silicon dioxide pillar hole defect in proximity to (within) a straight silicon waveguide: (a) Top view; (b) Side view. . . 23

2-3 Demonstration of a metal sphere defect in proximity to a straight silicon waveguide: (a) Top view; (b) Side view. . . . . . . . . . . . . . . 23

4-1 The impact of a silicon dioxide hole pillar defect of $r = 20$ nm (blue) and $r = 40$ nm (red) on phase shift of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). . . . . . . . . . . . . . . . . . . . . . . . 38

4-2 The impact of a silicon pillar defect of $r = 20$ nm (blue) and $r = 40$ nm (red) on phase shift of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39

4-3 The impact of a silicon dioxide hole pillar defect of $r = 20$ nm (blue) and $r = 40$ nm (red) on transmission of a straight waveguide, as a function of distance $d$, evaluated from direct simulation. . . . . . . . . 39

4-4 The impact of a silicon pillar defect of $r = 20$ nm (blue) and $r = 40$ nm (red) on transmission of a straight waveguide, as a function of distance $d$, evaluated from direct simulation. . . . . . . . . . . . . . . . . . . . . 40

4-5 The impact of a metal sphere defect of $r = 20$ nm on transmission of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple). . . . . . . . . . . . . . . . . . . . . . . . . . 41

4-6 The impact of a metal sphere defect of $r = 40$ nm on transmission of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple). . . . . . . . . . . . . . . . . . . . . . . . . . 41

4-7 The impact of a metal sphere defect of $r = 20$ nm on phase shift of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple). . . . . . . . . . . . . . . . . . . . . . . . . . 42

4-8 The impact of a metal sphere defect of $r = 40$ nm on phase shift of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple). . . . . . . . . . . . . . . . . . . . . . . . . . 42

4-9 The impact of a metal sphere defect of $r = 20$ nm on back reflection of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple). . . . . . . . . . . . . . . . . . . . . . . . . . 43

4-10 The impact of a metal sphere defect of $r = 40$ nm on back reflection of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple). . . . . . . . . . . . . . . . . . . . . . . . . . 43

4-11 The implementation of the adjoint method on y-splitter in Lumerical FDTD simulation. The yellow box in the middle is the added field monitor. If the output of interests are only $T_{12}$, $\phi_{12}$ and $R_2$, the port element at port 3 is optional, while it is still required in complete S-parameters analysis. . . . . . . . . . . . . . . . . . . . . . . . . 45

4-12 The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the upper branch transmission $T_{12}$ of the y-splitter structure, as a function of the spatial location of the defect. Zoom-in around the cavity. . . . . . . . . . . . . . . . . . . . . . . . . . . 46

4-13 The impact of a silicon dioxide hole pillar defect of $r = 20$ nm at selected location on the upper branch transmission $T_{12}$ of the y-splitter structure, evaluated from adjoint method (blue) and direct simulation (red). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

4-14 The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the upper branch phase shift $\phi_{12}$ of the y-splitter structure, as a function of the spatial location of the defect. Zoom-in around the cavity. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47

4-15 The impact of a silicon dioxide hole pillar defect of $r = 20$ nm at selected location on the upper branch phase shift $\phi_{12}$ of the y-splitter structure, evaluated from adjoint method (blue) and direct simulation (red). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47

4-16 The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the upper branch back reflection $R_2$ of the y-splitter structure, as a function of the spatial location of the defect. Zoom-in around the cavity. . . . . . . . . . . . . . . . . . . . . . . . . . . 48

4-17 The impact of a silicon dioxide hole pillar defect of $r = 20$ nm at selected location on the upper branch back reflection $R_2$ of the y-splitter structure, evaluated from adjoint method (blue) and direct simulation (red). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 48

4-18 The implementation of the adjoint method on directional coupler in Lumerical FDTD simulation. Only port elements at port 2 and port 4 are required for analysis of self-coupling coefficient. The field monitor is set to only record around the gap to reduce memory cost. . . . . . 49

4-19 The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the self-coupling coefficient $t = |S_{24}|$ of the directional coupler, as a function of the spatial location of the defect. Zoom-in around the gap (the yellow box in Figure 4-18). . . . . . . . . . . . . 50

4-20 The mapping of the impact of a silicon pillar defect of $r = 20$ nm on the self-coupling coefficient $t = |S_{24}|$ of the directional coupler, as a function of the spatial location of the defect. Zoom-in around the gap (the yellow box in Figure 4-18). . . . . . . . . . . . . . . . . . . . . . 50

5-1 Demonstration of using a combination of small field monitors in the y-splitter example. . . . . . . . . . . . . . . . . . . . . . . . . . . . 56

6-1 The transmission of a nominal Y-MZI (blue) and its curve-fitting result (dashed). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

6-2 Comparison of the transmission of a nominal Y-MZI (blue) and the Y-MZI with defect, evaluated from adjoint method (red) and direct simulation (yellow). . . . . . . . . . . . . . . . . . . . . . . . . . . 63

6-3 The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the average transmission power $a_0$ of the Y-MZI, as a function of the spatial location of the defect. Zoom-in around the cavity of the left y-splitter. . . . . . . . . . . . . . . . . . . . . . . . . 65

6-4 The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the average ripple amplitude $b_0$ of the Y-MZI, as a function of the spatial location of the defect. Zoom-in around the cavity of the left y-splitter. . . . . . . . . . . . . . . . . . . . . . . . . 65

6-5 The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the finesse of ring resonator $\mathcal{F}$ at $\lambda = 1550$ nm, as a function of the spatial location of the defect. Zoom-in around the gap. 66

6-6 The critical area $A_0$ of the Y-MZI for particle radius $r = 20$ nm, as a function of threshold $\Gamma$, where we treat $\Delta b_0 > \Gamma$ as failure. . . . . . . 67

# List of Tables

3.1 Analytic examples of $\gamma$ as a function of $\varepsilon_p$, $\varepsilon$ for different shapes. . . . 32

# Chapter 1

# Introduction

The semiconductor industry continues to improve on past integrated circuit (IC) designs, even after decades of development and growth. While efforts are made to build denser and more complex IC chips, packages, and systems, the demand for faster interconnects and larger data transmission bandwidth prompts the new and emerging field of silicon-based photonics. By replacing the electrical wires with optical links which transmit signals through light, silicon photonics allows vastly higher data bandwidth and thus large-scale data transport, due to the fundamental differences in physics between electrons and photons [2]. Indeed, a microprocessor that uses optical waveguides to communicate information around the chip has been prototyped [3]. This device demonstrates the feasibility and advantages of using the same silicon fabrication process for both electronic based digital circuits and photonic based optical functions.

Silicon photonics has also shown great potential in the field of sensing. For example, an on-chip digital Fourier transform (dFT) spectroscopic system has been fabricated recently based on silicon photonics [1, 4], which achieves high spectral resolution in a relatively small design by making use of Mach-Zehnder interferometers and thermal-optical switches (Figure 1-1).

Such designs, however, may suffer from process variations aring during manufacturing. It is necessary to model such variations and analyze their impact on performance in order to increase the product yield and efficiency of fabricated silicon wafers, and indeed, such techniques for modeling manufacturing process variations

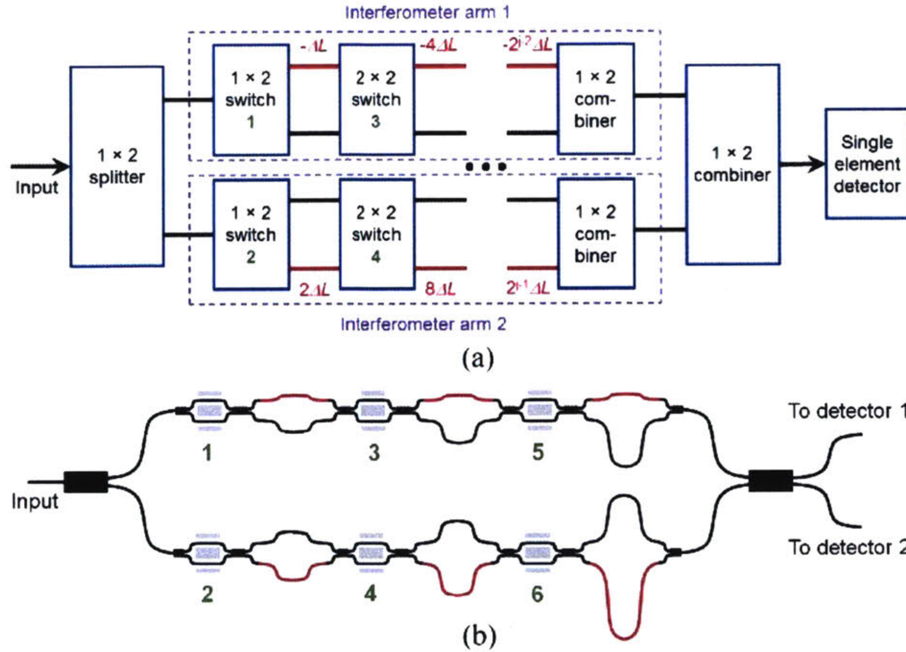



Figure 1-1: (a) Block diagram and (b) schematic layout of the dFT spectrometer design from [1].

have been developed for ICs [5], including creating physical models and predictive tools based on empirical data collected from wafer and process fabrication. However, silicon photonics does not yet have mature variation models for the existing IC and photonics process steps and device components; this lack presents a key challenge for design in this emerging industry.

Several research efforts have been focusing on building process and device variation models, both in simulation and experiment, for silicon photonics. For example, the impact of line-edge-roughness, which arises on the sidewall of the structure in the photolithography and etch process, has been studied on waveguides [6, 7], Bragg-gratings [8], and y-splitters [9]. There are also research efforts focusing on mathematical methods for developing these models, including the polynomial chaos expansion (PCE) method for uncertainty quantification [10, 11, 12], and the response surface model (RSM) for compact modeling [13, 14]. There is still great academic interest in exploring the impact of process variations, and thus develop design tools that account for those different types of variations.

## 1.1 Motivation

As one part of the goal of developing process variation models for silicon photonics, this thesis focuses on the impact of defects in silicon photonics that can arise in photolithography, deposition, etch, and other processes. The model and results will be used to help generate layout design rules and critical area extraction methods, and predicting and optimizing yield of complex silicon photonic devices and circuits for tomorrow's silicon photonics designers, just as IC designers do today.

Although doing direct finite-difference time-domain (FDTD) simulation by adding a defect particle somewhere near the device structure seems simple and easy to implement, it often requires an extremely fine mesh grid in FDTD as the particle becomes smaller, and requires a great number of simulations to sample and cover all the possible particle locations when the device becomes complicated. Instead, here we apply the adjoint method, which has been widely used recently in photonics inverse design [15, 16, 17], to both increase accuracy and accelerate simulation speed. In this thesis, the method will be applied on device components first, and later the simulated model will be used for photonic circuit simulation to produce a circuit-level model. Then critical area extraction will be performed for yield estimation.

## 1.2 Thesis Structure

Chapter 2 provides necessary background information that will be used throughout the thesis. The widely-used adjoint method is then adapted for the purpose of particle defect modeling in Chapter 3. Chapter 4 applies the method at the device component level, and shows implementation examples on straight waveguide, y-splitter, and directional coupler devices. Chapter 5 solves problems around wavelength dependence and frequency sampling. Chapter 6 applies the methodology at the photonic circuit level, explaining how to go from variation in component S-parameters to system output features. We also discuss yield modeling using results from circuit analysis. Chapter 7 provides conclusions and possible future research directions.

# Chapter 2

# Background

In this chapter, some necessary background information is presented. We start with an overview of the effects of defect study in silicon photonic circuits. Then we present the concept of S-parameters and the adjoint method in computational electromagnetics.

## 2.1 Silicon Photonics Process

There are different types of potential defects in silicon photonic circuits. Each defect type might have many causes, including foreign material, bubbles, pinholes in photo resists, and crystalline dislocations.

Yield management and analysis for ICs has been developed and implemented by taking defect sensitivities into account [18, 19]. The defect size distributions have been observed, with the conclusion that the probability distribution function is inversely proportional to the cube (or empirically, some other power $p$) of the defect size. The concept of critical area, where the center of a defect must fall to cause a circuit failure, has been introduced. Yield modeling for IC circuits has been further built upon the spatial distribution and the size distribution of defects [20].

Although there are few studies about the defect yield modeling for silicon photonics, considering the great similarities between the CMOS process and silicon photonics fabrication processes, we propose that the concept and methodology used in yield modeling for IC circuits can be transplanted to the silicon photonics area. In particular, we focus on two types of particle defects in this thesis: one is the pillar-shaped

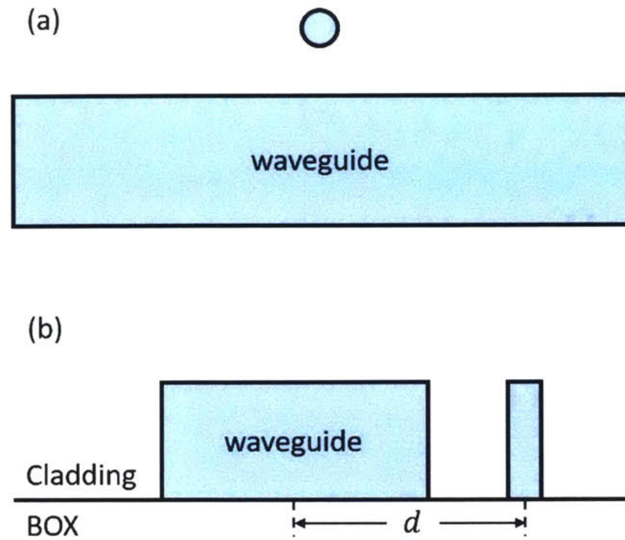



Figure 2-1: Demonstration of a silicon pillar defect in proximity to a straight silicon waveguide: (a) Top view; (b) Side view.

particle introduced during the photo-lithography process, the other is the ball-shaped particle introduced by foreign metal.

In the thesis, we always set the $z$-axis to be perpendicular to the circuit plan, and the location of the particle only depends on $x$ and $y$ while its $z$-position is fixed. For example, the pillar-shaped particle has the same height and same $z$-position as the silicon waveguide (Figure 2-1, Figure 2-2); the ball-shaped particle lies on the surface of the buried oxide (BOX) (Figure 2-3). The demonstrations and results in this thesis all are based on a silicon waveguide with width 500 nm and height 220 nm, though the methods can be applied to other geometries.

## 2.2 S-Parameters

Scattering parameters (S-parameters) are commonly used to describe the behaviour of a linear time-invariant network. Traditionally, this concept is used to describe the response of electrical devices, but they can be used to describe optical devices as well.

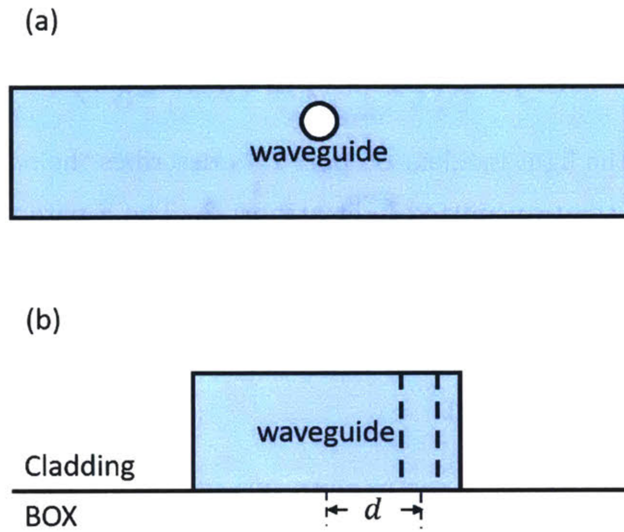



Figure 2-2: Demonstration of a silicon dioxide pillar hole defect in proximity to (within) a straight silicon waveguide: (a) Top view; (b) Side view.

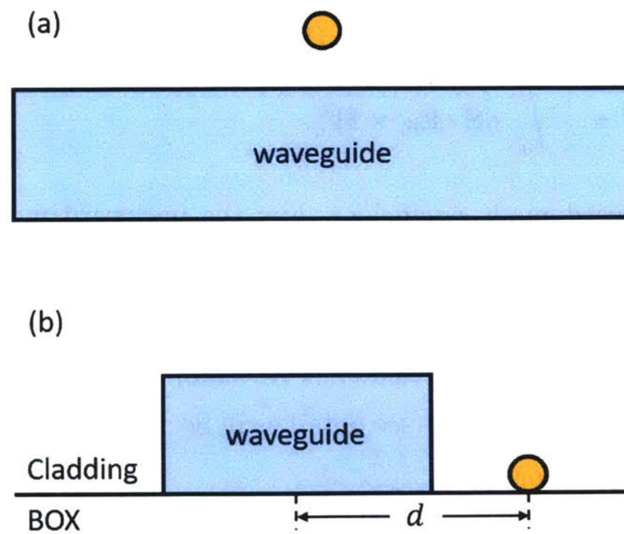



Figure 2-3: Demonstration of a metal sphere defect in proximity to a straight silicon waveguide: (a) Top view; (b) Side view.

For example, a two-optical-port device can be described as a matrix [2]:

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = S \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{2.1}$$

where $a_1$ describes the light incident on port 1, $b_1$ describes the reflected light at port 1, and $b_2$ describes the transmitted light at port 2. The S-parameters are generally complex, including an amplitude response and a phase response.

### 2.2.1 Mode Expansion

The S-parameters can be extracted from FDTD simulations, but a mode expansion is necessary to separate the waveguide modes of interest from the simulation field. Fortunately, in non-absorbing waveguides, the waveguide modes are power orthogonal, which gives the mode expansion coefficients for an arbitrary electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$, and modal fields $\mathbf{E}_{\mathrm{m}}$ and $\mathbf{H}_{\mathrm{m}}$, both on the surface $\partial\Omega$ [21]:

$$a = \frac{1}{4N}\left(\int_{\partial\Omega} \mathrm{d}\mathbf{S}\cdot\mathbf{E}\times\mathbf{H}_{\mathrm{m}}^{*} + \int_{\partial\Omega} \mathrm{d}\mathbf{S}\cdot\mathbf{E}_{\mathrm{m}}^{*}\times\mathbf{H}\right), \tag{2.2}$$

$$b = \frac{1}{4N}\left(\int_{\partial\Omega} \mathrm{d}\mathbf{S}\cdot\mathbf{E}\times\mathbf{H}_{\mathrm{m}}^{*} - \int_{\partial\Omega} \mathrm{d}\mathbf{S}\cdot\mathbf{E}_{\mathrm{m}}^{*}\times\mathbf{H}\right), \tag{2.3}$$

$$N = \frac{1}{2}\int_{\partial\Omega} \mathrm{d}\mathbf{S}\cdot\mathbf{E}_{\mathrm{m}}\times\mathbf{H}_{\mathrm{m}}^{*}, \tag{2.4}$$

where $a$ is the forward mode coefficient, $b$ is the backward mode coefficient, and $N$, corresponding to the power of the mode, is real for non-absorbing materials. S-parameters can be calculated from an S-parameter matrix sweep simulation combined with these mode expansion coefficients. For example, $S_{12}$ can be evaluated by a simulation where the source mode (see detail as in Section 2.2.2) is injected from port 1 and the monitor is at port 2:

$$S_{12} = \frac{a_{\text{monitor}}\sqrt{N_{\text{monitor}}}}{a_{\text{source}}\sqrt{N_{\text{source}}}}, \tag{2.5}$$

and appropriate $a$ or $b$ coefficient will be used for $a_{\text{monitor}}$ and $a_{\text{source}}$ depending on the direction of the mode.

It is straightforward to notice that $N$ only depends on the port and does not change through different simulations. Similarly, since the source mode is injected from every port exactly once in an S-parameter sweep (e.g., $S_{12}$, $S_{13}$, ... are evaluated from the same simulation in which the source mode is injected from port 1), the corresponding $a_{\text{source}}$ only depends on the port as well. Thus, we denote $N_i$, $s_i$ as $N$ coefficient, $a_{\text{monitor}}$ at port $i$ and $a_{ij}$, $\mathbf{E}_i$ as $a_{\text{monitor}}$ at port $j$; and electrical field from simulation where source mode is injected from port $i$, respectively in the rest of the thesis. Using this notation, (2.5) can be rewritten in general as

$$S_{ij} = \frac{a_{ij}\sqrt{N_j}}{s_i\sqrt{N_i}}. \tag{2.6}$$

### 2.2.2 Equivalent Source Theory

Another question around S-parameters simulations is how these source modes are generated. In Maxwell's equations, sources are represented in either dipole form (i.e., polarization $\mathbf{P}$ and magnetization $\mathbf{M}$) or current form (i.e., electric current density $\mathbf{J}$ and magnetic current density $\mathbf{K}$), which are connected through the continuity equation:

$$\mathbf{J} = \frac{\partial \mathbf{P}}{\partial t} = -\mathrm{i}\omega\mathbf{P}, \qquad \mathbf{K} = \mu_0\frac{\partial \mathbf{M}}{\partial t} = -\mathrm{i}\omega\mu_0\mathbf{M}, \tag{2.7}$$

where $\omega = 2\pi f$ is the angular frequency. So the question is: what is the dipole source ($\mathbf{P}$, $\mathbf{M}$) or current source ($\mathbf{J}$, $\mathbf{K}$) that generates fields ($\mathbf{E}_S = s\mathbf{E}_{\text{m}}$, $\mathbf{H}_S = s\mathbf{H}_{\text{m}}$) in some region $\Omega$?

Equivalent source theory implies that for a source-free field ($\mathbf{E}_S$, $\mathbf{H}_S$) in region $\Omega$ and no field outside $\Omega$, the corresponding current source is the surface current on $\partial\Omega$ [22]:

$$\begin{aligned} \mathbf{J}_S &= -\left(\mathbf{n} \times \mathbf{H}_S\right)\delta(\partial\Omega), \\ \mathbf{K}_S &= \left(\mathbf{n} \times \mathbf{E}_S\right)\delta(\partial\Omega), \end{aligned} \tag{2.8}$$

or equivalently in dipole form using (2.7),

$$\begin{aligned} \mathbf{P}_S &= -\frac{\mathrm{i}}{\omega}\left(\mathbf{n} \times \mathbf{H}_S\right)\delta(\partial\Omega), \\ \mathbf{M}_S &= \frac{\mathrm{i}}{\omega\mu_0}\left(\mathbf{n} \times \mathbf{E}_S\right)\delta(\partial\Omega), \end{aligned} \tag{2.9}$$

where $\mathbf{n}$ is the normal unit vector of the surface $\partial\Omega$, and $\delta(\partial\Omega)$ is the Dirac delta distribution that equals zero everywhere except for the surface $\partial\Omega$ and has the integral $\int \phi(\mathbf{x})\delta(\partial\Omega)\,\mathrm{d}^3\mathbf{x} = \int_{\partial\Omega} \phi(\mathbf{x})\,\mathrm{d}S$ for any continuous test function $\phi(\mathbf{x})$. This provides reference and understanding of how the source mode is generated in FDTD simulations.

## 2.3 Adjoint Method

The adjoint method (or adjoint state method) is a numerical method for efficiently computing the gradient of a function in a numerical optimization problem or for sensitivity analysis [23]. It has several different forms in different fields of application, but one can find a general derivation from the form of differential-algebraic equations in [24]. Specifically, the matrix form used for linear system analysis is easy to demonstrate [25]. Suppose we want to find the derivative of the output $y(p)$ of a linear system with variation parameters $p$:

$$A(p)x(p) = b, \qquad y(p) = c^T x(p). \tag{2.10}$$

Then it is easy to find that

$$\frac{\mathrm{d}y(p)}{\mathrm{d}p} = -c^T A(p)^{-1}\frac{\mathrm{d}A(p)}{\mathrm{d}p}x(p). \tag{2.11}$$

Thus the derivative at certain point $p_0$ can be computed using two system solves: $A(p_0)x(p_0) = b$ and $A(p_0)^T v = c$, and

$$\left.\frac{\mathrm{d}y}{\mathrm{d}p}\right|_{p_0} = -v^T \left.\frac{\mathrm{d}A}{\mathrm{d}p}\right|_{p_0} x(p_0), \tag{2.12}$$

which can be very efficient if the number of variation parameters is much greater than the number of outputs.

### 2.3.1 Adjoint Method in Electromagnetics

Maxwell's equations is an example of linear system analysis, where $x(p)$ is electromagnetic field ($\mathbf{E}(\mathbf{x})$, $\mathbf{H}(\mathbf{x})$), $A(p)$ is related to permittivity $\varepsilon(\mathbf{x})$, and $b$ is the source dipole ($\mathbf{P}(\mathbf{x})$, $\mathbf{M}(\mathbf{x})$) or source current ($\mathbf{J}(\mathbf{x})$, $\mathbf{K}(\mathbf{x})$). The output (real-valued) merit function

$$F(\mathbf{E},\mathbf{H}) = \int_{\chi} f(\mathbf{E}(\mathbf{x}),\mathbf{H}(\mathbf{x}))\,\mathrm{d}^3\mathbf{x} \tag{2.13}$$

is nonlinear in general, but can always be linearized around nominal point:

$$\delta F = 2\,\mathrm{Re}\int_{\chi}\left[\frac{\partial f}{\partial \mathbf{E}}(\mathbf{x})\cdot\delta\mathbf{E}(\mathbf{x}) + \frac{\partial f}{\partial \mathbf{H}}(\mathbf{x})\cdot\delta\mathbf{H}(\mathbf{x})\right]\mathrm{d}^3\mathbf{x}. \tag{2.14}$$

Thus $c$ is ($\partial f/\partial\mathbf{E}(\mathbf{x})$, $\partial f/\partial\mathbf{H}(\mathbf{x})$) in this form. We almost have a glimpse of the final formula for the adjoint method in electromagnetics by this analysis; indeed, it is possible to obtain the formula from (2.12). But here we take a different approach that gives us more physical insight into the problem, which was first proposed and implemented for optimization in [15].

Firstly, the effect of an infinitesimal perturbation of permittivity $\varepsilon(\mathbf{x}')$ in region $\psi$ is equivalent to introducing a dipole source

$$\mathbf{P}^{\mathrm{ind}}(\mathbf{x}') = \delta\varepsilon(\mathbf{x}')\mathbf{E}(\mathbf{x}') \tag{2.15}$$

in the same region, for a first-order approximation. Thus the perturbation ($\delta\mathbf{E}(\mathbf{x})$, $\delta\mathbf{H}(\mathbf{x})$) from $\mathbf{P}(\mathbf{x}')$ is

$$\begin{aligned}\delta\mathbf{E}(\mathbf{x}) &= \int_{\psi} G^{\mathrm{EP}}(\mathbf{x};\mathbf{x}')\mathbf{P}^{\mathrm{ind}}(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}' = \int_{\psi} G^{\mathrm{EP}}(\mathbf{x};\mathbf{x}')\delta\varepsilon(\mathbf{x}')\mathbf{E}(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}',\\ \delta\mathbf{H}(\mathbf{x}) &= \int_{\psi} G^{\mathrm{HP}}(\mathbf{x};\mathbf{x}')\mathbf{P}^{\mathrm{ind}}(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}' = \int_{\psi} G^{\mathrm{HP}}(\mathbf{x};\mathbf{x}')\delta\varepsilon(\mathbf{x}')\mathbf{E}(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}'.\end{aligned} \tag{2.16}$$

Notice that plugging (2.16) into (2.14) produces terms that look like $\mathbf{P}_1(\mathbf{x})\cdot G^{\mathrm{EP}}(\mathbf{x};\mathbf{x}')\mathbf{P}_2(\mathbf{x}')$ and $\mathbf{M}_1(\mathbf{x})\cdot G^{\mathrm{HP}}(\mathbf{x};\mathbf{x}')\mathbf{P}_2(\mathbf{x}')$, thus we can swap the order of the

inner product using reciprocity of Maxwell's equations:

$$
\begin{aligned}
\mathbf{P}_1(\mathbf{x}) \cdot G^{\mathrm{EP}}(\mathbf{x};\mathbf{x}')\mathbf{P}_2(\mathbf{x}') &= \mathbf{P}_2(\mathbf{x}') \cdot G^{\mathrm{EP}}(\mathbf{x}';\mathbf{x})\mathbf{P}_1(\mathbf{x}), \\
\mu_0\mathbf{M}_1(\mathbf{x}) \cdot G^{\mathrm{HP}}(\mathbf{x};\mathbf{x}')\mathbf{P}_2(\mathbf{x}') &= -\mathbf{P}_2(\mathbf{x}') \cdot G^{\mathrm{EM}}(\mathbf{x}';\mathbf{x})\mathbf{M}_1(\mathbf{x}).
\end{aligned}
\tag{2.17}
$$

Taking (2.16) and (2.17) into (2.14) and properly rearranging the order of integral, we get

$$
\delta F = 2\,\mathrm{Re} \int_{\psi} \delta\varepsilon(\mathbf{x}')\mathbf{E}(\mathbf{x}') \cdot \mathbf{E}^A(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}' \tag{2.18}
$$

where

$$
\mathbf{E}^A(\mathbf{x}') = \int_{\chi} \left[ G^{\mathrm{EP}}(\mathbf{x}';\mathbf{x}) \frac{\partial f}{\partial \mathbf{E}}(\mathbf{x}) - G^{\mathrm{EM}}(\mathbf{x}';\mathbf{x}) \frac{\partial f}{\partial \mathbf{H}}(\mathbf{x}) \right] \mathrm{d}^3\mathbf{x} \tag{2.19}
$$

is the eletric field from adjoint simulation for the same structure with dipole sources $(\mathbf{P}, \mathbf{M}) = (\partial f/\partial \mathbf{E}, -(1/\mu_0)\partial f/\partial \mathbf{H})$ in the region $\chi$. Since the adjoint field $\mathbf{E}^A$ only depends on the merit function but not the variation of permittivity, one single adjoint simulation is enough to compute any possible infinitesimal permittivity variation. Moreover, the adjoint simulation is based on the original struture without any perturbation, thus it does not require a fine mesh grid for high accuracy, as the direct method does.

# Chapter 3

# Adjoint Method on Particle Defects

In this chapter, we discuss how to apply the adjoint method on the particle defect problem. In particular, we focus on the impact of small particle defects on the S-parameters of photonic device components. To begin, we adapt the adjoint method for S-parameters in Section 3.1. Next, we discuss why particle defect is a special type of permittivity perturbation and needs special care in Section 3.2. Section 3.3 discusses and summarizes the effectiveness of the method that we develop.

## 3.1 Adjoint Method for S-Parameters

Although an implementation of the adjoint method wrapped around Lumerical FDTD has been proposed in [15], the figure of merit (FoM) was presumed to be real (for the purpose of optimization design), while the S-parameters we are modeling are complex. Despite the fact that it is feasible to model the amplitude response and the phase response separately, a more concise solution is to derive the adjoint method in the same path as [15] for S-parameters exclusively.

Since any S-parameter is proportional to either $a$ in (2.3) or $b$ in (2.4), here we demonstrate the adjoint method applied on the $a$ coefficient, and the same results can be easily derived for $b$ and thus $S$. In general, a complex merit function $F(\mathbf{E}, \mathbf{H})$ has much more complicated forms of derivatives form than that in (2.14). However,

for a holomorphic function $f(\mathbf{E}, \mathbf{H})$ where $\partial f/\partial \mathbf{E}^* = \partial f/\partial \mathbf{H}^* = 0$, the result is straightforward as

$$\delta F = \int_\chi \left[ \frac{\partial f}{\partial \mathbf{E}}(\mathbf{x}) \cdot \delta \mathbf{E}(\mathbf{x}) + \frac{\partial f}{\partial \mathbf{H}}(\mathbf{x}) \cdot \delta \mathbf{H}(\mathbf{x}) \right] \mathrm{d}^3\mathbf{x}. \tag{3.1}$$

The differences between (2.14) and (3.1) are just the factor of 2 and the real part operator. Thus the result from the adjoint method is similar as well,

$$\delta F = \int_\psi \delta\varepsilon(\mathbf{x}')\mathbf{E}(\mathbf{x}') \cdot \mathbf{E}^A(\mathbf{x}')\, \mathrm{d}^3\mathbf{x}', \tag{3.2}$$

with the same adjoint field as in (2.19).

Fortunately this is the case for the $a$ coefficient, thus it is easy to obtain the adjoint dipole sources:

$$\begin{aligned} \mathbf{P}(\mathbf{x}) &= -\frac{1}{4N} (\mathbf{n} \times \mathbf{H}_\mathrm{m}^*)\, \delta(\partial\Omega), \\ \mathbf{M}(\mathbf{x}) &= -\frac{1}{4N\mu_0} (\mathbf{n} \times \mathbf{E}_\mathrm{m}^*)\, \delta(\partial\Omega). \end{aligned} \tag{3.3}$$

Compared with (2.9), this is actually the equivalent current sources of the mode source

$$\mathbf{H}_S = -\frac{\mathrm{i}\omega}{4N}\mathbf{H}_\mathrm{m}^*, \qquad \mathbf{E}_S = \frac{\mathrm{i}\omega}{4N}\mathbf{E}_\mathrm{m}^*, \tag{3.4}$$

which is exactly the backward-propagating mode with amplitude $\mathrm{i}\omega/4N$.

It is unnecessary to specify the amplitude of the mode source in FDTD simulation; instead, the mode source is generated with an arbitrary amplitude $s$, which later can be obtained from the mode expansion monitor. In such scenario, we can write the variation in $a$ as

$$\delta a = \frac{\mathrm{i}\omega}{4sN} \int_\psi \delta\varepsilon(\mathbf{x}')\mathbf{E}(\mathbf{x}') \cdot \mathbf{E}^A(\mathbf{x}')\, \mathrm{d}^3\mathbf{x}', \tag{3.5}$$

where the adjoint field $\mathbf{E}^A$ comes from a backward mode with amplitude $s$.

Now we try to expand the case for S-parameters. If (3.5) is used in the calculation of $a_{ij}$ for $S_{ij}$, then the forward field $\mathbf{E}$ is $\mathbf{E}_i$ from simulation for $S_{ij}$, and the adjoint field $\mathbf{E}^A$ is $\mathbf{E}_j$ from simulation for $S_{ji}$. This suggests that the adjoint simulation of one S-parameter simulation is just the reciprocal S-parameter simulation. Therefore,

an S-parameter matrix sweep is able to cover all of the S-parameter simulations with their adjoint simulation, and no additional simulation is needed for adjoint analysis. From (2.6), we have

$$\delta S_{ij} = \frac{\mathrm{i}\omega}{4 s_i s_j \sqrt{N_i N_j}} \int_\psi \delta\varepsilon(\mathbf{x}')\mathbf{E}_i(\mathbf{x}') \cdot \mathbf{E}_j(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}', \tag{3.6}$$

or written in the form of transmission $T = |S|^2$ and phase $\phi = \arg(S)$,

$$\delta T_{ij} = -\frac{\omega}{2|s_i|^2 N_i} \mathrm{Im}\, \frac{a_{ij}^*}{s_j} \int_\psi \delta\varepsilon(\mathbf{x}')\mathbf{E}_i(\mathbf{x}') \cdot \mathbf{E}_j(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}', \tag{3.7}$$

$$\delta \phi_{ij} = \frac{\omega}{4N_j} \mathrm{Re}\, \frac{1}{a_{ij} s_j} \int_\psi \delta\varepsilon(\mathbf{x}')\mathbf{E}_i(\mathbf{x}') \cdot \mathbf{E}_j(\mathbf{x}')\,\mathrm{d}^3\mathbf{x}'. \tag{3.8}$$

## 3.2 Polarizability

While it may not be intuitive, the dipole approximation in (2.15) does not hold for finite permittivity variation. The reason is the discontinuous interface introduced by particle.

We consider a simple example first: a dielectric sphere (permittivity $\varepsilon_1$) in vacuum $\varepsilon_0$ with uniform electric field $\mathbf{E}$. While the relationship (2.15) still holds for overall total field:

$$\mathbf{P}^{\mathrm{ind}} = (\varepsilon_1 - \varepsilon_0)\mathbf{E}_{\mathrm{tot}}, \tag{3.9}$$

the polarization charge on the surface of sphere will induce an extra electric field inside the sphere and thus change the total field:

$$\mathbf{E}_{\mathrm{tot}} = \mathbf{E} - \frac{\mathbf{P}^{\mathrm{ind}}}{3\epsilon_0}. \tag{3.10}$$

Combining (3.9) and (3.10), we have

$$\mathbf{P}^{\mathrm{ind}} = \frac{3\varepsilon_0}{\varepsilon_1 + 2\varepsilon_0}(\varepsilon_1 - \varepsilon_0)\mathbf{E} = \gamma\Delta\varepsilon\mathbf{E}. \tag{3.11}$$

When $\varepsilon_1 \to \varepsilon_0$, $\gamma = 3\varepsilon_0/(\varepsilon_1 + 2\varepsilon_0)$ reaches 1, and thus (3.11) goes back to (2.15). In general, $\mathbf{P}^{\mathrm{ind}}$ can be not parallel to $\mathbf{E}$, and can even be non-uniform. But for the case where the particle is small, what matters most is not the polarization $\mathbf{P}^{\mathrm{ind}}$ but

the dipole moment

$$\mathbf{p}^{\text{ind}} = \int \mathbf{P}^{\text{ind}}(\mathbf{x}')\, \mathrm{d}^3\mathbf{x}'. \tag{3.12}$$

The relationship between the dipole moment of a particle and the electric field is known as: [26]

$$\mathbf{p}^{\text{ind}} = \alpha \mathbf{E}, \tag{3.13}$$

where the polarizability $\alpha$, inspired by (3.11), can be expressed as

$$\alpha = \gamma V \Delta\varepsilon, \tag{3.14}$$

where $V$ is the volume of the particle, the factor $\gamma$ can be scalar or tensor, and $\gamma$ depends on the shape of the particle, as well as the permittivity of the particle $\varepsilon_p$ and its environment $\varepsilon$, but it always symmetric and satisfies

$$\lim_{\Delta\varepsilon \to 0} \gamma = 1. \tag{3.15}$$

We list a few analytic examples of $\gamma$ as a function of $\varepsilon_p$, $\varepsilon$ for different shapes in Table 3.1. In general, we always have $\gamma < 1$ when $\varepsilon_p > \varepsilon$, and $\gamma > 1$ when $\varepsilon_p < \varepsilon$. This means that a silicon particle in the oxide cladding is less significant than an oxide hole of same volume in the silicon waveguide, even under the same amplitude of electric field – we will see validated examples in the next chapter.

Table 3.1: Analytic examples of $\gamma$ as a function of $\varepsilon_p$, $\varepsilon$ for different shapes.

| Shape | Sphere | Cylinder (tall approximation) | Cylinder (flat approximation) |
| --- | --- | --- | --- |
| $\gamma$ | $\dfrac{3\varepsilon}{\varepsilon_p + 2\varepsilon}$ | $\begin{pmatrix} 2\varepsilon/(\varepsilon_p+\varepsilon) & & \\ & 2\varepsilon/(\varepsilon_p+\varepsilon) & \\ & & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & & \\ & 1 & \\ & & \varepsilon/\varepsilon_p \end{pmatrix}$ |

Under the small particle approximation, (3.2) can be written for a particle introduced at $\mathbf{x}_p$:

$$\Delta F = (V\Delta\varepsilon)\gamma \mathbf{E}(\mathbf{x}_p) \cdot \mathbf{E}^A(\mathbf{x}_p), \tag{3.16}$$

and when the FoM is S-parameters,

$$\Delta S_{ij} = \frac{\mathrm{i}\omega V \Delta\varepsilon}{4 s_i s_j \sqrt{N_i N_j}} \gamma \mathbf{E}_i(\mathbf{x}_p) \cdot \mathbf{E}_j(\mathbf{x}_p). \tag{3.17}$$

## 3.3 Limitations

Before we jump into the implementation of the adjoint method on different photonic devices, we want to make a summary of all the approximations in the method we have made so far, and their possible implications in terms of potential errors in real-life implementations. This will serve as explanation for part of the validation result in next chapter.

The dipole approximation used in (2.15) and (3.13) is a commonly used first-order approximation in electromagnetics, which ignores higher order effects, e.g., magnetic dipole, quadrupole, etc. Such higher order effects can be significant for large perturbation – this is less important in the particle defect problem since large-sized particle rarely occurs. But they can also be significant when the first order effect is zero, which means that if $\Delta F$ calculated from the adjoint method is 0, the real variation $\Delta F$ is likely to be non-zero (although it is small in most cases) because of the higher order effects.

Specifically for particle defects, the approximation used in (3.13) and (3.16) also assumes that the electric field around the particle is uniform, which suggests that the size of particle should be much smaller than the length scale of change of the electric field. However, some particle defects do not satisfy such criteria, e.g., the silicon pillar defect from photo-lithography has same height as the waveguide, which is comparable to the length scale of change of the electric field in that direction. It is also hard to determine an exact location ($\mathbf{x}_p$ in (3.16)) for the particle in such case, so we will try to average the field over the region within the particle $\psi_p$:

$$\gamma \mathbf{E}(\mathbf{x}_p) \cdot \mathbf{E}^A(\mathbf{x}_p) \approx \frac{1}{V} \int_{\psi_p} \gamma \mathbf{E}(\mathbf{x}) \cdot \mathbf{E}^A(\mathbf{x}) \, \mathrm{d}^3\mathbf{x}. \tag{3.18}$$

The averaging approach potentially introduces higher order error, but it could be small when the field is varying slowly. Such assumption can be even more problematic

for metal particles, since the field tends to gather near the surface of conductive material at high frequency, which is known as the skin effect. Thus the size of particle needs to be even smaller for proper dipole approximation. Even when these criteria are satisfied, the factor $\gamma$ usually has some deviation from its theoretical value.

When the particle defect intersects with the material interface (e.g., the waveguide sidewall in most cases), either the actual shape of the particle changes, which makes it difficult to evaluate the shape-dependent factor $\gamma$, or the electric field is discontinuous, which breaks the uniformity assumption. We simply neglect these transition areas in our application, as they are extremely small for small particles, and we can approximate the behavior through interpolation if needed.

In general, a nonlinear merit function can introduce extra error in the linearization process in (2.14) or (3.1); but S-parameters happen to be linear functions, adn are thus free from these potential error. In the next chapter, we consider examples where the change in S-parameters is large compared to its original value, but the adjoint method still gives accurate enough results. However, since the transmission and phase are non-linear functions of S-parameters, the adjoint method in these forms introduces extra error and can be inaccurate when the nominal $S_{ij}$ is close to zero.

# Chapter 4

# Implementation on Photonic Device Components

In this chapter, we will show the implementation of the adjoint method developed in Chapter 3 on several different photonic device components: a straight waveguide in Section 4.1, a y-splitter in Section 4.2, and a directional coupler in Section 4.3. We will explain the simulation setup in each section, and validate the adjoint method by comparing to results from direct simulation. For complicated components like the y-splitter and directional coupler, a point-to-point validation is infeasible, so we will only consider a number of signature locations for direct simulation. The operation wavelength is 1550 nm for these examples; some figures show results for a spectrum ranging from 1500 to 1600 nm; we will discuss these additional wavelengths in Chapter 5.

## 4.1 Straight Waveguide

We set the straight waveguide to propagate light along the $x$ axis, and put port 1 at $x = 0$ and port 2 at $x = L$. From the propagating theory, $S_{12} = \exp(\mathrm{i}n_e kL)$ and

$$\mathbf{E}_1(x, y, z) = s_1 \mathbf{E}_\mathrm{m}(y, z) e^{\mathrm{i}n_e kx}, \qquad \mathbf{E}_2(x, y, z) = s_2 \mathbf{E}^*_\mathrm{m}(y, z) e^{\mathrm{i}n_e k(L-x)} \tag{4.1}$$

where $n_e$ is the effective index of the waveguide mode, and $k = 2\pi/\lambda$ is the angular wavenumber in vacuum. Using (3.17), we have

$$\Delta S_{11} = \frac{\mathrm{i}\omega V \Delta\varepsilon}{4N} \gamma \mathbf{E}_{\mathrm{m}}(y_p, z_p) \cdot \mathbf{E}_{\mathrm{m}}(y_p, z_p) e^{2\mathrm{i}n_e k x_p}, \tag{4.2}$$

$$\Delta S_{12} = \frac{\mathrm{i}\omega V \Delta\varepsilon}{4N} \gamma \mathbf{E}_{\mathrm{m}}(y_p, z_p) \cdot \mathbf{E}^*_{\mathrm{m}}(y_p, z_p) e^{\mathrm{i}n_e k L}. \tag{4.3}$$

We define transmission $T = T_{12}$, reflection $R = T_{11}$, and phase shift $\phi = \phi_{12}$; then

$$\Delta T = -\frac{\omega V}{2N} \operatorname{Im} \Delta\varepsilon\gamma \mathbf{E}_{\mathrm{m}}(y_p, z_p) \cdot \mathbf{E}^*_{\mathrm{m}}(y_p, z_p), \tag{4.4}$$

$$\Delta \phi = \frac{\omega V}{4N} \operatorname{Re} \Delta\varepsilon\gamma \mathbf{E}_{\mathrm{m}}(y_p, z_p) \cdot \mathbf{E}^*_{\mathrm{m}}(y_p, z_p), \tag{4.5}$$

$$\Delta R = \frac{\omega^2 V^2}{16N^2} \left|\Delta\varepsilon\gamma \mathbf{E}_{\mathrm{m}}(y_p, z_p) \cdot \mathbf{E}_{\mathrm{m}}(y_p, z_p)\right|^2. \tag{4.6}$$

Before making any assumption about the type of particle defect, there are several things we can observe from these expressions. Firstly, $x_p$ does not appear in any of these variations. Since $z_p$ is a fixed number for a certain type of particle with certain size, it suggests that there is only one degree of freedom in this problem, which is the distance $d$ from the center of the particle to the center line of the waveguide (as shown in Figure 2-1, Figure 2-2, and Figure 2-3). This is consistent with our intuition. Secondly, $\Delta R \propto V^2|\Delta\varepsilon|^2$ is a second order effect that is always very small compared to $\Delta T$ and $\Delta\phi$. Moreover, all we need from the simulation is the mode profile $\mathbf{E}_{\mathrm{m}}$ and its power $N$, which can be obtained from an eigenmode solver rather than FDTD simulator. Finally, for the fundamental TE mode of a silicon waveguide specifically, $E_{\mathrm{m}z}$ is small and negligible, and $E_{\mathrm{m}x}E^*_{\mathrm{m}y}$ is always a pure imaginary number[1], thus

$$\gamma \mathbf{E}_{\mathrm{m}}(y_p, z_p) \cdot \mathbf{E}^*_{\mathrm{m}}(y_p, z_p) = \gamma_{xx}|E_{\mathrm{m}x}(y_p, z_p)|^2 + \gamma_{yy}|E_{\mathrm{m}y}(y_p, z_p)|^2. \tag{4.7}$$

Therefore, only two diagonal elements of the tensor $\gamma$ contribute to the variation in transmission and phase shift.[2]

[1]These statements are not necessarily true for complicated devices. In the Lumerical MODE and FDTD solver, $E_{\mathrm{m}y}$ and $E_{\mathrm{m}z}$ is always set to be real-valued, and thus $E_{\mathrm{m}y}$ is pure imaginary.

[2]This also suggests a possible approach to evaluate $\gamma$ from simulation by rotating the particle.

### 4.1.1 Silicon Pillar and Silicon Dioxide Hole

For a non-absorbing particle, $\Delta R$ can be too small to detect in simulation; thus, here we mainly focus on the transmission and phase shift induced by the defect. Since $\Delta\varepsilon$ and $\gamma$ are real for non-absorbing materials, and also using (4.7),

$$\Delta T = 0, \tag{4.8}$$

$$\Delta\phi = \frac{\omega V \Delta\varepsilon}{4N}\left(\gamma_{xx}|E_{\mathrm{m}x}(y_p, z_p)|^2 + \gamma_{yy}|E_{\mathrm{m}y}(y_p, z_p)|^2\right). \tag{4.9}$$

For a cylinder-shaped pillar, $\gamma_{xx} = \gamma_{yy} = \gamma$,[3] thus

$$\Delta\phi = \frac{\omega V \gamma \Delta\varepsilon}{4N}|\mathbf{E}_{\mathrm{m}}(y_p, z_p)|^2, \tag{4.10}$$

where $\gamma = 2\varepsilon_{\mathrm{clad}}/(\varepsilon_{\mathrm{clad}} + \varepsilon_{\mathrm{si}}) = 0.3$ for a silicon cylinder pillar in oxide cladding, and $\gamma = 2\varepsilon_{\mathrm{si}}/(\varepsilon_{\mathrm{clad}} + \varepsilon_{\mathrm{si}}) = 1.7$ for an oxide cylinder hole in the silicon waveguide. (4.8) does not mean there is no transmission loss in a straight waveguide, but its first order effect is zero. On the other hand, from conservation of energy, we must have $\Delta T + \Delta R < 0$, which gives $\Delta T < 0$. Therefore, it would be more appropriate if we write it as

$$\Delta T = O(|V\Delta\varepsilon|^2). \tag{4.11}$$

As mentioned in Section 3.3, the height of the pillar $h = 220$ nm is too large to determine the value of $z_p$, thus we average over the $z$-axis:

$$\Delta\phi = \frac{\omega V \gamma \Delta\varepsilon}{4Nh}\int_0^h |\mathbf{E}_{\mathrm{m}}(y_p, z)|^2 \, \mathrm{d}z. \tag{4.12}$$

We test the adjoint method on both a small particle ($r = 20$ nm) and a large particle ($r = 40$ nm). For the oxide hole, the phase shift predicted by adjoint method shows good consistency with direct simulation (Figure 4-1), and even for the large particle, the error is still in a reasonable range. On the other hand, the same method does not work so well for the silicon pillar (Figure 4-2) since the electric field is varying too much near the interface, which happens to be the position with maximum impact.

[3] A cylinder is not the only shape that satisfies this requirement; a square pillar, an octagon pillar, etc. also have $\gamma_{xx} = \gamma_{yy} = \gamma$, although the value of $\gamma$ is different.

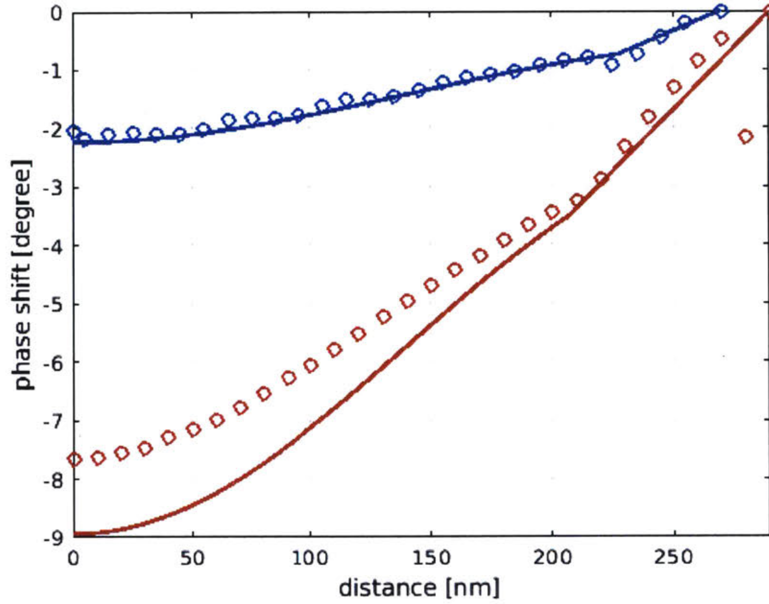



Figure 4-1: The impact of a silicon dioxide hole pillar defect of $r = 20$ nm (blue) and $r = 40$ nm (red) on phase shift of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line).

However, since the impact of the silicon pillar defect is much smaller than the impact of the dioxide hole of the same size, the absolute error from the adjoint method for the silicon pillar defect is still acceptable.

We also show the transmission impact from direct simulation in Figure 4-3 and Figure 4-4. These are the higher order effects that are ignored in the adjoint method, and here we see they are indeed negligible.

### 4.1.2 Metal Sphere

For a sphere-shaped particle, $\gamma$ is a scalar. Thus

$$\Delta T = -\frac{\omega V}{2N}|\mathbf{E}_{\mathrm{m}}(y_p, z_p)|^2 \ \mathrm{Im}\ \gamma\Delta\varepsilon, \tag{4.13}$$

$$\Delta\phi = \frac{\omega V}{4N}|\mathbf{E}_{\mathrm{m}}(y_p, z_p)|^2 \ \mathrm{Re}\ \gamma\Delta\varepsilon, \tag{4.14}$$

$$\Delta R = \frac{\omega^2 V^2}{16N^2}\left|\mathbf{E}_{\mathrm{m}}(y_p, z_p)\cdot\mathbf{E}_{\mathrm{m}}(y_p, z_p)\right|^2 |\gamma\Delta\varepsilon|^2. \tag{4.15}$$

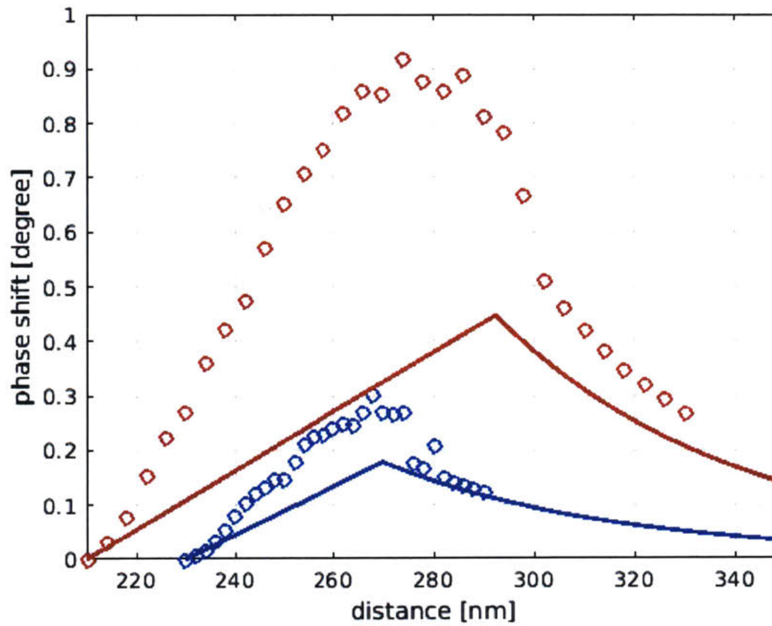



Figure 4-2: The impact of a silicon pillar defect of $r = 20$ nm (blue) and $r = 40$ nm (red) on phase shift of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line).

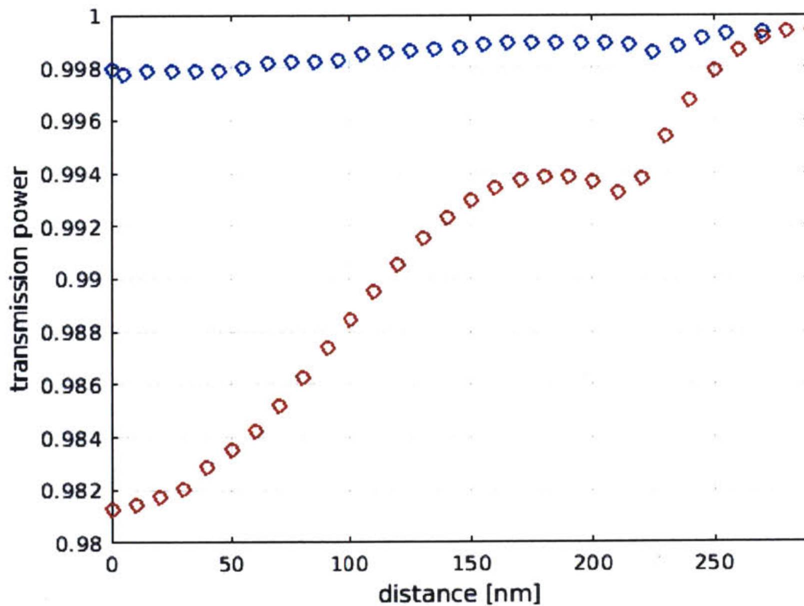



Figure 4-3: The impact of a silicon dioxide hole pillar defect of $r = 20$ nm (blue) and $r = 40$ nm (red) on transmission of a straight waveguide, as a function of distance $d$, evaluated from direct simulation.

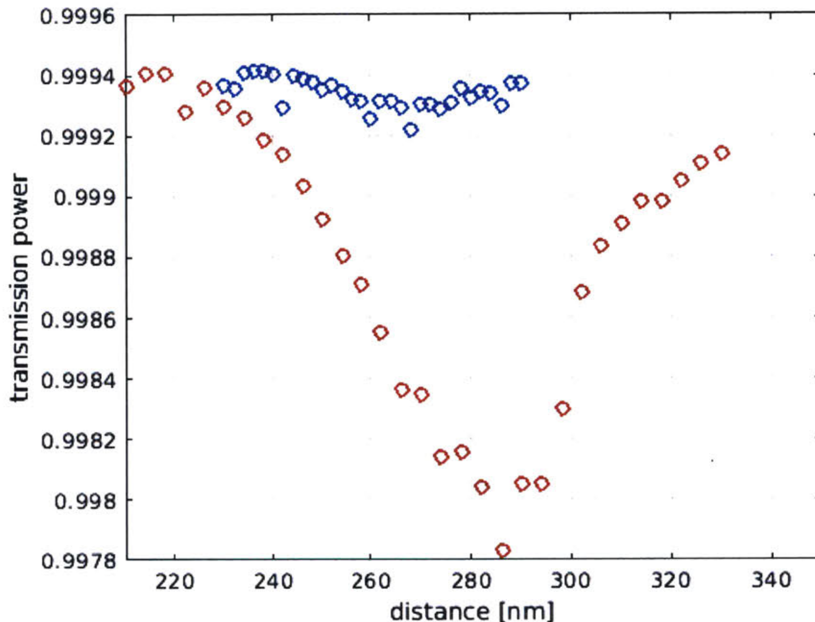



Figure 4-4: The impact of a silicon pillar defect of $r = 20$ nm (blue) and $r = 40$ nm (red) on transmission of a straight waveguide, as a function of distance $d$, evaluated from direct simulation.

where $\gamma = 3\varepsilon/(\varepsilon_p + 2\varepsilon)$. We test four different metal materials that are commonly used in integrated circuit fabrication: Aluminium, Copper, Tungsten, and Titanium. As with the pillar defect, we test both small particle ($r = 20$ nm) and large particle ($r = 40$ nm) cases.

We begin our analysis with the small particle. The adjoint method captures the impact on phase shift (Figure 4-7) relatively well, but for highly conductive materials like Aluminium and Copper, it seems to underestimate the transmission impact (Figure 4-5). This can be understood as error in the polarization factor $\gamma$; thus we can try to make a correction to match the result from direct simulation. For the impact on back reflection (Figure 4-9), the adjoint method and direct simulation show similar trends but different values; however, we cannot guarantee the accuracy of direct simulation since the value is too small.

However, the adjoint method does not produce reasonable results for the large particle case. There is a trend in direct simulation that obviously cannot be matched by simple correction on $\gamma$, e.g., the impact of Tungsten defect on phase shift (Figure 4-8). And comparing with the results for the small particle, we see many values that

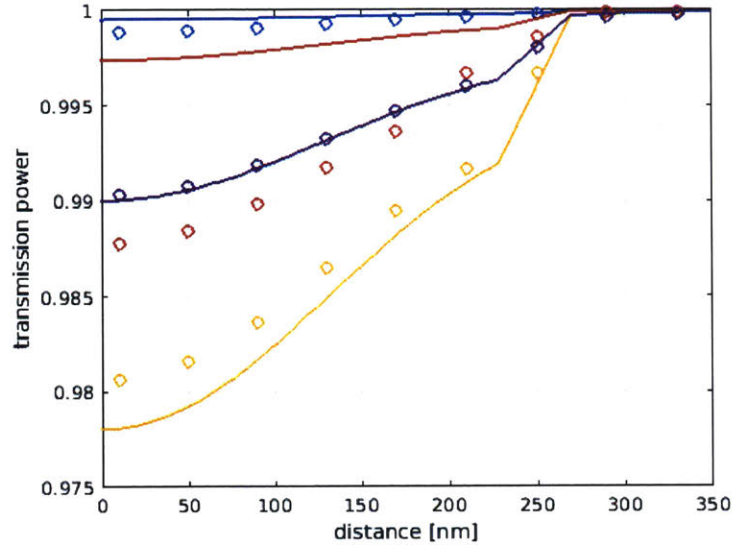



Figure 4-5: The impact of a metal sphere defect of $r$ = 20 nm on transmission of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple).

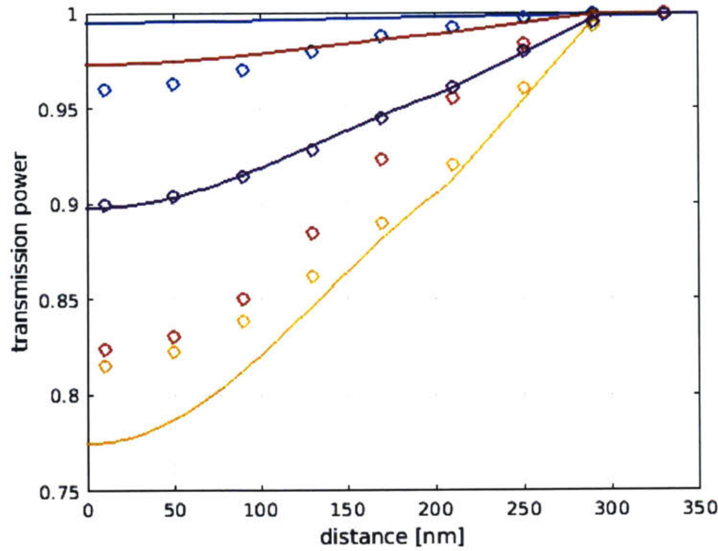



Figure 4-6: The impact of a metal sphere defect of $r$ = 40 nm on transmission of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple).

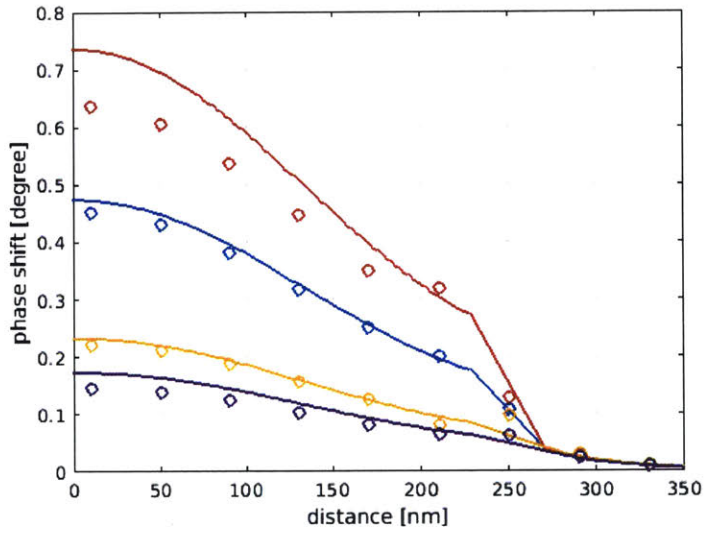



Figure 4-7: The impact of a metal sphere defect of $r = 20$ nm on phase shift of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple).

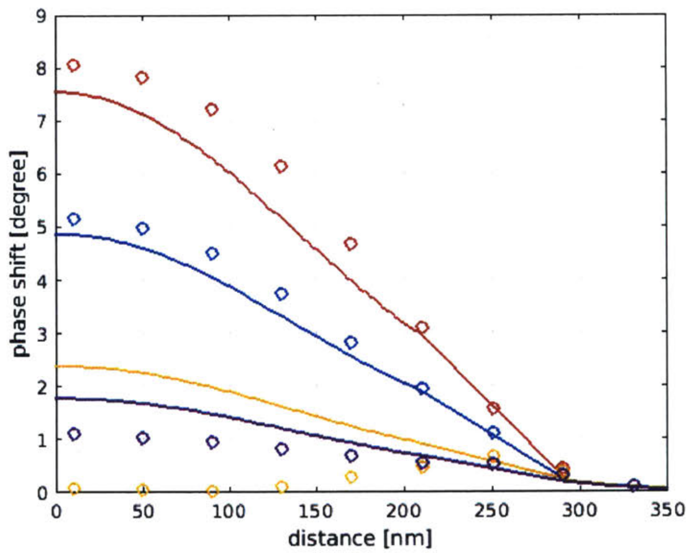



Figure 4-8: The impact of a metal sphere defect of $r = 40$ nm on phase shift of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple).

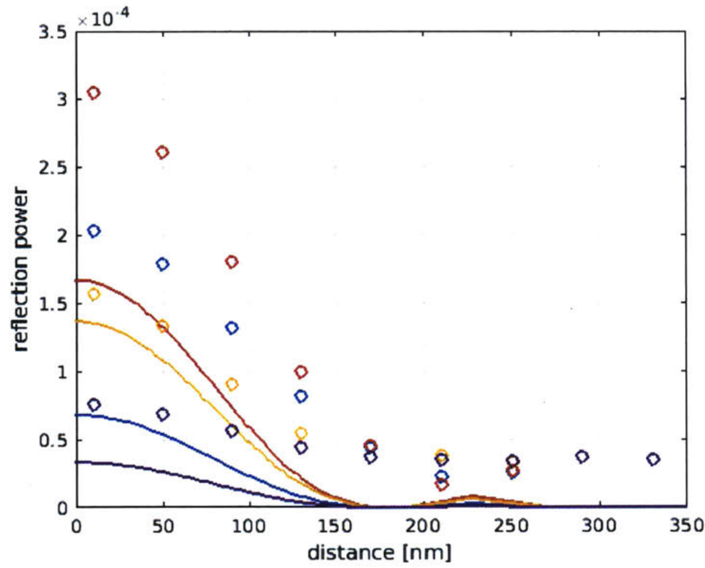



Figure 4-9: The impact of a metal sphere defect of $r = 20$ nm on back reflection of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple).

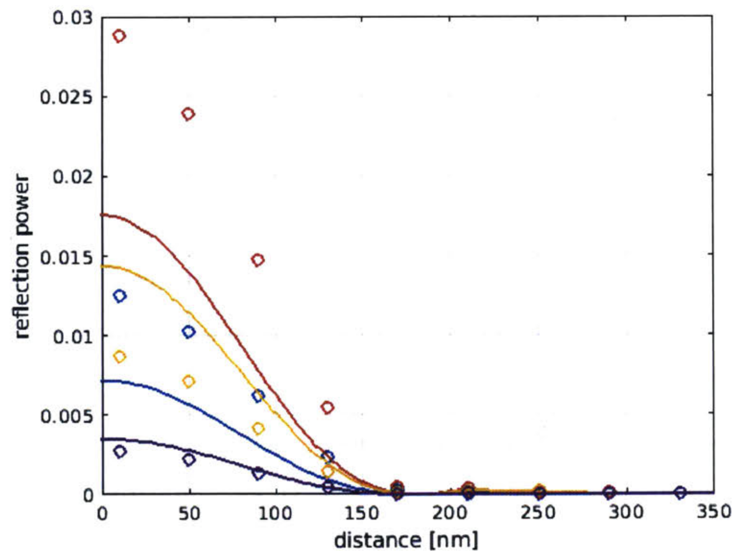



Figure 4-10: The impact of a metal sphere defect of $r = 40$ nm on back reflection of a straight waveguide, as a function of distance $d$, evaluated from direct simulation (circle) and adjoint method (solid line). Four different materials are tested: Aluminum (blue), Copper (red), Tungsten (yellow), and Titanium (purple).

break the law of the adjoint method where the impact is always proportional to the volume of the particle.

So far we see that for small particles, silicon dioxide hole pillar defects have the most impact on a straight waveguide, which happens to be the most accurate to predict from the adjoint method. For large particles, metal sphere defects can have stronger impact, but out of the appropriate range of the adjoint method. In the following sections and chapters, we thus only show examples of silicon dioxide hole pillar defects with $r = 20$ nm; the same methodology can be used for other types of defects, but we expect it to have either weaker impact (for small particles) or lower accuracy (for large particles).

## 4.2 Y-Splitter

In this section, we choose to model the impact of particle defects on the power-optimized Y-splitter in [27]. We label the input port as port 1, and the output ports as port 2 and 3. Based on the symmetry of the nominal design, we have

$$\begin{aligned} E_{3x}(x, y, z) &= -E_{2x}(x, -y, z) \\ E_{3y}(x, y, z) &= E_{2y}(x, -y, z) \\ E_{3z}(x, y, z) &= -E_{2z}(x, -y, z) \end{aligned} \tag{4.16}$$

and $S_{13} = S_{12}$, $S_{33} = S_{22}$. Therefore, a complete S-parameter sweep need only include simulation from port 1 and port 2, and the results for port 3 can be obtained by flipping the field of port 2. However, the particle defect analysis cannot skip port 3, as the introduced particle breaks the symmetry.

We will do a complete S-parameter analysis, but in this section we only focus on a few of them: the upper branch transmission $T_{12}$, phase shift $\phi_{12}$, and its self-reflection $R_2 = T_{22}$. The setup in FDTD is just the normal S-parameter sweep setup plus the field monitor (Figure 4-11); the port elements are used for both S-parameter calculation and necessary mode expansion coefficients for the adjoint method. In order to reduce memory cost, the field monitor is set to record only electric field and only at the wavelength of interest $\lambda = 1550$ nm.

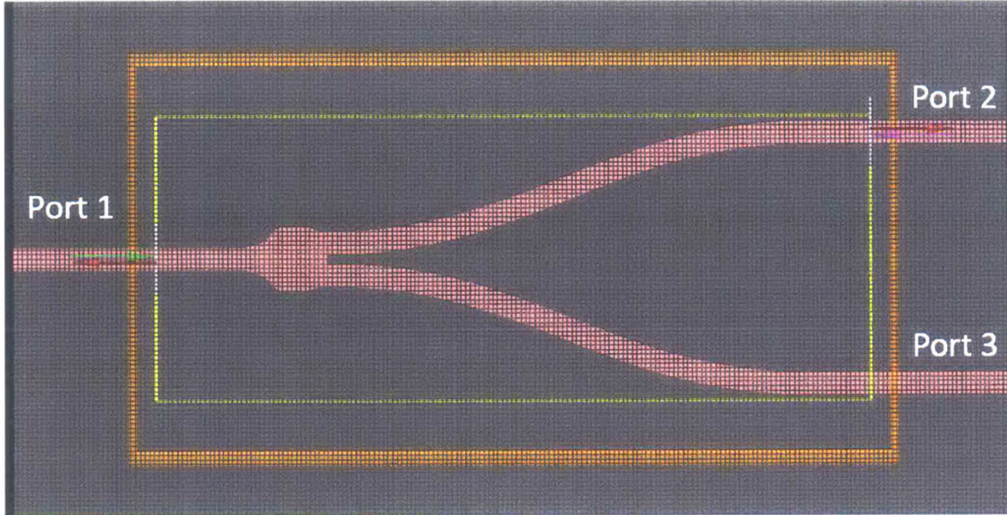



Figure 4-11: The implementation of the adjoint method on y-splitter in Lumerical FDTD simulation. The yellow box in the middle is the added field monitor. If the output of interests are only $T_{12}$, $\phi_{12}$ and $R_2$, the port element at port 3 is optional, while it is still required in complete S-parameters analysis.

We show the impact of a silicon dioxide hole pillar defect of $r = 20$ nm in Figure 4-12, Figure 4-14, and Figure 4-16. As we would expect, the particle in the upper branch contributes mostly to $\Delta\phi_{12}$ and less to $\Delta T_{12}$, which is similar to the case of straight waveguide. The particle in the cavity introduces asymmetry and thus impacts transmission distribution. Finally, the lower branch has little contribution to all three outputs of interest in the upper branch.

We choose the point at the cavity taper where the adjoint analysis shows maximum transmission impact in Figure 4-12 to do a direct simulation validation. We intentionally show the zero-impact axis in the comparison plot (Figure 4-13, Figure 4-15, and Figure 4-17) to get a better sense of relative error. Overall, the adjoint method shows excellent consistency with direct simulation; even for the back reflection of upper branch, where the impact has the same magnitude as the nominal value, the prediction from the adjoint method is still highly accurate.

## 4.3 Directional Coupler

We next consider a directional coupler (Figure 4-18) with coupling length $L_c = 4.5$ $\mu$m and gap width $w_g = 200$ nm which we will use to build a ring resonator in Chapter 6.

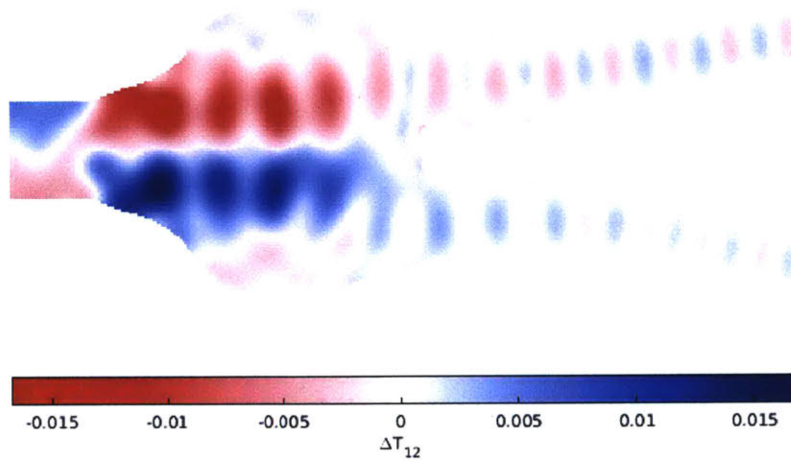



Figure 4-12: The mapping of the impact of a silicon dioxide hole pillar defect of $r$ = 20 nm on the upper branch transmission $T_{12}$ of the y-splitter structure, as a function of the spatial location of the defect. Zoom-in around the cavity.

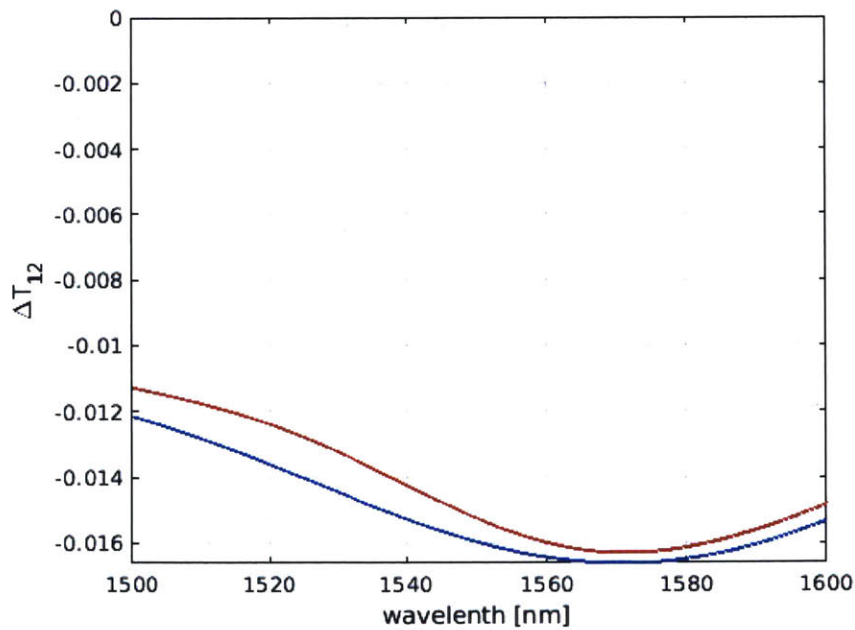



Figure 4-13: The impact of a silicon dioxide hole pillar defect of $r$ = 20 nm at selected location on the upper branch transmission $T_{12}$ of the y-splitter structure, evaluated from adjoint method (blue) and direct simulation (red).

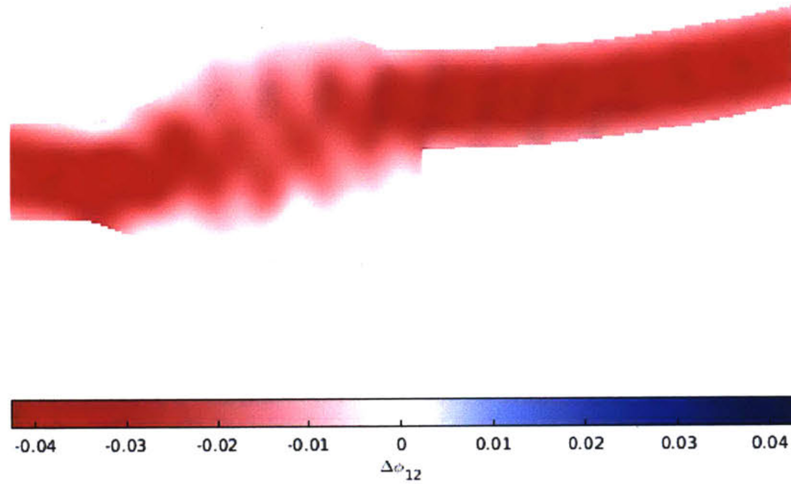



Figure 4-14: The mapping of the impact of a silicon dioxide hole pillar defect of $r$ = 20 nm on the upper branch phase shift $\phi_{12}$ of the y-splitter structure, as a function of the spatial location of the defect. Zoom-in around the cavity.

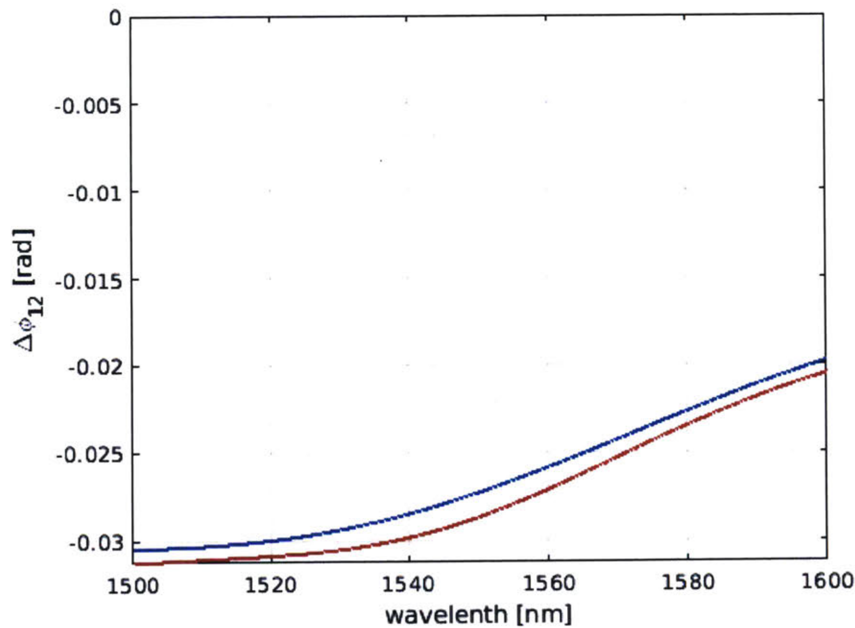



Figure 4-15: The impact of a silicon dioxide hole pillar defect of $r$ = 20 nm at selected location on the upper branch phase shift $\phi_{12}$ of the y-splitter structure, evaluated from adjoint method (blue) and direct simulation (red).

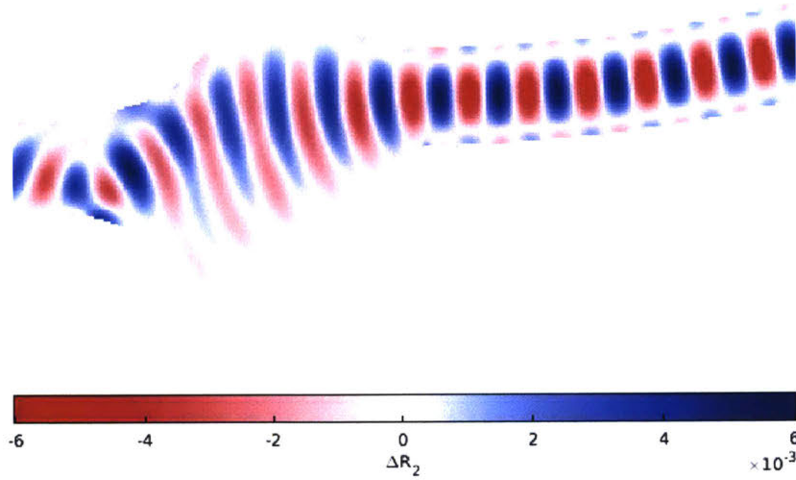



Figure 4-16: The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the upper branch back reflection $R_2$ of the y-splitter structure, as a function of the spatial location of the defect. Zoom-in around the cavity.

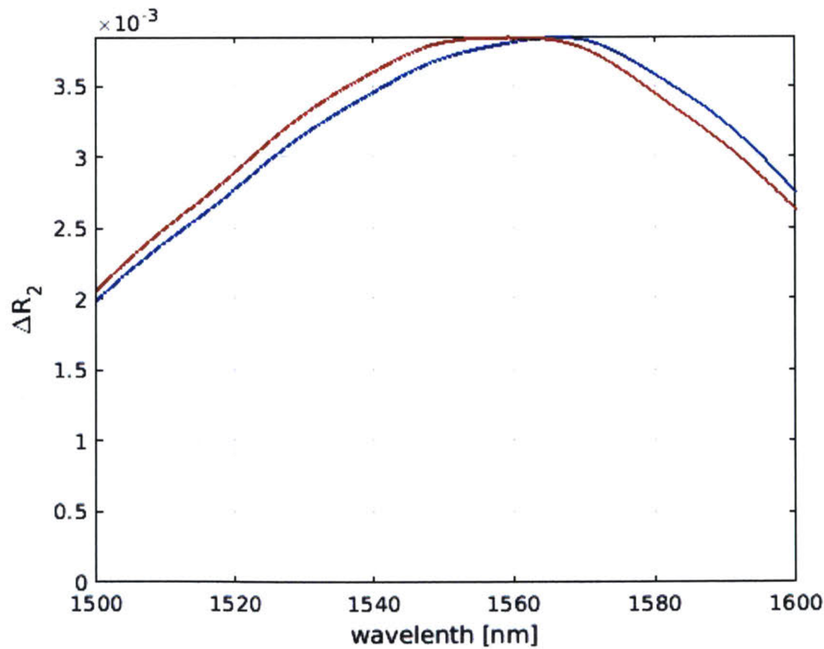



Figure 4-17: The impact of a silicon dioxide hole pillar defect of $r = 20$ nm at selected location on the upper branch back reflection $R_2$ of the y-splitter structure, evaluated from adjoint method (blue) and direct simulation (red).

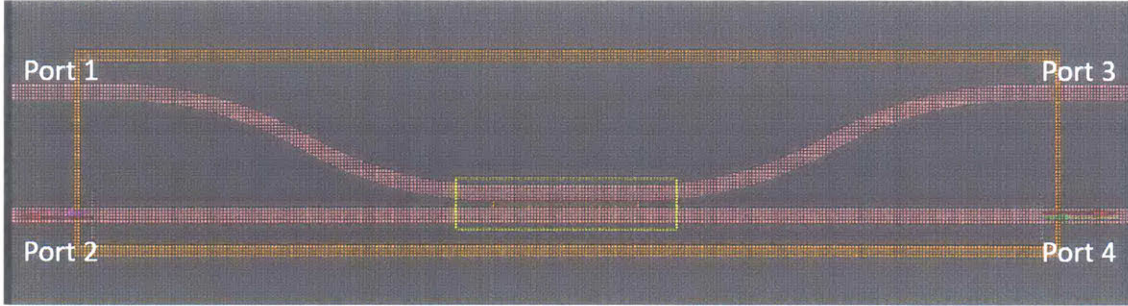



Figure 4-18: The implementation of the adjoint method on directional coupler in Lumerical FDTD simulation. Only port elements at port 2 and port 4 are required for analysis of self-coupling coefficient. The field monitor is set to only record around the gap to reduce memory cost.

We focus on the through-port transmission, also known as the self-coupling coefficient $t = |S_{24}|$ in this thesis. Similar to the y-splitter case, here we have symmetry in $x$-axis:

$$\begin{aligned} E_{4x}(x, y, z) &= -E_{2x}(-x, y, z) \\ E_{4y}(x, y, z) &= E_{2y}(-x, y, z) \\ E_{4z}(x, y, z) &= E_{2z}(-x, y, z) \end{aligned} \tag{4.17}$$

Therefore, we only need one simulation to obtain both $S_{24}$ and $\Delta S_{24}$.

We show the impact of a silicon dioxide hole pillar defect of $r = 20$ nm in Figure 4-19. The impact is much smaller compared to the y-splitter case in Section 4.2; it is most likely that the most sensitive area is in the gap, where the hole defect does not exist. However, a pillar defect or metal defect in the gap has relatively modest impact (Figure 4-20), for the reasons we discussed in Section 4.1.

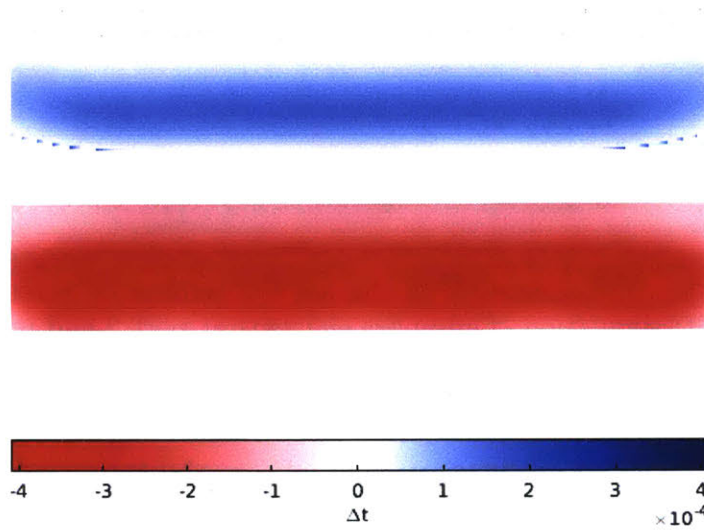



Figure 4-19: The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the self-coupling coefficient $t = |S_{24}|$ of the directional coupler, as a function of the spatial location of the defect. Zoom-in around the gap (the yellow box in Figure 4-18).

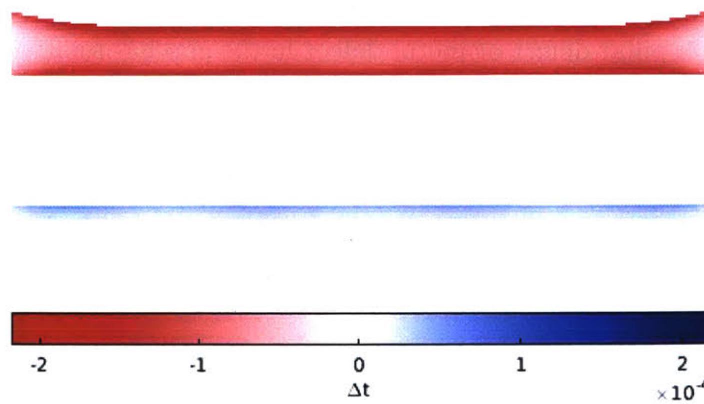



Figure 4-20: The mapping of the impact of a silicon pillar defect of $r = 20$ nm on the self-coupling coefficient $t = |S_{24}|$ of the directional coupler, as a function of the spatial location of the defect. Zoom-in around the gap (the yellow box in Figure 4-18).

# Chapter 5

# Wavelength Dependence

Before moving on to photonic circuit-level variation, we want to have a discussion about the implementation of the adjoint method over a range of wavelengths. This is necessary for evaluation of photonic devices and circuits, since most of the properties, e.g., free spectral range (FSR) and quality factor of a resonator, do not rely only on the performance at a single operating wavelength. The spectral behavior of photonic circuits can be quite complicated, and generally hundreds of frequency points are needed for frequency response interpolation, which requires excessively large amounts of memory if the adjoint method is used separately at all of these wavelengths. However, it is possible to break the circuit down into smaller components where each component has a simpler frequency response behavior, and thus fewer frequency points are needed for interpolation. As shown in Chapter 4, we interpolate the amplitude and argument of $\Delta S_{ij}$ separately. This will introduce phase wrapping problem for sparse frequency sampling, which will be discussed next in Section 5.1 and Section 5.2. Then we focus on how to further reduce the memory cost of the adjoint method in Section 5.3.

## 5.1 Group Delay

We start from a straight waveguide. For the electric field $\mathbf{E}_1$ in (4.1), we have its phase at a certain point $(x, y, z)$

$$\phi_1(\omega) = \arg\left(\mathbf{E}_1(x, y, z)\right) = n_e k x + \phi_0(\omega), \tag{5.1}$$

where $\phi_0(\omega) = \arg(s_1 \mathbf{E}_\mathrm{m}(y,z))$. Later we will see that our interest is actually in the normalized field $\mathbf{E}_1/s_1$, and $\arg(\mathbf{E}_\mathrm{m}(y,z))$ does not depend on wavelength, thus we focus on

$$\Delta\phi_1(\omega) = \phi_1(\omega) - \phi_0(\omega) = n_e k x = \frac{n_e \omega x}{c}. \tag{5.2}$$

This phase shift $\Delta\phi_1$ is a function of frequency $\omega$, not only because of the explicit $\omega$ in the equation, but also because $n_e$ depends on wavelength. Linear approximation gives

$$\Delta\phi_1(\omega) \approx \Delta\phi_1(\omega_0) + \tau_g(\omega_0)(\omega - \omega_0), \tag{5.3}$$

where the group delay[1]

$$\tau_g(\omega) = \frac{\mathrm{d}\Delta\phi_1}{\mathrm{d}\omega} \tag{5.4}$$

is also the transit time required for a narrow-band signal at $\omega$ to travel the distance $x$. Note that

$$\tau_g(\omega) = \left(n_e + \omega\frac{\mathrm{d}n_e}{\mathrm{d}\omega}\right)\frac{x}{c} = \frac{n_g x}{c}, \tag{5.5}$$

where $n_g$ is the group effective index, which has the value $n_g = 4.18$ at wavelength of 1550 nm for our straight waveguide with width 500 nm and height 220 nm. Thus the slope of $\Delta\phi_1(\omega)$ is easy to estimate without any simulation. For convenience, we use $\lambda$ instead of $\omega$ in the spectral analysis, and

$$\frac{\mathrm{d}\Delta\phi_1}{\mathrm{d}\lambda} = \tau_g(\lambda)\frac{\mathrm{d}\omega}{\mathrm{d}\lambda} = -\frac{2\pi n_g x}{\lambda^2}. \tag{5.6}$$

The above analysis is only accurate for straight waveguides, but it serves as a good estimation for all waveguide-like components, e.g., bends, splitters, and even couplers, by replacing $x$ with the minimal propagation length $l$ from the source to the observation point. For S-parameters $S_{ij}$, we have

$$\tau_{ij} = \frac{\mathrm{d}}{\mathrm{d}\omega}\phi_{ij} \approx \frac{n_g}{c} l_{ij}, \tag{5.7}$$

where $l_{ij}$ is the propagation length from port $i$ to port $j$. And for the perturbation

[1] The definition of group delay here has a difference of a negative sign compared to the one used in signal processing, which is because we choose the time-varying complex sinusoid to be $e^{-\mathrm{i}\omega t}$ in electromagnetics.

of S-parameters by particle defect in (3.17), [2]

$$\tau_{\Delta ij} = \frac{\mathrm{d}}{\mathrm{d}\omega}\arg(\Delta S_{ij}) \approx \frac{n_g}{c}(l_i + l_j), \tag{5.8}$$

where $l_i$ is the minimal propagation length from port $i$ to the location of the particle defect.

## 5.2 Phase Unwrapping

The periodic property of the phase causes multivalue in function $\arg(\cdot)$:

$$\arg(z) = \{\mathrm{Arg}(z) + 2\pi n \mid n \in \mathbf{Z}\}, \tag{5.9}$$

where $\mathrm{Arg}(z) \in (-\pi, \pi]$ is its principle value. Despite the arbitrariness in $\arg(\cdot)$, the phase relationship over the spectrum should be fixed, which means that

$$\phi(\omega) = \arg(S(\omega)) = \{\tilde{\phi}(\omega) + 2\pi n \mid n \in \mathbf{Z}\}. \tag{5.10}$$

The choice of $\tilde{\phi}(\omega)$ is not unique, but none of them satisfies $\tilde{\phi}(\omega) = \mathrm{Arg}(S(\omega))$: the range of $\tilde{\phi}(\omega)$ can be much larger than $2\pi$, but $\mathrm{Arg}(\cdot)$ always wrap it inside the small region $(-\pi, \pi]$. To obtain the appropriate $\tilde{\phi}(\omega)$, we need to unwrap $\mathrm{Arg}(S(\omega))$.

The most common approach to unwrap is to make use of the fact that $\tilde{\phi}(\omega)$ is continuous and assume the frequency sampling is dense so that the phase difference between consecutive frequencies is always smaller than $\pi$. Algorithm 1 shows how this approach works.

**Algorithm 1** Unwrap array $\phi_{1:N}$

**for** $i = 2$ to $N$ **do**
  $p \leftarrow \mathrm{round}((\phi_i - \phi_{i-1})/2\pi)$
  $\phi_i \leftarrow \phi_i - 2\pi p$
**end for**

To guarantee the algorithm to work properly, we need

$$|\delta\phi| = \tau_g |\delta\omega| < \pi, \tag{5.11}$$

[2] Note that $\arg(\Delta S_{ij}) \neq \Delta\arg(S_{ij}) = \Delta\phi_{ij}$.

which gives the restriction on frequency sampling,

$$|\delta\omega| < \frac{\pi}{\tau_g}, \tag{5.12}$$

or expressed in terms of wavelength,

$$|\delta\lambda| < \frac{\lambda^2}{2\tau_g c}. \tag{5.13}$$

For calculation of $S_{ij}$, this means

$$|\delta\lambda| < \frac{\lambda^2}{2n_g l_{ij}}. \tag{5.14}$$

For the y-splitter in Section 4.2, max $l_{ij} \approx 25$ $\mu$m, which gives $|\delta\lambda|$ approximately equal to 10 nm. This is easy to achieve in FDTD simulation, where typically $\delta\lambda <$ 1 nm. However, for the same structure, $\max(l_i + l_j) \geq \max l_{ij}$, thus the requirement is even stricter for $\delta S_{ij}$, which makes it impossible to cover a spectral range of 100 nm in just a few sample points.

However, we can make use of the group delay in the unwrap process if we have an estimation of it. As shown in Algorithm 2, for every frequency point, we can estimate its phase $\phi'$ based on the previous frequency point and the estimated group delay $\hat{\tau}_g$; then we can unwrap the current frequency point based on this estimation instead of the phase at previous frequency points.

**Algorithm 2** Unwrap array $\phi_{1:N}$ with $\omega_{1:N}$ based on estimated group delay $\hat{\tau}_g$

**for** $i = 2$ to $N$ **do**
  $\phi'_i \leftarrow \phi_{i-1} + (\omega_i - \omega_{i-1})\hat{\tau}_g$
  $p \leftarrow \text{round}((\phi_i - \phi'_i)/2\pi)$
  $\phi_i \leftarrow \phi_i - 2\pi p$
**end for**

For this improved algorithm, the limit of frequency sampling depends on how good the estimate $\hat{\tau}_g$ is:

$$|\phi - \phi'| = |(\hat{\tau}_g - \tau_g)\delta\omega| < \pi, \tag{5.15}$$

thus

$$|\delta\omega| < \frac{\pi}{|\hat{\tau}_g - \tau_g|}, \qquad |\delta\lambda| < \frac{\lambda^2}{2c|\hat{\tau}_g - \tau_g|}, \tag{5.16}$$

or if we fix the number of frequency points, this is the requirement on the error of estimation:

$$e = \frac{|\hat{\tau}_g - \tau_g|}{\tau_g} < \frac{\lambda^2}{2\tau_g c|\delta\lambda|} = \frac{|\delta\lambda|_{\min}}{|\delta\lambda|}, \tag{5.17}$$

where $|\delta\lambda|_{\min} = \lambda^2/2\tau_g c$ is actually the maximum value we can achieve for normal unwrap method (5.13); we denote it with the minimum because we can always have $0 \leq e \leq 1$, thus $|\delta\lambda| \geq |\delta\lambda|_{\min}$. We still use the y-splitter example where $|\delta\lambda|_{\min} \approx$ 10 nm: if we choose only three sample points, thus $\delta\lambda = 50$ nm, then the maximum error tolerance is about 20%; if we choose five sample points and $\delta\lambda = 25$ nm, then the maximum error tolerance is above 40%, which is easy to achieve.

## 5.3 Memory Reduction

So far in this chapter, we have sought to reduce frequency sample points in order to reduce the memory cost from the adjoint method. In this section, we consider another approach to reduce memory cost at each wavelength.

Most of the memory cost in the adjoint method arises from recording the electrical field profile, especially when the structure is large. However, the electrical field far outside the structure is almost zero, thus it is unnecessary to record these spaces. In FDTD simulations, the box-shaped field monitor can include substantial regions of such "useless" space. Using many small boxes instead of a large box to cover the structure can thus reduce memory size; we show an example of the implementation on y-splitter in Figure 5-1, which reduces about half of the original memory cost. From (3.17) we know that the impact of a particle defect in the adjoint method is localized; therefore, such splitting does not introduce any computational difficulty since each small box can be processed independently.

For pillar particles, the memory cost is extremely high because of the use of 3D monitors. For thickness of 220 nm and the highest mesh accuracy setting, there are 19 points in the $z$ direction to record, which means the memory cost is about 20 times higher than that for 2D monitors. In practice, such memory cost is so large that the simulation can be heavily slowed down or fail to complete.

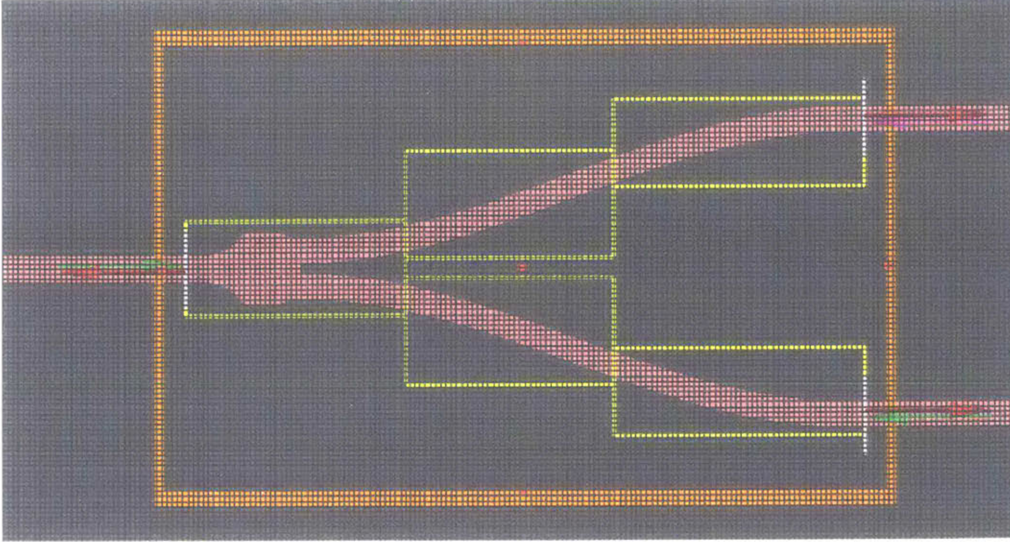


Figure 5-1: Demonstration of using a combination of small field monitors in the y-splitter example.

From the former discussion, data at different $z$ is used in the averaging process:

$$\gamma\mathbf{E}(x_p, y_p) \cdot \mathbf{E}^A(x_p, y_p) \approx \frac{1}{h}\int_0^h \gamma\mathbf{E}(x_p, y_p, z) \cdot \mathbf{E}^A(x_p, y_p, z)\,\mathrm{d}z. \tag{5.18}$$

In the sense of finite difference, this is actually a summation under the trapezoidal rule in our numerical computation:

$$\frac{1}{h}\int_0^h f(z)\,\mathrm{d}z \approx \frac{1}{36}\left(f_0 + 2\sum_{j=1}^{17} f_j + f_{18}\right). \tag{5.19}$$

Here we use notation $f(z) = \gamma\mathbf{E}(x_p, y_p, z)\cdot\mathbf{E}^A(x_p, y_p, z)$ and $f_j = f(j\Delta z)$, and assume that the sample points are evenly distributed in the $z$ axis. If we try to down sample in the $z$ axis, the accuracy of the trapezoidal rule will drop; but using higher order integral rules, we can minimize the error. For example, if we choose to down sample by 3 and use Simpson's rule, then

$$\frac{1}{h}\int_0^h f(z)\,\mathrm{d}z \approx \frac{1}{18}(f_0 + 4f_3 + 2f_6 + 4f_9 + 2f_{12} + 4f_{15} + f_{18}). \tag{5.20}$$

Noting the symmetry in $z$ axis for a slab waveguide, $f_j = f_{18-j}$, it can be further

written as

$$\frac{1}{h}\int_0^h f(z)\,\mathrm{d}z \approx \frac{1}{9}(f_0 + 4f_3 + 2f_6 + 2f_9), \tag{5.21}$$

which only needs four points in the $z$ direction. Such down-sampling scheme thus further reduces about 80% of the memory cost.

# Chapter 6

# Circuit Level Variation Analysis and Yield Modeling

In this chapter, we show an application of the adjoint method at the photonic circuit level, and how to use the result for yield modeling. Most photonic circuits can be split into small pieces of device components, which can be represented by their S-parameters. Thus the circuit-level simulation is accomplished by connecting component S-parameters; in other words, the output of the circuit is a function of the component S-parameters. Such function can be explicit, or implicit through INTERCONNECT simulation and even curve-fitting. Here we show an adjoint-based variation analysis methodology for two small-size "circuits": a Y-Mach-Zehnder interferometer (Y-MZI) and a ring resonator. These are often treated as device components in photonic circuit design, but can be broken down to smaller components; this enables us to use the same methodology asin larger photonic circuits in order to demonstrate the concept. We use the Y-MZI as an implicit example in Section 6.1, and the ring resonator as an explicit example in Section 6.2. The application to yield modeling is then shown in Section 6.3.

## 6.1 Y-MZI: Implicit Example

In this section we work with a simple balanced Y-MZI, where we connect two y-splitters shown in Section 4.2 back to back. We use the superscript to differentiate

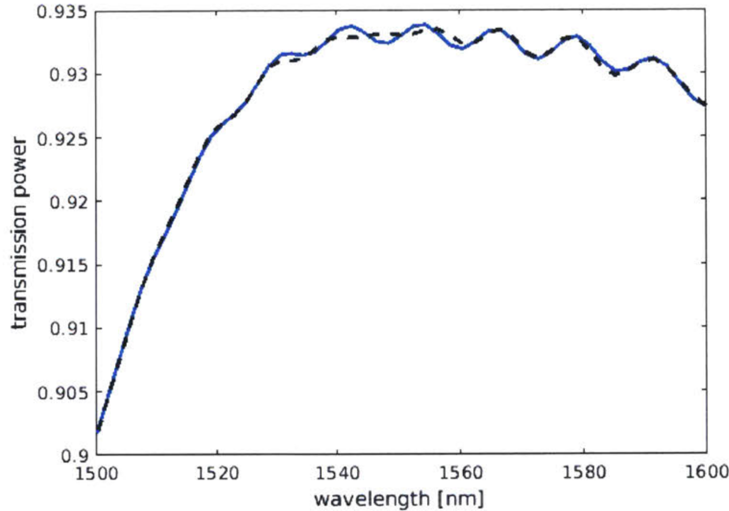



Figure 6-1: The transmission of a nominal Y-MZI (blue) and its curve-fitting result (dashed).

the two components, e.g., we use $S^{(1)}$ to represent S-parameters of the left y-splitter, and $S^{(2)}$ for the right y-splitter. As shown in Figure 6-1, the output transmission of the Y-MZI consist of a "smooth" curve that is similar to the transmission of y-splitter $S_{12}+S_{13}$, and a "ripple" component which is the result of the back-reflection, $S_{22}$, $S_{33}$, $S_{23}$. Our goal is to estimate the impact of particle defects on the average transmission power of the "smooth" component and the average amplitude of the "ripple" component, both over the wavelength range from 1500 to 1600 nm.

### 6.1.1 Adjoint Method for INTERCONNECT

We start by focusing on a single wavelength. We have already derived how to calculate $\Delta S^{(1)}(\lambda)$ and $\Delta S^{(2)}(\lambda)$ in the last two chapters, and now we seek to find the output transmission $\Delta T(\lambda)$. However, the relationship between $T$, $S^{(1)}$ and $S^{(2)}$ is implicit where $T(\lambda)$ is solved by INTERCONNECT simulation – which happens to be another linear system. Although applying direct simulation (i.e., do INTERCONNECT simulation for every possible particle location) might be feasible given that INTERCONNECT simulation is typically fast, it will be much more convenient if the adjoint method is applied on this level as well.

We first consider what equations the INTERCONNECT simulation is solving. For

the $k$-th component with $n$ ports, from the definition of S-parameters (2.1),

$$b^{(k)} = S^{(k)} a^{(k)}, \tag{6.1}$$

where $a^{(k)} = \left[a_1^{(k)}, \dots, a_n^{(k)}\right]^{\mathrm{T}}$ and $b^{(k)} = \left[b_1^{(k)}, \dots, b_n^{(k)}\right]^{\mathrm{T}}$. Combining all the components together, we define the grand input and output vector

$$a = \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(N)} \end{pmatrix}, \qquad b = \begin{pmatrix} b^{(1)} \\ \vdots \\ b^{(N)} \end{pmatrix}, \tag{6.2}$$

and the grand block-diagonal S-parameters matrix

$$S = \begin{pmatrix} S^{(1)} & & \\ & \ddots & \\ & & S^{(N)} \end{pmatrix}. \tag{6.3}$$

Then from (6.1) we simply have

$$b = Sa. \tag{6.4}$$

The connection between ports is equivalent to assigning output signal to input. For example, if we connect port $i$ of the $k$-th component to port $j$ of the $l$-th component, then $a_j^{(l)} = b_i^{(k)}$ and $a_i^{(k)} = b_j^{(l)}$. We can write this relationship as

$$a = Pb + f \tag{6.5}$$

where $P$ is the connection matrix, and $f$ is the source vector determining the input port of the circuit. From our example, we see that $P$ is symmetric: $P^{\mathrm{T}} = P$. $S$ is also symmetric because of the reciprocity of S-parameters. Combining (6.4) and (6.5), we have

$$\begin{pmatrix} I & -P \\ -S & I \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}. \tag{6.6}$$

The output $y$ is linear combination of $b$, thus

$$y = c^{\mathrm{T}} b = \begin{pmatrix} 0 \\ c \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} a \\ b \end{pmatrix}. \tag{6.7}$$

Applying adjoint method, we see that

$$\Delta y = -\begin{pmatrix} b_A \\ a_A \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} 0 & 0 \\ -\Delta S & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = a_A^{\mathrm{T}} (\Delta S) a \tag{6.8}$$

where

$$\begin{pmatrix} I & -S \\ -P & I \end{pmatrix} \begin{pmatrix} b_A \\ a_A \end{pmatrix} = \begin{pmatrix} 0 \\ c \end{pmatrix}, \tag{6.9}$$

or by switching rows and columns,

$$\begin{pmatrix} I & -P \\ -S & I \end{pmatrix} \begin{pmatrix} a_A \\ b_A \end{pmatrix} = \begin{pmatrix} c \\ 0 \end{pmatrix} \tag{6.10}$$

is exactly the same system but the source switched to the output. (6.8) can be written as a summation over components:

$$\Delta y = \sum_{k=1}^{N} a_A^{(k)\,\mathrm{T}} \left(\Delta S^{(k)}\right) a^{(k)}. \tag{6.11}$$

If there is only one particle defect in the whole circuit, only the $\Delta S$ term of the component with the defect is not zero; thus the calculation is independent for every component. For example, if we are only interested in the particle defect at the left part of the Y-MZI, then

$$\Delta y = a_A^{(1)\,\mathrm{T}} \left(\Delta S^{(1)}\right) a^{(1)} \tag{6.12}$$

which does not have $S^{(2)}$ explicitly. The other components only contribute in the calculation (simulation) of $a$ and $a_A$, which needs to be done only once for all the components.

We test the method by putting the silicon dioxide hole at the same location as in Section 4.2 on the left y-splitter, and compare the result $\Delta T = \Delta |y|^2$ from adjoint

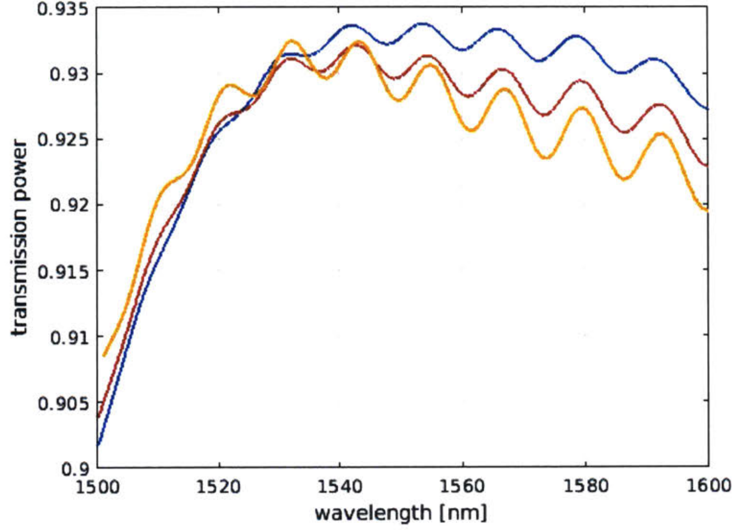



Figure 6-2: Comparison of the transmission of a nominal Y-MZI (blue) and the Y-MZI with defect, evaluated from adjoint method (red) and direct simulation (yellow).

method and direct simulation. We use the same $\Delta S^{(1)}$ in both methods to avoid error introduced from the FDTD-level adjoint method. Shown in Figure 6-2, the performance of the adjoint method approximates but does not closely match the direct simulation in this example, most likely because the balanced Y-MZI almost reaches its maximum transmission, thus the first order effect is already weak and the higher order terms become significant. The large variation in back-reflection we saw in Chapter 4.2 may also contribute to higher order effects.

### 6.1.2 Curve Fitting

We are sometimes more interested in the impact on some physical features based on the frequency spectrum (for example, free spectral range and Q factor in resonators) than in the response at every frequency point itself. A common approach to extract these features is through curve fitting. For example, we fit the Y-MZI spectrum with

$$f(\lambda;\beta) = \sum_{i=0}^{4} a_i P_i(\tilde{\lambda}) + \left(\sum_{i=0}^{3} b_i P_i(\tilde{\lambda})\right) \sin\left(\sum_{i=0}^{3} c_i P_i(\tilde{\lambda})\right) \tag{6.13}$$

where $\tilde{\lambda} = (\lambda - 1550 \text{ nm})/50 \text{ nm}$ is the normalized wavelength, and $P_i$ is the $i$-th order Legendre polynomial. Then $a_0$ represents the average transmission power of the "smooth" component, and $b_0$ represents the average amplitude of the "ripple" component.

We use the non-linear least-square fitting, which chooses $\beta = (a, b, c)$ that minimizes sum of squared error:

$$\beta^* = \arg\min_{\beta} \sum_i |f(\lambda_i; \beta) - T(\lambda_i)|^2 . \tag{6.14}$$

By linearizing $f(\lambda; \beta)$ near $\beta^*$, we get

$$\Delta\beta^* = (J^{\mathrm{T}} J)^{-1} \Delta T \tag{6.15}$$

where vector $\Delta T = [\Delta T(\lambda_1), \ldots, \Delta T(\lambda_K)]^{\mathrm{T}}$, and $J$ is the Jacobian matrix:

$$J_{ij} = \left. \frac{\partial f(\lambda_i; \beta)}{\partial \beta_j} \right|_{\beta=\beta^*} . \tag{6.16}$$

Notice that the linearization process further introduces error in the estimation, so in the worst case this may only serve as a rough approximation. The estimated impact on average transmission power $a_0$ and average ripple amplitude $b_0$ is shown in Figure 6-3 and Figure 6-4.

## 6.2 Ring Resonator: Explicit Example

The process is much easier if the physical feature is known as an explicit function of component S-parameters. For example, the extinction ratio $\mathcal{E}$ and finesse $\mathcal{F}$ of a ring resonator are functions of self-coupling coefficient $t$ and loss per cycle $\alpha$: [28]

$$\mathcal{E} = \left[ \frac{(\alpha + t)(1 - \alpha t)}{(\alpha - t)(1 + \alpha t)} \right]^2 , \tag{6.17}$$

$$\cos \frac{\pi}{\mathcal{F}} = \frac{2\alpha t}{1 + \alpha^2 t^2} . \tag{6.18}$$

The loss per cycle $\alpha$ is introduced mostly from waveguide propagation loss and

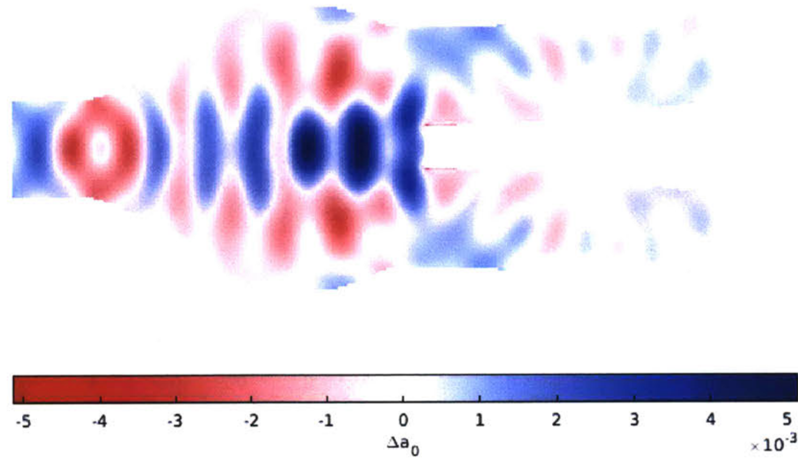



Figure 6-3: The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the average transmission power $a_0$ of the Y-MZI, as a function of the spatial location of the defect. Zoom-in around the cavity of the left y-splitter.

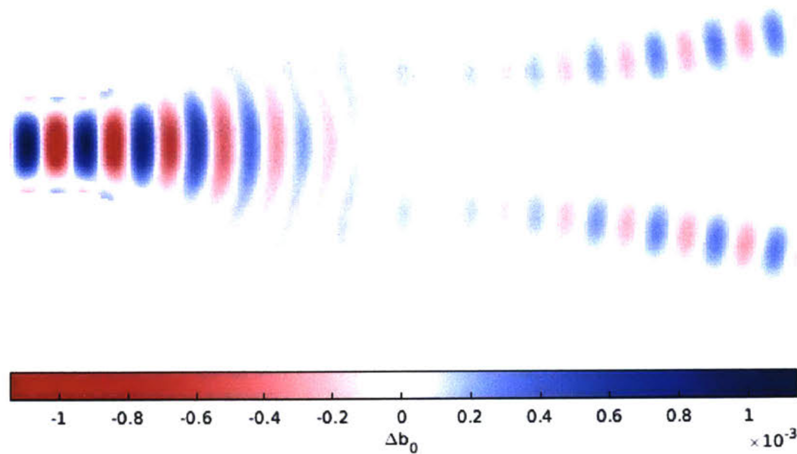



Figure 6-4: The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the average ripple amplitude $b_0$ of the Y-MZI, as a function of the spatial location of the defect. Zoom-in around the cavity of the left y-splitter.

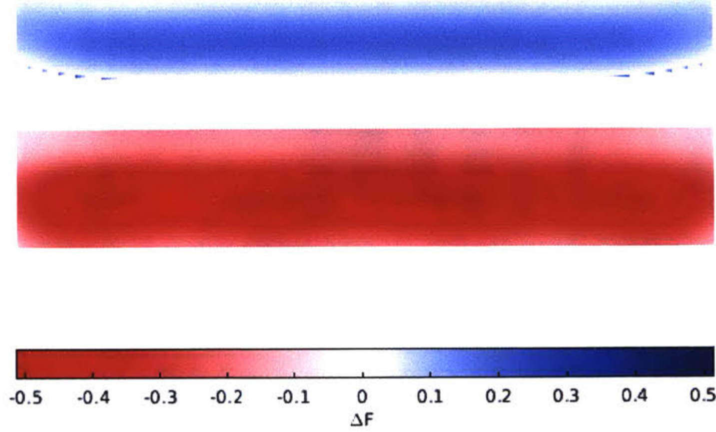



Figure 6-5: The mapping of the impact of a silicon dioxide hole pillar defect of $r = 20$ nm on the finesse of ring resonator $\mathcal{F}$ at $\lambda = 1550$ nm, as a function of the spatial location of the defect. Zoom-in around the gap.

scatter loss, and we assume that the hole defect does not significant impact $\alpha$. We also assume that $\alpha$ is very close to 1; so approximately

$$\cos\frac{\pi}{\mathcal{F}} = \frac{2t}{1+t^2}. \tag{6.19}$$

From the impact on self-coupling coefficient $\Delta t$ in Section 4.3, it is straight-forward to calculate the impact on finesse $\Delta F$, as shown in Figure 6-5. Since $\Delta t$ is small, the impact on finesse is also negligible.

## 6.3 Yield Modeling

There are typically metrics based on the important features of the output response used to evaluate the performance of the circuit; based on these, a threshold can be set to determine if the circuit is functional or not. For the particle defect, such threshold will enable us to extract the critical area $A_0$ in which the defect occurring will cause failure of the system.

As an example, we want to control the amplitude of the "ripple" in the Y-MZI

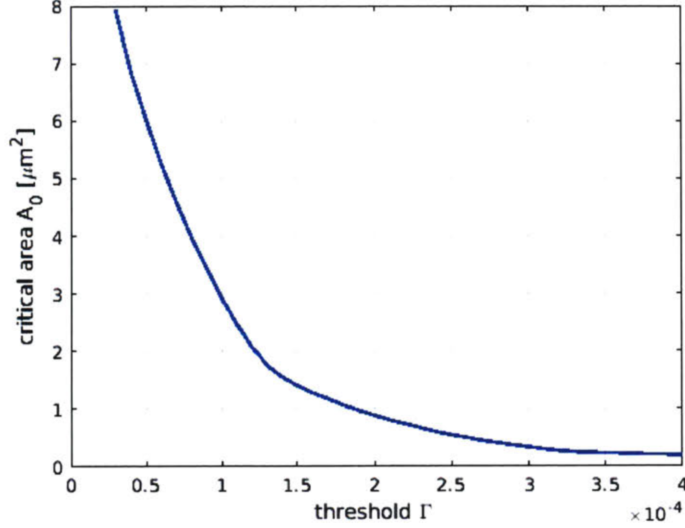



Figure 6-6: The critical area $A_0$ of the Y-MZI for particle radius $r = 20$ nm, as a function of threshold $\Gamma$, where we treat $\Delta b_0 > \Gamma$ as failure.

transmission spectrum, i.e., we treat $\Delta b_0 > \Gamma$ as failure. The critical area $A_0$ is thus a function of particle radius $r$ and threshold $\Gamma$. We plot the case of $r = r_0 = 20$ nm in Figure 6-6. Since the impact is always proportional to the volume of the particle in the adjoint method approximation, we have

$$A_0(r, \Gamma) = A_0\left(r_0, \frac{V_0}{V}\Gamma\right) \tag{6.20}$$

where $V \propto r^2$ for pillar defect and $V \propto r^3$ for sphere defect. This shows that cases of other particle radius can be obtained by stretching the plot in Figure 6-6, without redoing the calculation.

After the critical area extraction, the defect yield can be modeled based on defect density $D_0$ and critical area $A_0$: [18, 19]

$$Y = f(A_0, D_0), \tag{6.21}$$

where the function $f$ depends on the defect distribution in the model. For example, the simplest Possion model gives

$$Y = e^{-A_0 D_0}. \tag{6.22}$$

# Chapter 7

# Conclusions and Future Work

In conclusion, we adapt and implement the adjoint method for application to particle defect modeling in photonic devices and circuits, which provides relatively accurate impact prediction with modest computational cost. We demonstrate the methodology by modeling the impact on small device components as well as the impact on photonic circuits composed of smaller components. We also demonstrate application to model the yield impact of particle defects on the circuit. We find that the impact of different types of particle defects differ because of their shapes and material properties, and different components and different circuits can suffer very different degree of impact from the same particle defect.

The adjoint method based on S-parameters proposed in Section 3.1 can be used for not only particle defects, but many other kinds of variation in silicon photonics, such as line-edge-roughness. Although sometimes direct simulation is feasible for some variations, it only requires additional memory cost to add an adjoint analysis on top of direct simulation in order to gain first-order derivative information about the impact.

In Section 5.2 we present a condition on frequency sampling in order to correctly unwrap the phase results, and propose a method that enables phase unwrapping for even fewer frequency sample points. Both the condition and the new unwrapping method are useful for the wavelength-dependent design simulation in silicon photonics.

We also identify cases where the adjoint method is not perfect and causes large

error. The silicon pillar defect is hard to model by the adjoint method because the maximal impact happens to be at a condition where the dipole approximation breaks down. The relatively large metal defect is also troublesome because of the skin effect. Even when the defect is appropriate for analysis using the adjoint method, the output feature might have small first order derivatives, in which case higher order effects cannot be neglected. Fortunately, except for the limitation for large particles, large error only occurs when the actual impact is already small; thus the adjoint method is still a good approach to estimate the impact of small particle defects.

Future work could focus on more complicated yield modeling and design rule checking. In Section 6.3 we briefly discuss the yield loss caused by defects of a specific type with specific size, but in real situation we would want to bring all of the different defect types and different defect sizes together for yield modeling. The corresponding design rule checking needs further development for silicon photonics, as it is not so clear how the impact of defects changes with the design. However, as we find in Chapter 4 where the directional coupler is much more robust to particle defects than the y-splitters, there might be some simple rules taht can help to reduce the impact of particle defects.

As for the adjoint method, exploit its potential in analysis of other kinds of variation, and to consider how it might facilitate efficient analysis in the entire variation field.

# Bibliography


[1] Derek M. Kita, Hongtao Lin, Anu Agarwal, Kathleen Richardson, Igor Luzinov, Tian Gu, and Juejun Hu. On-chip infrared spectroscopic sensing: redefining the benefits of scaling. *IEEE Journal of Selected Topics in Quantum Electronics*, 23(2):340–349, 2016.

[2] Lukas Chrostowski and Michael Hochberg. *Silicon Photonics Design: From Devices to Systems.* Cambridge University Press, 2015.

[3] Chen Sun, Mark T. Wade, Yunsup Lee, Jason S. Orcutt, Luca Alloatti, Michael S. Georgas, Andrew S. Waterman, Jeffrey M. Shainline, Rimas R. Avizienis, Sen Lin, Benjamin R. Moss, Rajesh Kumar, Fabio Pavanello, Amir H. Atabaki, Henry M. Cook, Albert J. Ou, Jonathan C. Leu, Yu-Hsin Chen, Krste Asanović, Rajeev J. Ram, Miloš A. Popović, and Vladimir M. Stojanović. Single-chip microprocessor that communicates directly using light. *Nature*, 528(7583):534–538, 2015.

[4] Derek M. Kita, Brando Miranda, David Favela, David Bono, Jérôme Michon, Hongtao Lin, Tian Gu, and Juejun Hu. High-resolution on-chip digital fourier transform spectroscopy. In *CLEO: Science and Innovations*, pages SF1A–1. Optical Society of America, 2018.

[5] Michael Orshansky, Sani Nassif, and Duane Boning. *Design for Manufacturability and Statistical Design - A Constructive Approach.* Springer Science & Business Media, 2008.

[6] Kevin K. Lee, Desmond R. Lim, Hsin-Chiao Luan, Anuradha Agarwal, James Foresi, and Lionel C. Kimerling. Effect of size and roughness on light transmission in a Si/SiO2 waveguide: Experiments and model. *Applied Physics Letters*, 77(11):1617–1619, 2000.

[7] Francesco Morichetti, Antonio Canciamilla, Carlo Ferrari, Matteo Torregiani, Andrea Melloni, and Mario Martinelli. Roughness induced backscattering in optical silicon waveguides. *Physical Review Letters*, 104(3):1–4, 2010.

[8] A. D. Simard, N. Ayotte, Y. Painchaud, S. Bédard, and S. LaRochelle. Impact of sidewall roughness on integrated Bragg gratings. *Journal of Lightwave Technology*, 29(24):3693–3704, 2011.

[9] Sally I. El-Henawy, Ryan Miller, and Duane S. Boning. Effects of a random process variation on the transfer characteristics of a fundamental photonic integrated circuit component. In *Proc. of SPIE*, volume 10743, pages 107430O–1, 2018.

[10] Tsui-Wei Weng, Zheng Zhang, Zhan Su, Youssef Marzouk, Andrea Melloni, and Luca Daniel. Uncertainty quantification of silicon photonic devices with correlated and non-Gaussian random parameters. *Optics Express*, 23(4):4242, 2015.

[11] Géraud Blatman and Bruno Sudret. Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6):2345–2367, 2011.

[12] Dario Cassano, Francesco Morichetti, and Andrea Melloni. Statistical Analysis of Photonic Integrated Circuits Via Polynomial-Chaos Expansion. *Advanced Photonics 2013*, (2):JT3A.8, 2013.

[13] Germain Martinez. Methods for Compact Modeling of Process Variations in Silicon Photonics Devices. Master's thesis, Massachusetts Institute of Technology, 2018.

[14] Daniel H. Moon. Modeling Silicon Photonics Process Variations using Ring Resonator Devices. Master's thesis, Massachusetts Institute of Technology, 2018.

[15] Owen D. Miller. Photonic design: From fundamental solar cell physics to computational inverse design. *arXiv preprint arXiv:1308.0212*, 2013.

[16] Christopher M. Lalau-Keraly, Samarth Bhargava, Owen D. Miller, and Eli Yablonovitch. Adjoint shape optimization applied to electromagnetic design. *Optics Express*, 21(18):21693, sep 2013.

[17] Armand C. R. Niederberger, David A. Fattal, Nicolas R. Gauger, Shanhui Fan, and Raymond G. Beausoleil. Sensitivity analysis and optimization of subwavelength optical gratings using adjoints. *Optics Express*, 22(11):12971–12981, 2014.

[18] Gary S. May and Costas J. Spanos. *Fundamentals of Semiconductor Manufacturing and Process Control.* John Wiley & Sons, 2006.

[19] Charles H. Stapper and Raymond J. Rosner. Integrated circuit yield management and yield analysis: Development and implementation. *IEEE Transactions on Semiconductor Manufacturing*, 8(2):95–102, 1995.

[20] A. V. Ferris-Prabhu. Role of defect size distribution in yield modeling. *IEEE Transactions on Electron Devices*, 32(9):1727–1736, 1985.

[21] Allan W. Snyder and John D. Love. *Optical Waveguide Theory.* Springer US, Boston, MA, 1984.

[22] Ardavan Oskooi and Steven G. Johnson. Electromagnetic Wave Source Conditions. *arXiv preprint arXiv:1301.5366*, 2013.

[23] Ronald M. Errico. What Is an Adjoint Model? *Bulletin of the American Meteorological Society*, 78(11):2577–2591, nov 1997.

[24] Yang Cao, Shengtai Li, Linda Petzold, and Radu Serban. Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution. *SIAM Journal on Scientific Computing*, 24(3):1076–1089, 2003.

[25] F. Branin. Network sensitivity and noise analysis simplified. *IEEE Transactions on Circuit Theory*, 20(3):285–288, 1973.

[26] David J. Griffiths. *Introduction to electrodynamics.* Prentice Hall New Jersey, 1962.

[27] Yi Zhang, Shuyu Yang, Andy Eu-Jin Lim, Guo-Qiang Lo, Christophe Galland, Tom Baehr-Jones, and Michael Hochberg. A compact and low loss Y-junction for submicron silicon waveguide. *Optics Express*, 21(1):1310–6, 2013.

[28] W. R. McKinnon, D.-X. Xu, C. Storey, E. Post, A. Densmore, A. Delâge, P. Waldron, J. H. Schmid, and S. Janz. Extracting coupling and loss coefficients from a ring resonator. *Optics express*, 17(21):18971–18982, 2009.