A Hierarchical Approach to the Design and Optimization of Photonics

Andrew Michaels

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2019-163
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-163.html

December 1, 2019

A Hierarchical Approach to the Design and Optimization of Photonics

By

Andrew S Michaels

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eli Yablonovitch, Chair
Professor Ming Wu
Professor Javad Lavaei

Spring 2019

Abstract

A Hierarchical Approach to the Design and Optimization of Photonics

by

Andrew S Michaels

Doctor of Philosophy in Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Eli Yablonovitch, Chair

Over the last two decades, silicon photonics has rapidly matured, leading to a growing interest in building large complex systems consisting of thousands of integrated optical components. A direct consequence of this push towards large scale integration is the need for high efficiency silicon photonic building blocks.

In this work, we present a concrete path towards realizing these essential photonic building blocks. The foundation of our approach to designing photonic components is gradient-based shape optimization, which employs *boundary smoothing* based on high numerical precision polygons. In addition to helping us calculate accurate device sensitivities, this method affords us a great amount of flexibility when representing device geometries and enables us to incorporate design constraints directly into optimizations in a simple and intuitive way.

Our approach to gradient-based optimization shares an important similarity to other forms of shape and topology optimization employed in the nanophotonics community: on its own, it is not a complete solution to designing high performance and robust devices. Due to the inherently non-convex nature of electromagnetic optimization problems, we cannot expect convex optimization to universally yield good devices without outside input. In order to overcome this obstacle, we have systematized the process of providing "outside input" through our hierarchical approach to design and optimization. Using a strategic combination of simple physical analysis to find good starting geometries and optimization with both coarse and fine parameterizations, we show that efficient and robust devices can be designed with minimal guesswork.

Using this hierarchical approach, we demonstrate how a variety of silicon photonic components can be designed with superior performance. In particular, we design three port splitters, broadband four port splitters, fabrication-insensitive waveguide crossings, and a variety of efficient grating couplers which set a new standard for device performance. These components form the foundation for an optimized silicon photonic component library which will be important for demanding applications of the future.

To Alan Michaels and Raymond Steil

Contents

List of Figures

Acknowledgments

There are many people I have to thank for aiding me in my pursuit of this degree and my path through life as a whole. First and foremost, I must thank my parents for all of the love and support they have provided me throughout my studies. Without their advice and at-times-difficult-to-digest wisdom, it is clear to me that I would not be the person, researcher or otherwise, that I am today. From the beginning, both have served as the ultimate role models, and for that, I am forever indebted.

Next, I would like to thank my brother Aaron. While I doubt he has any real idea what I have been doing for the last five years, he has nonetheless always been a role model and balancing force to my overly scientific disposition.

Closer to the Bay Area, I would like to give a shout-out to Anne Hsiao and Efthymios Papageorgious. Everyone needs someone to commiserate with and vent to, and Anne and Phil have been there every step of the way.

Similarly, I must thank the members of the Yablonovitch group, past and present. Thay have all been a tremendous source of camaraderie and have always been a reliable source of support. Without their diverse perspectives, constant feedback, heated lunch time debates, and excessive coffee breaks, I am confident that my degree would not have been as fulfilling as it has.

When it comes to developing as a researcher, I do not believe I can overstate the impact my adviser, Eli Yablonovitch, has had on me. Eli has taught me how to think critically and independently. He has instilled in me the understanding that complexity and sophistication are often at odds with true comprehension. Finally, he has helped exposed me to the rich world that lies outside of academia and research. Each aspect of our interactions over the last five years has left a lasting mark, and for that I will always be grateful.

Lastly, I would like to thank the two individuals who have impacted me on this journey the most: my grandfathers Alan Michaels and Raymond Steil. My grandfather Alan, whom I knew only a short while as a child, imparted on me an intense desire to create and invent. Although growing up I understood little of his success as a professor and leader in industry, his legacy has left an immense impact which will forever push me to be the best that I can. My grandfather Ray, also an educator, is the hardest working and most genuine person I have ever met. I have always and will always aspire to follow his example.

Chapter 1

Introduction

As integrated photonics has matured, we have witnessed a growing desire for more compact and efficient passive optical components. At the same time, our ability to satisfy these demands by using either analytic methods or by tuning only a small set of device parameters is becoming increasingly difficult. To combat this trend, more sophisticated numerical optimization methods have been proposed.

Numerical optimization represents a significant departure from more conventional design methodologies. Whereas designing an electromagnetic device "by hand" involves intuiting the final device geometry based on well understood physics, optimization largely abstracts the design process into a *search* problem. In optimization, we answer the question: how do we modify our current design to improve its performance? Given some metric for performance and a description of the anatomy of the device, repeatedly answering this question yields a general way of designing electromagnetic devices which works for both simple and complex geometries alike.

Of the various optimization methods which are applicable to electromagnetics, gradient-based shape and topology optimization are particularly promising as they enable us to efficiently design devices with an almost arbitrary number of degrees of freedom. In both methods, a figure of merit is defined which quantifies the performance of the device, and then its gradient is computed in order to determine how the device should be modified. In the case of shape optimization, the boundaries between different materials composing an initial structure are modified in order to minimize the figure of merit [1]. Topology optimization operates in a similar manner; however, fewer constraints are placed on how the structure can evolve and in particular how its topology may be modified (e.g., the creation or elimination of holes) [2].

Shape and topology optimization have found extensive application in structural mechanics, and work on both methods dates back more than 30 years [1–4]. Despite the maturity of this field, less application of these methods has found its way into the optics and photonics community. Early work on gradient-based optimization of microwave devices [5, 6] demonstrated application of shape optimization

to electromagnetics. This work was followed by the application of topology optimization to photonic crystals [7] and later to a greater variety of passive photonic components [8]. Within the last five years, however, the photonics community has witnessed a steady rise in interest in these optimization techniques as demonstrated by the numerous optimizations of efficient splitters, couplers, etc [9–18].

A large portion of this prior work has focused on implementing topology optimization techniques and has emphasized problems in which the design space consists of thousands or even millions of parameters. In particular, the idea of choosing the permittivity at each point in a discretized domain as independent design variables has gained popularity. While this “grayscale" approach has proven useful for generating solutions without requiring significant intuition about the appearance of the final structure, it suffers a few disadvantages. First, in order to ensure that the device can be fabricated, a final post-processing step is often required to convert a grayscale material distribution to a binary material distribution [19]. As a result, the final solution may not be truly optimal or may require additional shape optimization steps. Next, these grayscale methods tend to produce structures with the same characteristic geometry consisting of holes and islands. This geometry may not be well matched to all problems; for example there are many problems in which either the geometry of the structure is constrained in some way (for example, a one dimensional grating coupler consisting of strictly rectangular segments) or for which topological changes are unnecessary. Finally, topology optimization inherently leverages a huge number of degrees of freedom. While this is generally viewed as a strength, it can be a double-edged sword, and there are many problems that may benefit from a smaller set of design parameters.

Shape optimization serves as a remedy to these problems. Shape optimization can be used to fine-tune the result of the topology optimization. For design problems in which the general shape of the structure is known beforehand, shape optimization alone is a natural fit. Shape optimization has an additional benefit in that it gives us considerable freedom over the choice of design space—we are free to choose arbitrary parameters, such as the length of a rectangle or the positions of polygon vertices, depending on the requirements of the problem.

In the first half of this work, we review the foundations of gradient-based optimization for electromagnetics and highlight some of the important caveats which are relevant to both shape and topology optimization. We then introduce an intuitive approach to shape optimization based on boundary smoothing techniques which is compatible with most design problems in electromagnetics.

Depending on the specific device we design, one implementation of shape or topology optimization may have advantages over another. Ultimately, however, gradient-based optimization is merely a tool, and the way in which we use that tool has tremendous influence on the final results we are able to achieve. Even in the simplest optimizations, we have to make at least three important choices which impact the final outcome: how device performance is defined, how the geometry is

defined, and the initial state of that geometry. From the standpoint of engineering a *good* device, it is essential that we approach this decision-making process in a systematic way.

In the latter half of this work, we introduce a hierarchical approach to design and optimization which systematizes the process of shape optimization. This approach relies on a strong physical understanding of Maxwell's equations to choose initial conditions for optimizations and leverages both small and large design spaces to ensure that high quality devices are generated. Based on this methodology, we demonstrate how a wide variety of important components in silicon photonics can be designed with record high efficiencies. Although we focus on silicon photonics, the same methods can be applied to many different electromagnetic platforms.

1.1 How to Read This Manuscript

This thesis is roughly divided into two parts. The first half of the document is dedicated to reviewing the optimization methods which underpin the work discussed in the second half of the thesis. This second half focuses on how to apply these optimization methods in order to design efficient nanophotonic components and presents a number of examples in silicon photonics.

If you would like a tutorial on the optimization methods, skim through Chapter 2 and Chapter 3. If you want a tutorial on how to apply existing optimization tools to nanophotonic components, skip ahead to Chapter 4. If you are only interested in efficient silicon photonic component designs, skip ahead to Chapter 5 and skim over the figures.

Chapter 2

Gradient-Based Shape Optimization of Electromagnetic Devices

Optimization (or "inverse design") of an electromagnetic device involves modifying an initial design in order to maximize (or minimize) a figure of merit which describes the device's performance. In order to modify the design, we specify design parameters which define the shapes of material boundaries and hence the structure of the device. There are many strategies for choosing the values of our design parameters in order to maximize our figure of merit; of particular interest are gradient-based minimization techniques which are very efficient at finding local optima assuming we can inexpensively compute the gradients of our figure of merit (i.e. the derivative of the figure of merit with respect to each design parameter). The simplest way to do this is by a brute force approach in which each design parameter is independently varied and a separate simulation is run in order to determine how the figure of merit changes. If M design variables specify the shape of the device, then this method requires $M + 1$ simulations per gradient calculation, which quickly becomes impractical for even modest values of M.

It turns out that through clever application of the chain rule, we can find the gradient of our figure of merit using only two simulations. This process is called the *adjoint method* or the *adjoint variable method* and has found widespread application in many fields, including electromagnetics [4–8, 10–12, 15–23]. In this chapter, we present a detailed derivation of the adjoint method and discuss how it can be used to compute accurate sensitivities of electromagnetic structures. Furthermore, we highlight the important caveats of the method and present examples to demonstrate how the methods can be applied to the optimization of nanophotonic devices.

2.1 Calculating Gradients with the Adjoint Method

Typically when "optimizing" an electromagnetic device, our goal is to minimize a figure of merit $F(\mathbf{E}, \mathbf{H})$ which depends on the electric ($\mathbf{E}$) and magnetic ($\mathbf{H}$) fields in some region of space. The fields are found by solving Maxwell's equations,

$$\nabla \times \mathbf{E} - i\omega\mu\mathbf{H} = \mathbf{M} \tag{2.1}$$

$$\nabla \times \mathbf{H} + i\omega\varepsilon\mathbf{E} = \mathbf{J} \tag{2.2}$$

which we have chosen to write in their time harmonic form as it simplifies the process of computing gradients. As is the case with most linear partial differential equations, we can discretize these equations on a rectangular grid and rewrite them in matrix form as

$$A\vec{x} = \vec{b} \tag{2.3}$$

where A is a matrix containing the discretized curls and the permittivity and permeability at each point in the discretized space, $\vec{x}$ is a vector containing the electric and magnetic fields at all points in space, and $\vec{b}$ contains the inhomogeneous electric and magnetic current density at all points in space. Depending on how unknowns are ordered when assembling the discretized system of equations, A, $\vec{x}$, and $\vec{b}$ might take the form

$$A = \begin{pmatrix} i\omega\varepsilon(\mathbf{r}) & \nabla\times \\ \nabla\times & -i\omega\mu(\mathbf{r}) \end{pmatrix},\ \vec{x} = \begin{pmatrix} \vec{E} \\ \vec{H} \end{pmatrix},\ \vec{b} = \begin{pmatrix} \vec{J} \\ \vec{M} \end{pmatrix} \tag{2.4}$$

where $\vec{E}$, $\vec{H}$, $\vec{J}$, and $\vec{M}$ are $N \times 1$ vectors containing the discretized field and current density values. In our discretized world, our figure of merit becomes a function of $\vec{x} = [\vec{E},\ \vec{H}]^T$ which we write as $F(\vec{x})$.

If we had direct control over the electric and magnetic fields in $\vec{x}$, finding the gradient of $F(\vec{x})$ would require only simple differentiation with respect to $\vec{E}$ and $\vec{H}$. Instead of directly controlling $\vec{E}$ and $\vec{H}$, we have control over the permittivity and permeability defined everywhere in space. In many cases, defining material distribution explicitly is not desirable; instead it is often convenient to define the boundaries of different material regions by using a structured set of design parameters (like the dimensions of shapes, the positions of shapes, the coordinates of polygon vertices, etc). If we write this set of design variables as $\vec{p} = [p_1, p_2, \cdots, p_M]^T$, then the gradient we are interested in is the set of derivatives of F with respect to each p_i, i.e.

$$\vec{\nabla}_p F = \left[\frac{dF}{dp_1}, \frac{dF}{dp_2}, \cdots, \frac{dF}{dp_M}\right] \tag{2.5}$$

To find these derivatives, we begin by applying chain rule when differentiating $F(\vec{x})$. Consider the i'th derivative for the simple case in which F is an explicit function of the electric and magnetic fields only:

$$\begin{aligned}\frac{dF}{dp_i} &= \frac{\partial F}{\partial \vec{x}}\frac{\partial \vec{x}}{\partial p_i} + \frac{\partial F}{\partial \vec{x}^*}\frac{\partial \vec{x}^*}{\partial p_i} \\ &= 2\text{Re}\left\{\frac{\partial F}{\partial \vec{x}}\frac{\partial \vec{x}}{\partial p_i}\right\}\end{aligned} \tag{2.6}$$

Because the fields in $\vec{x}$ are complex valued, we must be careful when taking their derivatives, hence the appearance of the $2\text{Re}\{\cdots\}$[1].

Notice that the derivative $\partial F/\partial \vec{x}$ is already known since the figure of merit is an explicit function of the electric and magnetic fields. The second term in Eq. (2.6), $\partial \vec{x}/\partial p_i$, remains to be found. We accomplish this by directly differentiating our system of equations given in Equation (2.3) with respect to p_i and left multiplying by A^{-1}, which yields

$$\begin{aligned}&\frac{\partial}{\partial p_i}(A\vec{x}) = \frac{\partial \vec{b}}{\partial p_i} \\ &\Rightarrow A\frac{\partial \vec{x}}{\partial p_i} = \frac{\partial \vec{b}}{\partial p_i} - \frac{\partial A}{\partial p_i}\vec{x} \\ &\Rightarrow \quad \frac{\partial \vec{x}}{\partial p_i} = A^{-1}\left(\frac{\partial \vec{b}}{\partial p_i} - \frac{\partial A}{\partial p_i}\vec{x}\right) \quad .\end{aligned} \tag{2.7}$$

Notice that the Maxwell operator A contains the distribution of permittivity and permeability in the system which is directly controlled by the design parameters $\vec{p}$. Therefore, $\partial A/\partial p_i$ is known or assumed to at least be easily computable. The derivative of the current sources with respect to the design parameters, $\partial \vec{b}/\partial p_i$, can be calculated, although in most cases we will assume that the inputs to the system are fixed and this term will be zero. In this case, Equation (2.7) becomes

$$\frac{\partial \vec{x}}{\partial p_i} = -A^{-1}\frac{\partial A}{\partial p_i}\vec{x} \quad . \tag{2.8}$$

[1]The appearance of the second derivative with respect to the complex conjugate x^* is a consequence of chain rule for Wirtinger derivatives. Because $F(\vec{x})$ and p_i are both real-valued, we see that the second term is the complex conjugate of the first, i.e.,

$$\left(\frac{\partial F}{\partial \vec{x}}\frac{\partial \vec{x}}{\partial p_i}\right)^* = \frac{\partial F}{\partial \vec{x}^*}\frac{\partial \vec{x}^*}{\partial p_i} \quad .$$

This results in the appearance of the $2\text{Re}\{\cdots\}$.

Substituting this expression for $\partial\vec{x}/\partial p_i$ in Equation (2.6), we find an expression for the i'th derivative of F in terms of known quantities:

$$\frac{dF}{dp_i} = -2\text{Re}\left\{\frac{\partial F}{\partial \vec{x}} A^{-1} \frac{\partial A}{\partial p_i}\vec{x}\right\} \tag{2.9}$$

This expression can be written in a more enlightening way by introducing a new vector given by

$$\vec{y}^T = \frac{\partial F}{\partial \vec{x}} A^{-1} \tag{2.10}$$

which we can rewrite in the form "$A\vec{x} = \vec{b}$" by right multiplying by A and taking the transpose of both sides:

$$A^T\vec{y} = \left(\frac{\partial F}{\partial \vec{x}}\right)^T \tag{2.11}$$

Solving this expression for y is *similar* to solving Maxwell's equations where $\partial F/\partial\vec{x}$ acts as the current sources. In general, the discretized Maxwell's equations are not symmetric, and therefore the forward and adjoint equations are not identical. Substituting Equation (2.10) into (2.9), we obtain a final expression for the derivative of our function F with respect to the design variables of the system:

$$\frac{dF}{dp_i} = -2\text{Re}\left\{\vec{y}^T \frac{\partial A}{\partial p_i}\vec{x}\right\} \ . \tag{2.12}$$

In order to solve for all M derivatives of F with respect to every p_i, we need to compute the physical electric and magnetic fields represented by $\vec{x}$ as well as a second set of non-physical "adjoint" fields represented by $\vec{y}$. We can intuitively think about these adjoint fields as the fields that are produced by injecting the desired output fields (which arise from adjoint current sources $\partial F/\partial\vec{x}$) into the system and running the whole system backwards. Solving for $\vec{x}$ and $\vec{y}$ each correspond to a single "forward" and "adjoint" simulation, respectively, and thus solving for the gradient of F requires two simulations, independent of the number of design variables. This is the great advantage of the adjoint method.

In addition to the forward and adjoint simulations, an essential component of the adjoint method is the accurate calculation of $\partial A/\partial p_i$ which contains the information about how the distribution of materials in the system is controlled by the design parameters $\vec{p}$. In particular, the discretized equations which are assembled into A can be ordered such that the permittivity and permeability values are contained in the diagonal blocks of A as indicated by Equation (2.4). Typically, when working on a rectangular grid (as we do in this manuscript), the discretization of the problem will remain unchanged as the design variables are modified. If this is true then the off-diagonal blocks of the system matrix will not change with respect to changes to the design variables and hence the off-diagonal blocks of dA/dp_i

will be zero. This leaves only the diagonal blocks which contain the derivatives of the permittivity and permeability tensors at each point in space. In this case, the derivatives of the figure of merit are greatly simplified to

$$\begin{aligned}\frac{dF}{dp_k} &= -2\text{Re}\left\{i\omega\vec{E}_{\text{adj}}^T\frac{\partial\varepsilon}{\partial p_k}\vec{E} - i\omega\vec{H}_{\text{adj}}^T\frac{\partial\mu}{\partial p_k}\vec{H}\right\} \\ &= 2\omega\ \text{Im}\left\{\vec{E}_{\text{adj}}^T\frac{\partial\varepsilon}{\partial p_k}\vec{E} - \vec{H}_{\text{adj}}^T\frac{\partial\mu}{\partial p_k}\vec{H}\right\}\end{aligned} \tag{2.13}$$

where the field quantities with the subscript “adj" are contained in $\vec{y} = [\vec{E}_{\text{adj}}, \vec{H}_{\text{adj}}]^T$. Based on this, calculating the gradient of our figure of merit using the adjoint method consists of computing forward fields, adjoint fields, and material gradients. Assuming we can calculate these three quantities, the adjoint method gives us a very efficient way to calculate the gradient of our figure of merit with respect to any number of variables.

These gradients are an essential component in both shape and topology optimization. As far as the adjoint method is concerned, the primary distinction between different forms of shape and topology optimization lies in the details of how $\varepsilon(\vec{p})$ and $\mu(\vec{p})$, and hence their derivatives $\partial\varepsilon/\partial p_i$ and $\partial\mu/\partial p_i$, are defined. In this work, we focus on shape optimization. This means that our design variables typically define the boundaries of regions of different ε and μ. The derivatives, meanwhile, relate changes in boundary positions to local changes in the permittivity and permeability. Handling these derivatives in careful way is essential to successful application of the adjoint method.

2.1.1 Including Explicit Dependence on the Design Parameters

Up until now, we have considered only figures of merit which depend exclusively on the electric and magnetic fields. In many cases, however, it will be advantageous to have the figure of merit include an explicit dependence on the design variables. In these cases, our definition of the derivative of our figure of merit will need to change slightly. Specifically, we must modify Equation (2.6) such that it is consistent with the definition of total derivatives:

$$\frac{dF}{dp_i} = 2\text{Re}\left\{\frac{\partial F}{\partial\vec{x}}\frac{\partial\vec{x}}{\partial p_i}\right\} + \frac{\partial F}{\partial p_i} \tag{2.14}$$

Here, the partial derivative $\partial F/\partial p_i$ is the derivative with respect to p_i, holding $\vec{x}$ and $p_{j\neq i}$ constant. Based on this simple modification, our expression for the gradient computed using the adjoint method becomes

$$\frac{dF}{dp_i} = -2\text{Re}\left\{\vec{y}^T\frac{\partial A}{\partial p_i}\vec{x}\right\} + \frac{\partial F}{\partial p_i}\ . \tag{2.15}$$

The addition of an explicit dependence on the design variables to our figure of merit results in an additional additive term to our gradient and does not require modification of the adjoint method itself. In appendix A, we demonstrate how this explicit dependence can be used to easily impose design constraints.

2.1.2 Discretizing and Differentiating Figures of Merit

The goal of gradient-based optimization is to minimize (or maximize) a figure of merit which describes a device's performance by changing the set of design parameters which define the shape of the device. Typically, this figure of merit will be contain a spatial integral of the electric and magnetic field, i.e., it will have the form

$$F(\mathbf{E}, \mathbf{H}) = \int_{\mathcal{D}} d\mathcal{D} f(\mathbf{E}, \mathbf{H}) \tag{2.16}$$

where $\mathcal{D}$ is some 1D, 2D, or 3D domain and $d\mathcal{D}$ is the differential of that domain. Before computing its gradient using the adjoint method, we need to make two modifications to our figure of merit. First, we need to represent it in a discretized domain which is conducive to numerical computation. Second, we need to take its derivative with respect to the field components which serves as the adjoint simulation sources.

In this work, we primarily work with finite difference methods which represent the fields on a rectangular grid. This makes it relatively straightforward to rewrite our integrals as discrete sums. For example, our integral expression in Equation (2.16) can be rewritten on a rectangular grid using a Riemann sum as

$$F(\vec{\mathbf{E}}, \vec{\mathbf{H}}) = \sum_q \Delta\mathcal{D} f(\mathbf{E}_q, \mathbf{H}_q) \tag{2.17}$$

where $\Delta\mathcal{D}$ is the spacing between field values in our grid and q denotes a spatial index. Furthermore, we have switched to vector notation to emphasize that the electric and magnetic fields are discretized on the rectangular grid and are now stored as vectors of discrete values. Depending on the dimensionality, the sum and indexing will change. In particular, in 1D, 2D, and 3D, the sums become:

$$\begin{aligned}
&\text{1D}: && \sum_q \Delta\mathcal{D} && \rightarrow && \sum_k \Delta x && \text{and} && q \rightarrow k \\
&\text{2D}: && && \rightarrow && \sum_j \sum_k \Delta x \Delta y && \text{and} && q \rightarrow jk \\
&\text{3D}: && && \rightarrow && \sum_i \sum_j \sum_k \Delta x \Delta y \Delta z && \text{and} && q \rightarrow ijk
\end{aligned} \tag{2.18}$$

Having written our integral figure of merit as a discrete sum, taking its derivative with respect to the field components is relatively straight forward. Consider, for example, taking the derivative of Equation (2.17) with respect to E_m^x:

$$\begin{aligned}\frac{\partial F}{\partial E_m^x} &= \frac{\partial}{\partial E_m^x}\sum_q \Delta\mathcal{D} f(\mathbf{E}_q, \mathbf{H}_q) \\ &= \sum_q \Delta\mathcal{D}\frac{\partial f(\mathbf{E}_q, \mathbf{H}_q)}{\partial E_m^x} \\ &= \sum_q \Delta\mathcal{D}\frac{\partial f(\mathbf{E}_m, \mathbf{H}_m)}{\partial E_m^x}\delta_{qm} \\ &= \Delta\mathcal{D}\frac{\partial f(\mathbf{E}_m, \mathbf{H}_m)}{\partial E_m^x}\end{aligned} \tag{2.19}$$

The derivative of our figure of merit with respect to the field components is simply the derivative of the integrand times the differential $\Delta\mathcal{D}$. It is important to note that this is only true when m corresponds to an index contained within the sum. For all m not contained in the sum, $\partial F/\partial E_m^x = 0$.

For most figures of merit that we are interested in, the process of calculating them numerically and differentiating them for use with the adjoint method will follow this blue print. Even when working with more sophisticated functions which involve products and quotients of integrated quantities, we can always follow this process of discretizing the integral and then differentiating.

2.1.3 Accounting for Spatial Interpolation

When solving Maxwell's equations using finite difference methods, we typically represent the electric and magnetic fields on a staggered grid. This means that the different field components that we solve for are not known at the same exact point in space. In order to properly calculate quantities like power and energy density which combine multiple field components, it becomes necessary to interpolate the field values such that every field component is known at the same points.

This process of interpolating the fields modifies our figure of merit. As a result, we need to be careful to account for the modification when computing the adjoint sources. As an example, let us consider the case of representing 2D TM fields on the staggered rectangular grid as depicted in Figure 2.1[2]. For this field configuration, we could define a set of linearly spatially-interpolated field values given by

[2]In this depiction, we do not use the half-step index numbering scheme as is frequently used with Yee cells/staggered grids. Instead, we group sets of H_z, E_x, and E_y and assign them the index (i, j). This is purely subjective and should not influence the results of this section.

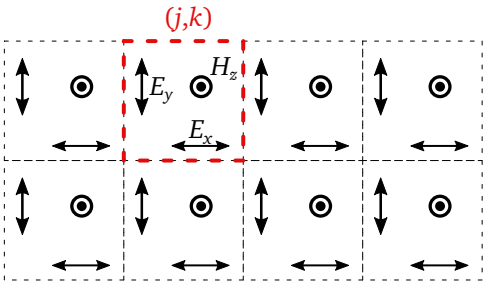

Figure 2.1: Example of a staggered grid used to represent 2D TM fields. Each (Yee) cell contains the three field components. This staggered grid ensures numerical stability.

$$\widetilde{H}^z_{j,k} = H^z_{j,k} \tag{2.20}$$

$$\widetilde{E}^x_{j,k} = \frac{1}{2}\left(E^x_{j,k} + E^x_{j+1,k}\right) \tag{2.21}$$

$$\widetilde{E}^y_{j,k} = \frac{1}{2}\left(E^y_{j,k} + E^x_{j,k+1}\right) \tag{2.22}$$

where $\widetilde{\square}$ refers to interpolated quantities, and the electric field components, $E^x_{j,k}$ and $E^y_{j,k}$, are interpolated at the position of $H^z_{j,k}$. In terms of these interpolated quantities, our figure of merit given in Equation (2.17) becomes

$$F(\vec{\widetilde{\mathbf{E}}}, \vec{\widetilde{\mathbf{H}}}) = \sum_{\mathcal{D}} \Delta\mathcal{D} f(\widetilde{\mathbf{E}}_q, \widetilde{\mathbf{H}}_q) \quad . \tag{2.23}$$

In order to apply the adjoint method to this *new* figure of merit, we need to calculate its derivative with respect to the *uninterpolated* field components (since $\vec{x}$ contains the uninterpolated values). In order to calculate these derivatives, we proceed exactly as we did in order to arrive at Equation (2.19). Let's consider first the simpler derivative with respect to $H^z_{m,n}$ which is given by

$$\frac{\partial F}{\partial H^z_{m,n}} = \sum_j \sum_k \Delta x \Delta y \frac{\partial f(\widetilde{\mathbf{E}}_{j,k}, \widetilde{\mathbf{H}}_{j,k})}{\partial H^z_{m,n}} \quad . \tag{2.24}$$

Notice that in this case, the figure of merit is written explicitly in terms of the interpolated fields. As such, we would like to be able to express the derivative of F in terms of the derivative of f with respect to the interpolated fields. We accomplish this by applying chain rule:

$$\frac{\partial F}{\partial H^z_{m,n}} = \sum_j \sum_k \Delta x \Delta y \frac{\partial f(\widetilde{\mathbf{E}}_{j,k}, \widetilde{\mathbf{H}}_{j,k})}{\partial \widetilde{H}^z_{j,k}} \frac{\partial \widetilde{H}^z_{j,k}}{\partial H^z_{n,m}} \tag{2.25}$$

Recalling that $\widetilde{H}^z_{j,k} = H^z_{j,k}$, we find that $\partial \widetilde{H}^z_{j,k}/\partial H^z_{m,n} = \delta_{jm}\delta_{kn}$ and hence

$$\frac{\partial F}{\partial H^z_{j,k}} = \Delta x \Delta y \frac{\partial f(\widetilde{\mathbf{E}}_{j,k}, \widetilde{\mathbf{H}}_{j,k})}{\partial \widetilde{H}^z_{j,k}} \tag{2.26}$$

From this we see that the derivative of our figure of merit with respect to $H^z_{j,k}$ is identical to the "naive" derivative with respect to the interpolated field.

This is not the case, on the other hand, for the derivative of F with respect to $E^x_{j,k}$ and $E^y_{j,k}$. Let us consider the former derivative. Just as we did in Equation (2.25), we differentiate with respect to $E^x_{m,n}$ and apply chain which yields

$$\frac{\partial F}{\partial E^x_{m,n}} = \sum_j \sum_k \Delta x \Delta y \frac{\partial f(\widetilde{\mathbf{E}}_{j,k}, \widetilde{\mathbf{H}}_{j,k})}{\partial \widetilde{E}^x_{j,k}} \frac{\partial \widetilde{E}^x_{j,k}}{\partial E^x_{m,n}} \quad . \tag{2.27}$$

In this case, $\widetilde{E}^x_{j,k} \neq E^x_{j,k}$ and so we must proceed cautiously. Consider the derivative of the interpolated $\widetilde{E}^x_{j,k}$ with respect to the uninterpolated $E^x_{m,n}$:

$$\begin{aligned}
\frac{\partial \widetilde{E}^x_{j,k}}{\partial E^x_{m,n}} &= \frac{\partial}{\partial E^x_{m,n}} \frac{1}{2} \left(E^x_{j,k} + E^x_{j+1,k} \right) \\
&= \frac{1}{2} \left(\frac{\partial E^x_{j,k}}{\partial E^x_{m,n}} + \frac{\partial E^x_{j+1,k}}{\partial E^x_{m,n}} \right) \\
&= \frac{1}{2} \left(\delta_{j,m}\delta_{k,n} + \delta_{j+1,m}\delta_{k,n} \right)
\end{aligned} \tag{2.28}$$

Substituting this result into Equation (2.27), yields for the derivative with respect to $E^x_{j,k}$

$$\frac{\partial F}{\partial E^x_{j,k}} = \Delta x \Delta y \frac{1}{2} \left(\frac{\partial f(\widetilde{\mathbf{E}}_{j,k}, \widetilde{\mathbf{H}}_{j,k})}{\partial \widetilde{E}^x_{j,k}} + \frac{\partial f(\widetilde{\mathbf{E}}_{j,k}, \widetilde{\mathbf{H}}_{j,k})}{\partial \widetilde{E}^x_{j-1,k}} \right) \tag{2.29}$$

where we have sneakily selected out the terms from the sum which correspond to the Kronecker delta functions in Equation (2.28). Notice that this expression is really just the average of two derivatives of the full figure of merit with respect to the *interpolated* fields, i.e.,

$$\frac{\partial F}{\partial E^x_{j,k}} = \frac{1}{2} \left(\frac{\partial F}{\partial \widetilde{E}^x_{j,k}} + \frac{\partial F}{\partial \widetilde{E}^x_{j-1,k}} \right) \quad . \tag{2.30}$$

From this, it follow that in order to calculate the derivative of our figure of merit with respect to the uninterpolated fields (which is the adjoint source), we can simply calculate the "naive" derivative with respect to the interpolated fields and then

average the spatially-adjacent derivatives! This provides us with an easy way to account for spatial interpolation of the fields in our adjoint source calculation.

This holds true for the derivative with respect to $E^y_{j,k}$,

$$\frac{\partial F}{\partial E^y_{j,k}} = \frac{1}{2}\left(\frac{\partial F}{\partial \widetilde{E}^y_{j,k}} + \frac{\partial F}{\partial \widetilde{E}^y_{j,k-1}}\right) \quad . \tag{2.31}$$

and a similar process can be followed for figures of merit which depend on 2D TE fields and 3D fields. In this sense, these results are very general and should be used when calculating adjoint sources to ensure that the resulting gradients account for the numerical details of the simulation.

2.1.4 Keys to Computing Accurate Gradients

Based on Equations (2.13), it is apparent that if we are able to calculate the forward fields, the adjoint fields, and the material gradients, then we should be able to efficiently calculate the gradient of our figure of merit. There are, however, a number of numerical pitfalls associated with implementing the adjoint method that we must be aware of. Not accounting for these details can negatively impact the accuracy of our gradients, and as a result, significantly impact the optimization process.

First, we must take care to ensure that all calculations (forward simulation, adjoint simulation, and material gradients) are consistent with one another. Any post processing of the fields generated by the forward simulation must be accounted for in the adjoint simulation. The most common example of this is (spatial or spectral) field interpolation: if our figure of merit is computed using interpolated field values (which is necessary when using finite-difference methods), it is important that the derivative of our figure of merit $\partial F/\partial \vec{x}$ account for the interpolation. Since this derivative serves as the source to the adjoint simulation, inconsistencies in the definition of $\vec{x}$ will lead to errors in the adjoint field.

In addition to field interpolation, it is important that our material gradients are computed in a way that is consistent with how the forward simulator handles its material distributions. Although this may seem trivial, guaranteeing this consistency can prove difficult when working with third party solvers which are effectively black boxes.

The material gradients pose an additional challenge. When working with finite difference methods (FDTD or FDFD) as is common in the nanophotonics community, simulations are constrained to represent data on a rectangular grid. This rectangular grid is typically not conducive to making smooth modifications to the material boundaries. Without this smooth behavior, the material distributions are not differentiable and hence $\partial \epsilon_j/\partial p_k$ and $\partial \mu_j/\partial p_k$ are ill-defined. In such cases, these derivatives can be approximated [10], however, such approximations may be prone to producing gradient errors in some applications.

Finally, the process of computing adjoint fields itself poses some challenges. Mathematically, the adjoint simulation, $A^T y = (\partial F/\partial \vec{x})^T$, consists of solving a transposed version Maxwell's equations. If we solve Maxwell's equations using an explicit matrix method like the finite difference frequency domain method or finite element method, the system matrix A can be transposed, and therefore running the adjoint simulation is straightforward. If, however, we prefer to use an alternative matrix-free method like the finite difference time domain method (which is typically faster and more scalable for dielectric structures), the adjoint simulation is more difficult to implement. In this case, we have two options: either we can work directly in the time domain as in [19, 24], requiring slight modification to the underlying equations, or we can work with the frequency domain fields (acquired, for example, using a Fourier transform) and assume that the forward and adjoint simulations solve an identical set of equations (with different sources). In both cases, we typically assume that the systems of equations that we solve are *symmetric*, which is not strictly true when working with non-diagonally anisotropic materials, non-uniform grids, perfectly matched layers (PML), and non-absorbing boundary conditions. In order to ensure that the gradients computed are accurate, we therefore either need to correctly account for anisotropies, non-uniform grids, and boundary conditions in our adjoint simulation (which may require modification to the FDTD equations) or avoid them altogether. In practice, perfectly matched layers do not significantly diminish gradient accuracy as long as all relevant calculations are restricted to domains which do not overlap the PML domains.

As long as we account for the nuances of the adjoint method, we can efficiently calculate gradients with high accuracy. In the following sections, we will explore some of the more important caveats we have just introduced, and in section 2.3 we will demonstrate the impact of neglecting these details.

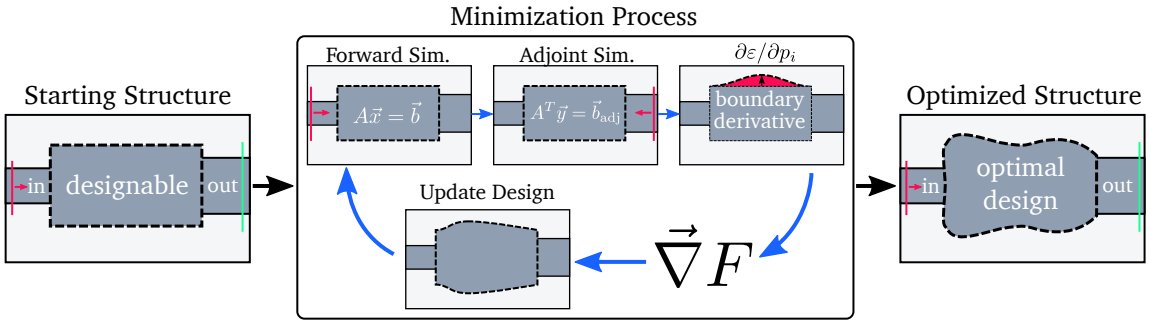

Figure 2.2: Overview of how the adjoint method fits into shape optimization.

2.1.5 Using the Adjoint Method for Shape Optimization

The adjoint method serves a single purpose: to calculate gradients of a function which depends implicitly on a larger set of variables. On their own, these gradients, which are also aptly referred to as “device sensitivities," are a useful measure of how sensitive a device’s performance is to changes in its anatomy. Their true power, however, is unlocked when supplied to a gradient-based minimization method. These methods use the gradient information in order to rapidly search for a minimum of the function (i.e., out figure of merit), which corresponds to an optimal device. Many such methods exist; in this work, we primarily employ the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm which works reliably in a wide range of problems.

The general process for using the adjoint method in shape optimization is shown in Figure 2.2. We begin with some initial design for our device and define design parameters (i.e., a list of numbers) which control the shape of the initial device’s boundaries. The goal of the optimization is to alter these design parameters in order to minimize our figure of merit. This process is iterative. First, we run a forward simulation of the device in order to calculate the current value of our figure of merit. The fields obtained from this forward simulation are then used to calculate the adjoint simulation sources. We then run the adjoint simulation to obtain the adjoint fields. After both sets of fields are known, we compute the derivative of the material distributions (typically just the permittivity distribution) with respect to each of the design variables. In many cases, the process for calculating these material derivatives involves perturbing the boundary in different ways and calculating how the permittivity distribution changes. Multiplying the adjoint field, forward field, and each of the material derivatives yields the gradient of our figure of merit with respect to the design variables. This gradient is then used by the minimization algorithm to update the design variables, yielding an improved device. The gradient-fueled update process is then repeated (potentially tens or hundreds of times) until we determine that the optimization has converged. The set of design parameters that achieve this minimum describes the optimized device geometry.

2.2 Shape Optimization Based on Boundary Smoothing

One of the key steps in computing device sensitivities is to calculate the material gradients. This becomes a challenge when combining shape optimization with finite difference methods which operate on rectangular grids. When trying to represent an abrupt interface between two materials, we inevitably run into the problem of staircasing: representing curved or diagonal boundaries between two different materials on a rectangular grid results in a jagged interface which conforms to the

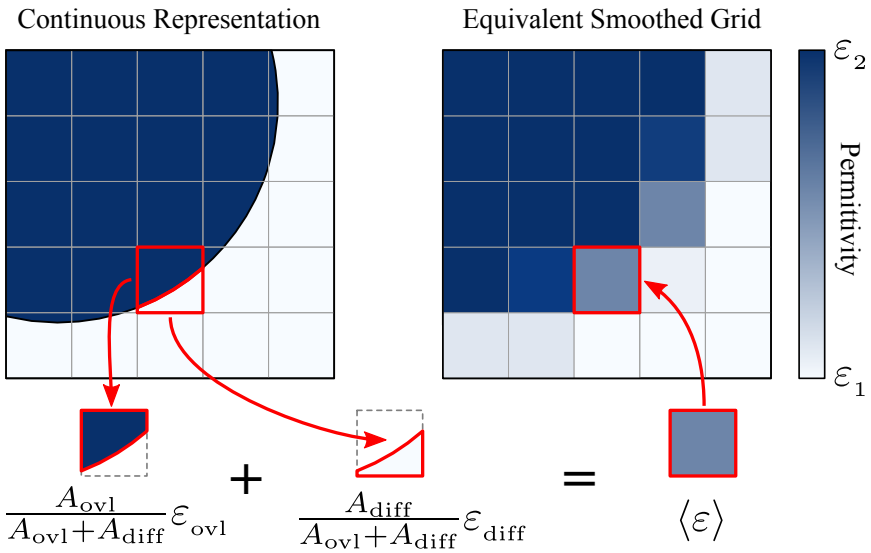

Figure 2.3: Visual depiction of boundary smoothing process. Internally, all material boundaries of the system are represented using polygons which are defined in a continuous domain. These shapes are then mapped onto a rectangular grid by computing the average value of permittivities and permeabilities which overlap with each cell in the grid. This mapping is achieved by computing the overlap area between grid cells and material domains.

underlying grid. In addition to compromising simulation accuracy, the rectangular nature of the grid poses significant challenges to calculating sensitivities since perturbations smaller than a grid cell are not possible.

A considerable amount of work has been done to improve the treatment of non-rectangular boundaries with finite difference methods for the purpose of improving simulation accuracy. In particular, modification of the FDTD equations have been successfully employed in order to improve the simulation accuracy [25, 26] and the introduction of effective permittivity at material interfaces [27–30] has been demonstrated to improve simulation accuracy in many situations. In many of these works, the process of computing effective intermediate material values, which we refer to as "boundary smoothing," relies on the calculation of a volume-averaged permittivity/permeability. Despite the central role this averaging plays, comparatively little work has been done on computing these averages in a way that produces effective material values that change smoothly with respect to underlying material boundaries. If this smooth behavior is achieved, this method of smoothing material interfaces becomes a powerful tool for shape optimization.

In order to demonstrate precisely this, we have implemented a simple form of boundary smoothing using weighted averages which is depicted in Figure 2.3. In a given grid cell, the effective permittivity in a 2D domain is given by

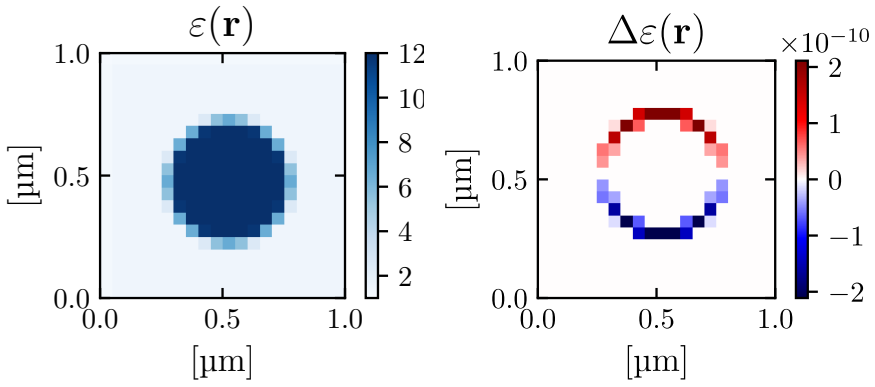

Figure 2.4: Demonstration of boundary smoothing for a 0.5 µm diameter dielectric circle. (a) shows the relative permittivity of the smoothed grid computed for the circle on an intentionally coarse grid in order to clearly show the averaging which occurs at the circle's boundary. (b) shows the difference in permittivity between the grid shown in (a) and the grid corresponding to the same circle which has been shifted in the y direction by 10^{-12} µm. The difference in permittivity is correspondingly small, highlighting the continuous nature of our boundary smoothing.

$$\langle \varepsilon(i,j) \rangle = \frac{1}{\Delta x \Delta y} \sum_k C_k \varepsilon_k(i,j) \tag{2.32}$$

where C_k is the overlap area between the k'th material domain and the grid cell at location i, j and Δx and Δy are the grid cell width and height, respectively. For the purpose of computing derivatives, it is essential that C_k be computable with high precision. We accomplish this by representing all boundaries in the system as piecewise linear functions (i.e., polygons) and then computing intersections between the material domains and the grid-cells that intersect the boundaries of those domains. Because these piecewise linear functions are stored with very high (or even arbitrary) numerical precision, infinitesimal modifications to the boundaries are reflected by infinitesimal modifications to the local effective permittivity and permeabilities on the grid. This process is contingent on our ability to efficiently find polygon intersections. Fortunately, this has long been a topic of great importance in computational geometry [31]. Due to the maturity of this field, efficient algorithms for finding the intersection between polygons are readily available, making this simple form of boundary smoothing relatively straightforward to implement.

A demonstration of this process is depicted in Figure 2.4. In Figure 2.4(a), the smoothed permittivity for a 0.5 µm diameter circle is shown on an intentionally coarse grid. As a result of the smoothing process, the grid cells at the boundary of the circle are filled with an effective permittivity whose value is between the

permittivity inside of the circle and the permittivity surrounding the circle. We then test the continuous nature of this smoothing by displacing the circle in the y direction by $10^{-12}\ \mu\mathrm{m}$. The change in the permittivity as a result of this displacement is shown in Figure 2.4(b). The small size of the displacement is reflected by the correspondingly small change in permittivity in the grid cells at the circle's outer boundary.

Figure 2.4 demonstrates that we can make infinitesimal changes to boundaries represented on a rectangular grid. These infinitesimal changes translate into the smooth changes in quantities calculated using the simulated fields in response to changes to the simulated structure. This behavior is demonstrated in greater detail in Figure 2.5 which plots the scattering cross section of an infinite dielectric cylinder as a function of the radius of the cylinder. In both plots, the cross section is calculated using FDTD with a grid spacing of $30\,\mathrm{nm}$. The cylinder in these simulations is represented using our boundary smoothing which is compared to a strictly binary Manhattan representation of the cylinder as well as to the theoretical value for the scattering cross section. Notice that in the left plot, the cross section of the smoothed cylinder varies smoothly over the full range of radii which spans two grid cells, closely matching the theoretical behavior. The Manhattan representation, meanwhile, varies in a jagged way with changing radius. This behavior is made more apparent by zooming in on a small range of radii (which spans much less than one grid cell) as shown in the left plot of Figure 2.5. Unlike the Manhattan representation which decreases in discontinuous steps, the cross section computed using our smoothed representation decreases smoothly, closely matching the theoretical cross section. This behavior is essential to the calculation of accurate gradients of a figure of merit.

With continuous boundary smoothing at our disposal, the derivatives $\partial A/\partial p_i$ are easily computed using finite differences. One at a time, each design variable of the system p_i is perturbed by a small amount (e.g., $10^{-8} \times \Delta x$) and the diagonals of a new matrix $A(p_i + \Delta p)$ is computed. The derivative is then given approximately by

$$\frac{\partial A}{\partial p_i} \approx \frac{A(p_i + \Delta p) - A(p_i)}{\Delta p} \tag{2.33}$$

which is accurate so long as Δp is sufficiently small, regardless of how coarse or fine the spatial discretization of the underlying simulation is. Note that although it is possible to compute this derivative (semi-) analytically without relying on finite differences, it turns out that approximating the derivative using finite differences is no more computationally intensive and has practical implementation benefits[3]

[3]In order to calculate the derivative analytically, we would apply chain rule:

$$\frac{dA}{dp_i} = \frac{\partial A}{\partial C_k}\frac{\partial C_k}{\partial x}\frac{\partial x}{\partial p_i} + \frac{\partial A}{\partial C_k}\frac{\partial C_k}{\partial y}\frac{\partial y}{\partial p_i}$$

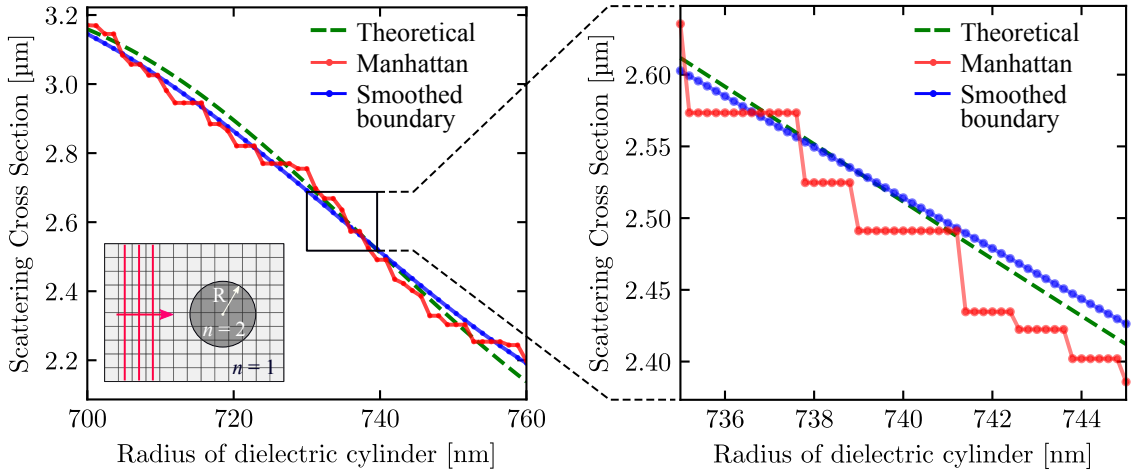

Figure 2.5: Demonstration of boundary smoothing applied to the calculation of the scattering cross section of an infinite dielectric cylinder (in 2D). The scattering cross sections computed using boundary smoothing (blue curve) and without using boundary smoothing (red curve) are compared to the theoretical value for a range of cylinder radii. On the left, the cylinder radius is varied by an amount equal to twice the grid spacing (30 nm). On the right, the radius is varied over a range much smaller than a single grid step. In both cases the cross-section of the grid-smoothed-cylinder evolves smoothly with changing radius while cross section of the cylinder represented on a strictly binary grid exhibits a step-like behavior. The smooth change in the cross section of the boundary-smoothed-cylinder is highly desirable when calculating sensitivities.

It is interesting to note that our boundary smoothing method is conceptually similar to a variety of different "fictitious domain methods" which are regularly employed in the field of structural engineering [32]. In particular, the volume-based material averaging we use is analogous to the "ersatz approach" to the level set method employed by Allaire et al [4] which has since been applied to other topology optimization methods (e.g., the method of moving morphable components [33]). However, in comparison to this level set approach, our method is in many ways easier to grasp conceptually and its implementation is comparatively straightforward. Furthermore, by representing the material values explicitly as polygons (which is supported by most if not all layout tools), translating our designs from a simulated to fabricated structure requires no addition effort. An additional consequence of this explicit polygon representation is the ease with which fabrication constraints may be incorporated into the optimization as we demonstrate in later chapters and discuss in detail in Appendix A. Finally, our boundary smoothing technique provides us with great flexibility in how we manipulate structures: it works equally well for

where C_k is the volume fraction, x is the x coordinate of each vertex in the polygon and y is the y coordinate of each vertex in the polygon. In many cases, $\partial x/\partial p_i$ and $\partial y/\partial p_i$ may be tedious to implement. Computing the derivatives using finite differences avoids this issue all together.

geometries which are heavily constrained (e.g., grating couplers which consist of an arrangement of different sized rectangles) as it does for geometries which evolve in a more "free form" manner (as is typically observed with level set methods). This combination of simplicity and flexibility makes our method an attractive alternative to other fictitious domain methods for electromagnetic optimization.

For these reasons, boundary smoothing provides a very flexible foundation for shape optimization. In the remainder of this work, we will demonstrate how shape optimization based on boundary smoothing can be used to optimize a wide range of devices in nanophotonics.

2.3 A Simple Example

In order to demonstrate the application of the adjoint method and highlight some of the key details discussed in this chapter, let us consider the problem of optimizing the simple 90° waveguide bend depicted in Figure 2.6. In this optimization, we excite the fundamental mode of a horizontal silicon waveguide and attempt to maximize the power which remains in the wavguide after passing through the bend by varying the inner and outer radii of the bend. In order to choose these radii, we will compute the gradient of our figure of merit (the output power) using the adjoint method, and then supply that gradient to the BFGS minimization algorithm. This example is intentionally simplistic as it allows us to work through the full optimization problem without having to deal with excessive and opaque math. In chapters 4 and 5, we will apply these same methods to more interesting 3D problems.

For the purpose of this example, we approximate a three dimensional silicon strip waveguide in two dimensions using the effective index method. Under this approximation, for a wavelength of $1550\,\mathrm{nm}$, the refractive index of the waveguide is set to 2.8 (corresponding to an approximate waveguide thickness of $220\,\mathrm{nm}$) and the the cladding index is set to 1.44 (corresponding to silicon dioxide). At the input and output, the waveguide has a fixed width of $500\,\mathrm{nm}$, however this width may change in the bend region depending on the inner and outer bend radii. The $500\,\mathrm{nm}$ input waveguide is excited with its fundamental TM mode with a wavelength of $1550\,\mathrm{nm}$. In order to achieve reasonable accuracy, a grid spacing of $30\,\mathrm{nm}$ is used during optimization. For the purpose of exploring the accuracy of our gradients, we will also consider other grid resolutions.

Our first step in optimizing the waveguide bend is to define the figure of merit that will guide us to the optimal design. In this case, we want to maximize the power which remains in the waveguide after the bend. A simple function which captures this is the power flux through a cross section of the waveguide, which is given by

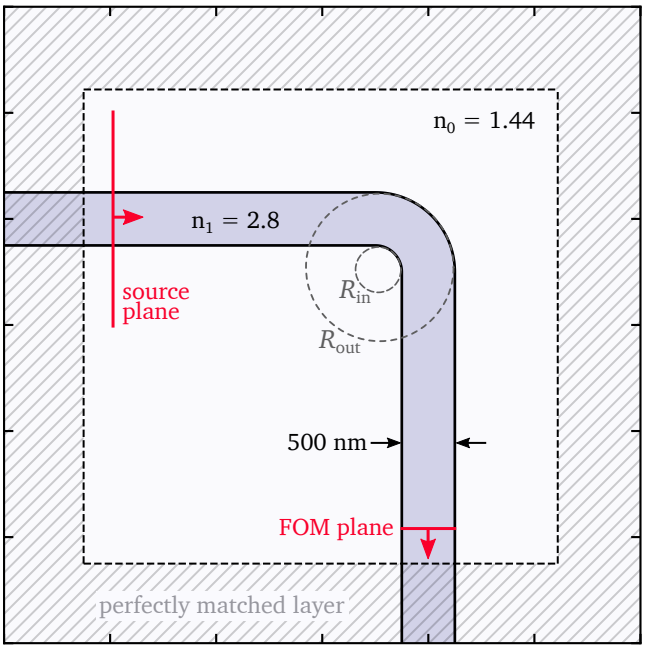

Figure 2.6: Problem setup for a simple demonstration of gradient-based optimization of a 90° silicon waveguide bend. The fundamental TM mode of a silicon waveguide (reduced to 2D using the effective index method) is injected in the top left of the simulation region (labeled “source plane"). The waveguide is bent 90° and the power leaving in the waveguide at the bottom of the simulation region is measured. The inner and outer radius of bend are chosen by the optimization in order to maximize the amount of power flowing through the “FOM plane."

$$\begin{aligned} F(\mathbf{E},\mathbf{H}) &= \frac{1}{P_{\text{src}}} \int_{\ell} d\ell(-\hat{y}) \cdot \frac{1}{2}\text{Re}\{\mathbf{E}\times\mathbf{H}^*\} \\ &= -\frac{1}{P_{\text{src}}} \int_{\ell} d\ell \, \frac{1}{2}\text{Re}\{E_x \times H_z^*\} \end{aligned} \tag{2.34}$$

where P_{src} is the total optical power injected into the simulation and ℓ is the line labeled “FOM plane" in Figure 2.6. Since we are simulating the TM fields, only the E_x, E_y, and H_z are relevant, and hence our figure of merit contains only E_x and H_z which are relevant to power flowing in the $\hat{y}$ direction.

In our figure of merit, we normalize the power flux with respect to the “source power." For the purposes of this example, we will calculate this source power in the first simulation and assume it remains constant for the rest of the optimization. In reality, due to the finite extent of the source, the source power can fluctuate by small amounts from one simulation to the next as the structure and fields change. If we

want to normalize with respect to the source power calculated for each simulation (which will yield the most accurate results), it is important that we account for this normalization in the derivatives of our figure of merit. These derivatives are discussed in detail in Appendix A.

Although we have defined our figure of merit, we still need to rewrite it in terms of the discretized fields. As described in the previous section, we can write our figure of merit as a function of the discretized fields by converting the integral to a sum over index field values:

$$F(\vec{E}_x, \vec{E}_y, \vec{H}_z) = -\frac{1}{P_{\text{src}}} \sum_{k=k_0}^{k_f} \Delta x \, \frac{1}{2} \text{Re}\{\widetilde{E}^x_{j0,k} \times \widetilde{H}^{z*}_{j_0,k}\} \tag{2.35}$$

Here, the vector notation indicates that the fields are stored as vectors of discrete values, $\widetilde{E}$ and $\widetilde{H}$ are the spatially-interpolated fields given in Equation (2.20) and Equation (2.21), k_0 to k_f is the range of indices along the x direction of the sum, and j_0 is the y index of the line where the function is evaluated.

Having expressed our figure of merit in terms of the discretized fields, we can now calculate the adjoint source $\partial F / \partial \vec{x}$. In this case, because we are working with the TM fields, the derivative of our figure of merit with respect to the field components will have the form

$$\frac{\partial F}{\partial \vec{x}} = \left[\frac{\partial F}{\partial \vec{E}_x}, \frac{\partial F}{\partial \vec{E}_y}, \frac{\partial F}{\partial \vec{H}_z} \right] \tag{2.36}$$

The derivatives with respect to the individual field components are relatively straightforward to derive. In particular, we find that the derivatives with respect to the interpolated field components are (see Appendix A.1.1 for details)

$$\begin{aligned} \frac{\partial F}{\partial \widetilde{E}^x_{j,k}} &= -\frac{1}{4P_{\text{src}}} \Delta x \; \widetilde{H}^{z*}_{j,k} \; \delta_{j,j0} \delta_{k \in [k_0, k_f]} \\ \frac{\partial F}{\partial \widetilde{E}^y_{j,k}} &= 0 \\ \frac{\partial F}{\partial \widetilde{H}^z_{j,k}} &= -\frac{1}{4P_{\text{src}}} \Delta x \; \widetilde{E}^{x*}_{j,k} \; \delta_{j,j0} \delta_{k \in [k_0, k_f]} \end{aligned} \tag{2.37}$$

where the Kronecker deltas indicate that the derivatives are only non-zero for the indices contained within the “FOM plane." These derivatives are with respect to the interpolated fields, however we need the derivatives with respect to the uninterpolated fields (i.e. the field values stored in $\vec{E}_x$, $\vec{E}_y$, and $\vec{H}_z$). These derivatives are easily acquired by noting that under our chosen interpolation scheme, $\widetilde{H}^z_{j,k} = H^z_{j,k}$ and the derivative with respect to $E^x_{j,k}$ can be obtained by applying Equation (2.30).

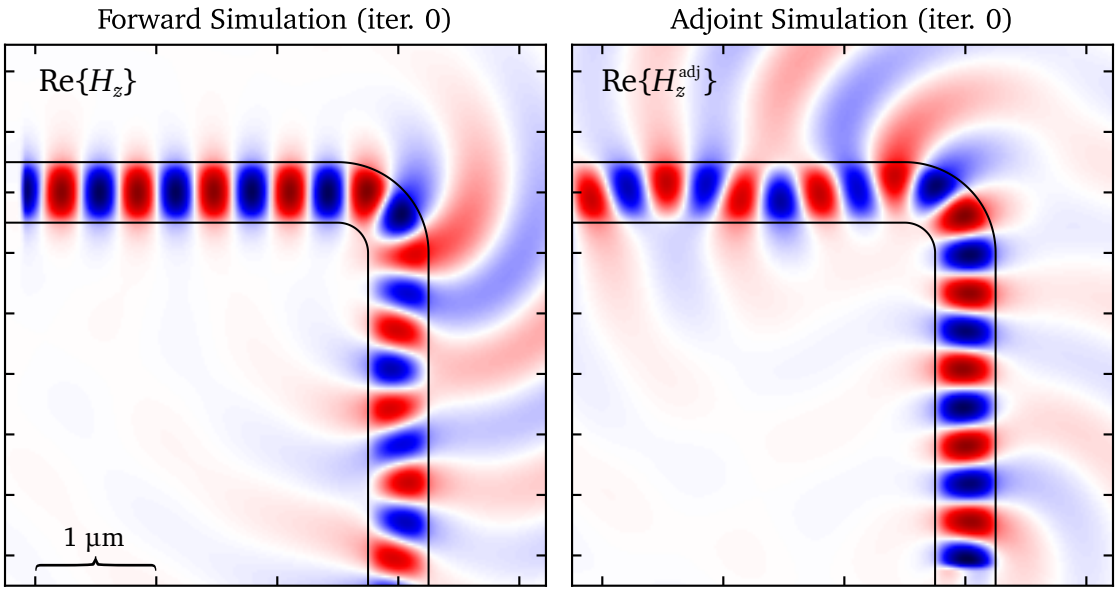

Figure 2.7: Plots of the forward (left) and adjoint (right) fields for the initial waveguide bend. In both cases the $\mathrm{Re}\{H_z\}$ component is plotted and overlaid with the outline of the bent waveguide. In the case of the forward simulation, the fields are injected in the top left and propagate "clockwise". In the case of the adjoint simulation, the fields are injected in the bottom right and propagate "counter-clockwise."

These derivatives define the source used for the adjoint simulation as discussed in section 2.1. As such, we are now in a position to simulate both the forward and adjoint fields. For this two dimensional example, we simulate the structure using the finite difference frequency domain (FDFD) method[4] which sets up and solves Maxwell's equations explicitly in the form $A\vec{x} = \vec{b}$. This makes it straightforward to calculate the adjoint fields since A is explicitly defined.

The simulated forward and adjoint fields are shown in Figure 2.7. In the case of the forward fields, light is injected in the top left and a significant amount of power is scattered at the $90°$ bend as indicated by the easily visible field propagating outside of the waveguide towards the right boundary of the simulation. In the case of the adjoint simulation, it appears as if the light is injected at the output waveguide towards the bend. Just as with the forward simulation, the adjoint fields are scattered by the bend. This example gives us an intuitive sense of what the adjoint fields mean: approximately, they are the fields that result from injecting the desired fields at the output of the system, and running the whole system backwards. This analogy works well when the figure of merit is a relatively simple quantity like power flux or energy density.

The forward and adjoint fields shown in Figure 2.7 are two of the three com-

[4]The solvers we employ for both 2D and 3D problems are discussed in detail in Chapter 3.

ponents in Equation (2.13) needed to calculate the gradient of our figure of merit using the adjoint method. The remaining component is the material gradient $\partial\vec{\epsilon}/\partial\vec{p}$. To ensure that the waveguide permittivity distribution changes smoothly with respect to changes in the bend radii, we represent the waveguide using a polygon with 200 points in the inner and outer arcs and generate the permittivity distribution using our boundary smoothing methods discussed in section 2.2. Any modifications we make to the underlying polygon (like changes to the bend radii) will be reflected in continuous changes to material distribution, allowing us to approximate the material gradients using Equation (2.33).

At this point, we are able to compute the gradient of our figure of merit,

$$\nabla F = \left[\frac{\partial F}{\partial R_{\text{in}}}, \frac{\partial F}{\partial R_{\text{out}}}\right] \tag{2.38}$$

which in this problem is the derivative with respect to the inner and outer bend radii. In order to evaluate the accuracy of the gradient computed using the adjoint method, we compare it to the gradient computed using a brute force finite difference, i.e.,

$$\begin{aligned}\frac{\partial F}{\partial R_{\text{in}}} &\approx \frac{F(R_{\text{in}} + \Delta R, R_{\text{out}}) - F(R_{\text{in}}, R_{\text{out}})}{\Delta R} \\ \frac{\partial F}{\partial R_{\text{out}}} &\approx \frac{F(R_{\text{in}}, R_{\text{out}} + \Delta R) - F(R_{\text{in}}, R_{\text{out}})}{\Delta R}\end{aligned}$$

and estimate the error in our calculation with

$$\text{error} = \frac{|\nabla_{\text{AM}} F - \nabla_{\text{FD}} F|}{|\nabla_{\text{FD}} F|} \tag{2.39}$$

where AM and FD refer to "adjoint method" and "finite differences". Figure 2.8 (a) shows the error in the gradient of our figure of merit as a function of the step size Δp used in the calculation of $\partial A/\partial p_i$ for a range of different grid resolutions. As desired, for small Δp, the gradient calculated using the adjoint method is exceptionally accurate, with error dropping below 10^{-6}, at which point numerical error begins to dominate (as reflected by the noise that appears for very small Δp). This demonstrates that our methods are capable of computing gradients both efficiently and with high accuracy.

In this simple example, there is little apparent correlation between gradient accuracy and grid resolution. We believe this to be due to the fact that changes in the design variables result in changes to the structure which are significantly larger than a single grid cell. In cases in which changes to the design parameter result in changes in only one or a few grid cells, we have observed [34] that the minimum error in the gradient has a polynomial dependence on the grid cell size.

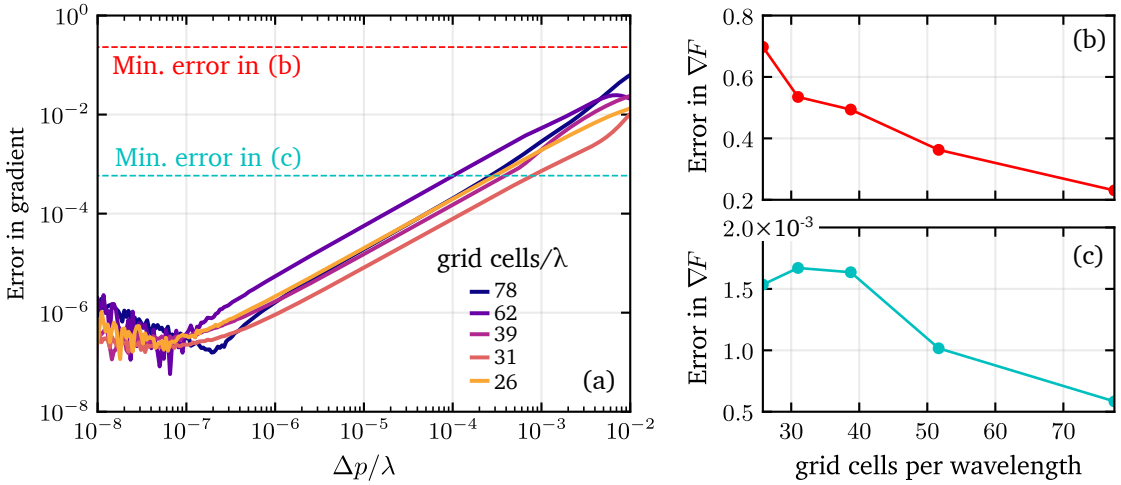

Figure 2.8: Plots of the error in the gradient computed using the adjoint method for the example waveguide bend. (a) Plot of the error in the gradient as a function of the step size Δp used in the calculation of $\partial A/\partial p_i$. (b) Plot of the error in the gradient vs grid resolution when spatial interpolation of the field is not accounted for in the adjoint source. (c) Plot of the error in the gradient vs grid resolution when non-constant source power is not accounted for in the adjoint source.

This simple example gives us an opportunity to explore some of the pitfalls of the adjoint method that we discussed earlier in this chapter. In particular, it is interesting to evaluate how much of an impact spatial interpolation has on gradient accuracy. Figure 2.8 (b) shows the error in the gradient as function of grid resolution when we do not account for spatial interpolation in the definition of the adjoint source (in other words, we compute the adjoint source assuming the interpolated and uninterpolated fields are the same). The error in the gradient as a result of this mistake is increased by at least five orders of magnitude as shown by the horizontal dashed line in Figure 2.8 (a). As the grid spacing is decreased (more grid cells per wavelength), this error steadily drops. This makes sense since the distinction between interpolated and uninterpolated field disappears as we tend towards a continuous (non-discretized) system. Unfortunately, in order to achieve a sufficiently accurate gradient for the purpose of optimization, we would need to use an unreasonably high resolution grid. As such, accounting for the exact form of spatial interpolation used by the Maxwell solver is essential to gradient accuracy.

A second potential source of error concerns the source power normalization. In this simple example, we used the same constant value for the source power throughout our optimization. In reality, the source power in the simulation changes by a small amount as the structure itself changes. When optimizing very high efficiency components, it is important that we normalize our figure of merit with respect to the source power computed for each separate simulation. This introduces an ex-

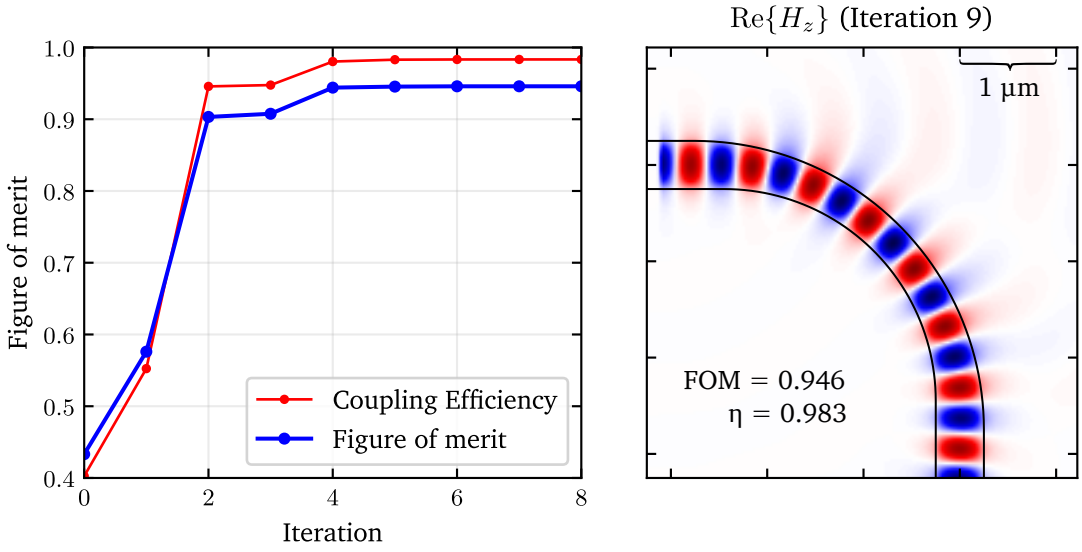

Figure 2.9: (left) The plot of the figure of merit (blue line) and coupling efficiency (red curve) of the bent waveguide as a function of iteration of the optimization. (right) Plot of the real part of the magnetic field corresponding to the optimized bent waveguide. The figure of merit for this optimized bend is 94.6% while the coupling efficiency is 98.3%.

plicit dependence of the source power on the electric and magnetic fields of the system, and therefore we must account for that dependence when differentiating our figure of merit and evaluating the adjoint source; not doing so will introduce unwanted errors in our gradients. Figure 2.8 (c) shows the error in the gradient as a function of grid resolution in the case where we normalize our figure of merit with respect to the actual source power of each simulation, but assume the source power is constant when computing the adjoint source. As shown by the dashed light blue line in Figure 2.8 (a), this increases the minimum error by at least two orders of magnitude. In some problems, we have observed this error to be even higher to the point that it impacts the optimization. For this reason, it is important that we know exactly how the source power is computed by our Maxwell solver and account for it in the derivative of our figure of merit with respect to the fields.

With these different sources of error in mind, we have verified that the gradients of our figure of merit of our simple waveguide bend problem are accurate. All that remains is to optimize the waveguide bend. To do this, we feed the computed gradient into a minimization algorithm. In particular, we use an off-the-shelf BFGS implementation which yields consistently fast convergence. Recall that our goal in this optimization is to *maximize* the transmitted power through our waveguide bend. We accomplish this by minimizing the negative of Equation (2.35).

The improvement in the figure of merit during the optimization is shown on the left side of Figure 2.9. Because the waveguide bend is defined using only two

design variables, convergence to the final value of the figure of merit is very rapid, requiring only nine iterations (which takes about 20 seconds on a 14 core Intel Xeon processor). In this optimization, the figure of merit improves from 0.433 to 0.946, indicating that the optimization has more than doubled the fraction of power transmitted through the bend. This improvement is further reflected by the mode matched coupling efficiency, shown by the red curve in Figure 2.9, which is improved to over 98%.

The optimized structure which achieves this result is depicted on the right side of Figure 2.9. Intuitively, the radius of the bend has increased significantly in order to reduce the amount of power leaked from the waveguide. In practice, we would expect the optimization to make the bend radius as large as possible. Because we are constrained to a fixed simulation domain, we set a bound of $3\,\mu\mathrm{m}$ on the maximum bend radius. The outer bend radius of the optimized structure is increased to $3\,\mu\mathrm{m}$. Interestingly, the inner radius of the optimized structure is not chosen to maintain a fixed waveguide width. Instead, the bend radius is smaller than we might expect, causing the waveguide to narrow slightly in the bend region. One explanation for this is that the slight narrowing into and out of the bend results in a better overlap between the modes in the different sections of the bend. This indicates that a simple bent waveguide can be improved by choosing bend radii which narrow the waveguide slightly.

This highlights a great strength of gradient-based optimization: it allows us to design more efficient devices than our intuition would typically allow. Our success in achieving this depends strongly on our ability to compute accurate gradients in an efficient manner. In this example, we have demonstrated that as long as we are careful to execute each step of the adjoint method correctly, we can successfully calculate gradients efficiently and with the desired accuracy. This process with minimal modification serves as an effective foundation for most optimizations of nanophotonic structures that we have pursued and that we will discuss in later chapters.

Chapter 3

Developing an Optimization Toolchain

Successful application of gradient-based optimization using the adjoint method hinges on our ability to execute forward and adjoint simulations and compute material gradients in a way that is self consistent. Accomplishing this using commercial or other third party Maxwell solvers for which the internals of the software are not well understood can be difficult.

In order to test and demonstrate the methods we have developed, a large portion of our work has been dedicated to building a complete toolchain for electromagnetic device optimization. The fruit of these efforts is an open source toolbox called EMopt [35] which provides all of the core components needed to simulate and optimize a diverse set of electromagnetic devices as depicted in Figure 3.1. EMopt provides stand-alone two dimensional and three dimensional Maxwell solvers, 1D and 2D mode solvers, an easily extensible adjoint method implementation, a boundary smoothing implementation, and a variety of tools like common figure of merit definitions and tools for aiding in the creation of complex geometries. As evidence of EMopt's capabilities, all optimized structures in this work have been generated using EMopt alone.

Optimizing electromagnetic devices, particularly in three dimensions, is a very computationally intensive task. In order to maximize performance, EMopt employs parallelism from the ground up based on the Message Passing Interface (MPI). This allows EMopt to run on any number of cores, whether it is 16 cores in a single server or 256 cores distributed across multiple servers connected by infiniband interconnects.

In order to further optimize performance, many of the core components of EMopt are written in C++. Higher level operations and all interaction with the toolbox, meanwhile, is performed using the python programming language (EMopt itself is wrapped as a python module). This allows the user to rapidly setup and debug optimizations while still benefiting from the raw performance of compiled

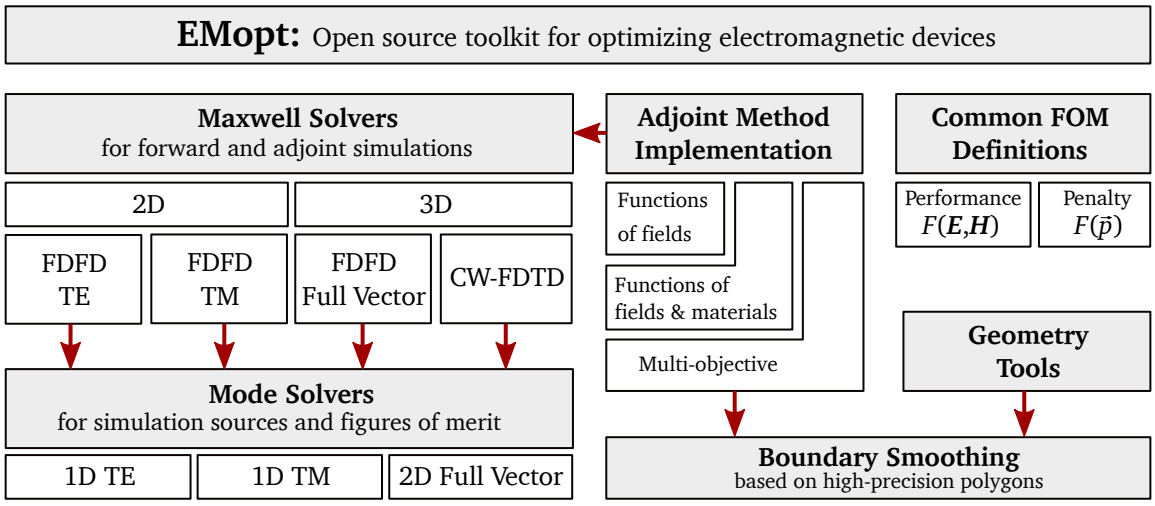

Figure 3.1: Graphical overview of the open source EMopt software toolbox. EMopt provides all of the simulators, mode solvers, adjoint method implementation, and geometry tools needed to optimize most complex electromagnetic structures.

languages.

In this chapter, we will present a brief overview of the functionality of each of the core components of EMopt, and highlight some of the details of our implementations which are important to the adjoint method. The code (and supporting documentation) relevant to each section is available in [35].

3.1 Maxwell Solvers

Perhaps the most important and most complex parts of EMopt are its Maxwell solvers. In order to simulate and optimize electromagnetic devices, it is important that we have an efficient simulator. To this end, EMopt provides a variety of different options. In two dimensions, EMopt implements a transverse magnetic finite difference frequency domain (FDFD) solver and a transverse electric FDFD solver. In three dimensions, EMopt implements a preconditioned iterative FDFD solver as well as a finite difference time domain (FDTD) solver.

In general, for 3D problems, EMopt's FDTD solver is faster, more memory efficient, and scales to larger numbers of cores better than its FDFD counterpart. For the majority of problems, the FDTD solver is preferred. In two dimensions, EMopts FDFD solvers based on LU factorization are fast and appropriate for most problems.

3.1.1 Finite Difference Frequency Domain Solvers

Perhaps the easiest Maxwell solver to implement is the finite difference frequency domain (FDFD) method [36–39]. FDFD involves discretizing the time-harmonic

Maxwell's equations on a rectangular grid using centered finite differences, assembling the resulting discretized equations in a matrix equation of the form $A\vec{x} = \vec{b}$, and then solving that system of equations using either a direct solver or preconditioned iterative solver.

In EMopt, a non-dimensionalized form of Maxwell's equations is used. This allows us to work in any length unit we choose and eliminates fundamental constants. In particular, the equations we solve are

$$\nabla \times \mathbf{H} + i\epsilon_r \mathbf{E} = \mathbf{J} \tag{3.1}$$

$$\nabla \times \mathbf{E} - i\mu_r \mathbf{H} = \mathbf{M} \tag{3.2}$$

which are derived in Appendix B. In these equations, $\mathbf{E}$, $\mathbf{H}$, $\mathbf{J}$, and $\mathbf{M}$ are the electric field, magnetic field, electric current density, and magnetic current density. The magnetic current density, though not observed in nature, is included to symmetrize the equations and enables us to inject power unidirectionally. Note that ϵ_r and μ_r are the unitless complex relative permittivity and permeability at the desired frequency. In order to define the frequency of the excitation, we specify the corresponding vacuum wavelength which is then accounted for in the non-dimensionalization of the equations. All subsequent physical lengths that we work with have the same units as the defined wavelength[1].

The FDFD method solves these equations by discretizing them on a rectangular grid. Unlike some implementations of the method [37, 39], we do not eliminate $\mathbf{H}$ from the equations, instead choosing to keep both Faraday's and Ampere's law separate[2]. In order to discretize these equations, it is convenient to first expand the curls and write six coupled equations:

$$\begin{aligned}
\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} + i\varepsilon_r^x E_x = J_x \qquad & \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} - i\mu_r^x E_x = M_x \\
\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} + i\varepsilon_r^y E_y = J_y \qquad & \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} - i\mu_r^y E_y = M_y \\
\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} + i\varepsilon_r^z E_z = J_z \qquad & \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} - i\mu_r^z E_z = M_z
\end{aligned} \tag{3.3}$$

Notice that here we have assumed that the permittivity and permeability are actually diagonally anisotropic. This adds no additional complexity to the equations and

[1] This means that if we want to excite our simulation with a source whose vacuum wavelength is $1.55\,\mu\mathrm{m}$ and we want to define the structures in our simulation in units of micron, we would set $\lambda = 1.55$. If we then set some length $L = 1.0$, then L corresponds to a length of $1\,\mu\mathrm{m}$.

[2] Elimination of the magnetic field appears to be, at least in part, done to mimic the common variational formulation of Maxwell's equations used in finite element analysis. In our experience, not combining the equations actually yields a matrix with lower condition number and a similar number of non zero matrix elements.

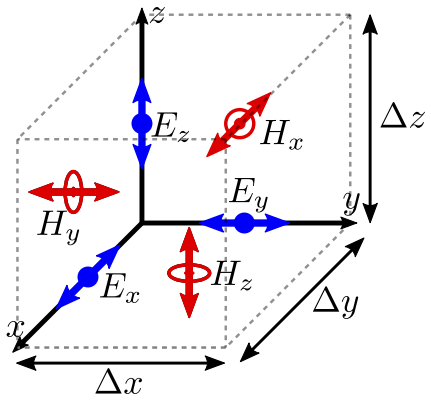

Figure 3.2: Three dimensional Yee cell configuration used in our finite difference frequency domain and finite difference time domain solvers.

generalizes them to a small degree. In these equations, the fields are continuous functions of x, y, and z. Numerically, however, we will express the fields on a set of six staggered grids. One “grid cell" (also known as a Yee cell) which makes up these staggered grids is depicted in Figure 3.2. This staggering lends itself to computing centered finite differences of the fields and also ensures numerical stability.

Based on this choice of grid cell layout, we can discretize the spatial derivatives in Equation (3.3) using centered finite differences. Assuming the location of each field component is described by three spatial indices i, j, k corresponding to the z, y, and x axes, respectively, the new discretized equations are given by

$$\frac{1}{\Delta y}\left(H^z_{i,j,k} - H^z_{i,j-1,k}\right) - \frac{1}{\Delta z}\left(H^y_{i,j,k} - H^y_{i-1,j,k}\right) + i\varepsilon^x_{i,j,k}E^x_{i,j,k} = J^x_{i,j,k} \tag{3.4}$$

$$\frac{1}{\Delta z}\left(H^x_{i,j,k} - H^x_{i-1,j,k}\right) - \frac{1}{\Delta x}\left(H^z_{i,j,k} - H^z_{i,j,k-1}\right) + i\varepsilon^y_{i,j,k}E^y_{i,j,k} = J^y_{i,j,k} \tag{3.5}$$

$$\frac{1}{\Delta x}\left(H^y_{i,j,k} - H^y_{i,j,k-1}\right) - \frac{1}{\Delta y}\left(H^x_{i,j,k} - H^x_{i,j-1,k}\right) + i\varepsilon^z_{i,j,k}E^z_{i,j,k} = J^z_{i,j,k} \tag{3.6}$$

$$\frac{1}{\Delta y}\left(E^z_{i,j+1,k} - E^z_{i,j,k}\right) - \frac{1}{\Delta z}\left(H^y_{i+1,j,k} - H^y_{i,j,k}\right) - i\mu^x_{i,j,k}H^x_{i,j,k} = M^x_{i,j,k} \tag{3.7}$$

$$\frac{1}{\Delta z}\left(E^x_{i+1,j,k} - E^x_{i,j,k}\right) - \frac{1}{\Delta x}\left(E^z_{i,j,k+1} - E^z_{i,j,k}\right) + i\mu^y_{i,j,k}H^y_{i,j,k} = M^y_{i,j,k} \tag{3.8}$$

$$\frac{1}{\Delta x}\left(E^y_{i,j,k+1} - E^y_{i,j,k}\right) - \frac{1}{\Delta y}\left(E^x_{i,j+1,k} - E^x_{i,j,k}\right) + i\mu^z_{i,j,k}H^z_{i,j,k} = M^z_{i,j,k} \tag{3.9}$$

where Δx is the grid spacing along the x direction, Δy is the grid spacing along the y direction, and Δz is the grid spacing along the z direction. These equations are equally valid at every index in our 3D grid, with the exception of grid cells which lie along the simulation boundaries. Depending on our boundary conditions, certain field components in the derivative terms will be zero or multiplied by a constant.

To solve these equations, we first assemble them in a large sparse matrix equation,

$$\begin{bmatrix} i\varepsilon_x & 0 & 0 & 0 & -D_{z-} & D_{y-} \\ 0 & i\varepsilon_y & 0 & D_{z-} & 0 & -D_{x-} \\ 0 & 0 & i\varepsilon_z & -D_{y-} & D_{x-} & 0 \\ 0 & -D_{z+} & D_{y+} & -i\mu_x & 0 & 0 \\ D_{z+} & 0 & -D_{z+} & 0 & -i\mu_y & 0 \\ -D_{y+} & D_{x+} & 0 & 0 & 0 & -i\mu_z \end{bmatrix} \begin{bmatrix} \vec{E}_x \\ \vec{E}_y \\ \vec{E}_z \\ \vec{H}_x \\ \vec{H}_y \\ \vec{H}_z \end{bmatrix} = \begin{bmatrix} \vec{J}_x \\ \vec{J}_y \\ \vec{J}_z \\ \vec{M}_x \\ \vec{M}_y \\ \vec{M}_z \end{bmatrix} \tag{3.10}$$

which has the form $A\vec{x} = \vec{b}$. Here, the blocks matrices $D_{i\pm}$ correspond to the derivatives of E and H, and the $\pm$ subscript refers to the difference in how the electric and magnetic fields are indexed in the derivative terms. The vector $\vec{x}$ is assembled from stacking field vectors, where each field vector contains the the field component values at every point in the grid. The source vector $\vec{b}$ is defined in a similar way. Notice that based on our previous discretized Equations (3.4)-(3.9), our matrix A has only five non zero elements per row.

With the matrix built, we can solve for the unknown field values contained in $\vec{x}$. To do so, EMopt employs the PETSc library [40] which provides parallelized direct and iterative solvers for large sparse systems. In 3D, a direct solver based on LU factorization is impractical due to fill-in (which is on the order of 100 for these equations). Instead, we employ a preconditioned Krylov subspace method. At the time of writing this manuscript, we use a multigrid preconditioner [41] which uses a preconditioned GMRES solver for its smoothers and uses LU decomposition at the coarsest level. As the preconditioner for the smoothers, we multiply by the conjugate transpose of our matrix. The largest eigenvalues of the product A^*A are real valued and tightly clustered, leading to an accelerated convergence of the iterative Krylov subspace solvers.

Our discussion up until now has focused on the full three dimensional form of Maxwell's equations. EMopt also implements 2D FDFD solvers for both transverse electric and transverse magnetic fields. The process for setting up and solving the 2D equations is nearly identical to the 3D equations. In 2D, EMopt assumes that the structure is infinitely extruded in the z direction. This means that all $\partial/\partial z$ terms in Equation (3.3) are zero. This leads to two independent sets of coupled equations given by:

$$\begin{aligned}
\frac{\partial H_z}{\partial y} + i\varepsilon_r E_x &= J_x & \frac{\partial E_z}{\partial y} - i\mu_r H_x &= M_x \\
-\frac{\partial H_z}{\partial x} + i\varepsilon_r E_y &= J_y & -\frac{\partial E_z}{\partial x} - i\mu_r H_y &= M_y \\
\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} - i\mu_r H_z &= M_z & \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} + i\varepsilon_r E_z &= J_z
\end{aligned} \tag{3.11}$$

The equations on the left correspond to the transverse magnetic (TM) equations while the fields on the right correspond to the transverse electric (TE) equations. In EMopt, these equations are discretized on a grid resembling Figure 2.1, assembled into a matrix, and solved using the PETSc library.

Unlike in 3D, the 2D matrix exhibits minimal fill-in ($\lesssim 10$) during LU factorization. This combined with the fact that the 2D matrices are significantly smaller in general makes direct solvers based on LU factorization an excellent choice for solving the system of equations. Specifically, EMopt uses MUMPS [42, 43] for its 2D TE and TM Maxwell solvers.

Using a direct solver based on LU factorization has a significant benefit from the standpoint of implementing the adjoint method. Running a forward simulation involves solving $A\vec{x} = LU\vec{x} = \vec{b}$, where L and U are the factored lower and upper triangular matrices. Running an adjoint simulation, meanwhile, involves solving $A^T\vec{y} = \vec{c}$. Notice that $A^T = (LU)^T = U^T L^T$. This tells us that the L and U factors we computed during the forward simulation can be *reused* for the adjoint simulation. For the forward and adjoint simulations, the LU factorization is the most computationally expensive operation; once these factors are known, the fields are found using inexpensive forward and back substitution. As a result, when solving our 2D system of equations using LU factorization, the adjoint simulation is effectively *for free*. This can lead to a nearly two times improvement in speed when running two dimensional optimizations.

3.1.2 Finite Difference Time Domain Solver

The finite difference frequency domain, while fast and acceptably memory efficient in 2D, does not scale well in 3D. In our experience, problems which have more than $\sim 125^3$ grid elements become slow enough that FDFD ceases to be an attractive method. This makes it largely incompatible with a lot of interesting problems in nanophotonics where devices can be more than $10\,\mu\mathrm{m}$ long and require relatively fine grid spacing to get accurate results.

To get around this issue, EMopt implements a simple finite difference time domain (FDTD) solver which is popular in the electromagnetics community [44, 45]. For even modest sized problems, FDTD is typically faster and less memory than FDFD. This is a direct result of the fact that in FDFD we not only need to store a copy

of the fields ($\vec{x}$), we also need to store a matrix which contains five non zero values per field value[3]. The gap in memory requirements between FDFD and FDTD is further widened when using Krylov subspace methods which require storing additional basis vectors which have the same length as the field vector $\vec{x}$. In FDTD, nominally we only need to store the fields and material distributions, which is equivalent to storing a two copies of $\vec{x}$. In terms of computational complexity, FDTD executes a single "matrix-vector multiplication" per time step. Iterative matrix solvers, on the other hand, require at least one matrix-vector product and multiple vector-vector products per iteration (in the case of Krylov subspace methods, many matrix-vector products may be required). Furthermore, compared to published results on preconditioned FDFD solvers [39], we have found that FDTD requires nearly an order of magnitude fewer iterations (i.e., time steps) in some examples.

In our implementation of FDTD, the electric and magnetic field are solved on a uniform rectangular grid[4] consisting of the Yee cells shown in Figure 3.2. A key benefit of the FDTD algorithm is that it is easily parallelizable and scales very well on distributed computing architectures. In our implementation, the 3D grid is broken up into rectangular "chunks" which are distributed across multiple cores/processors with the help of MPI and PETSc. This allows the solver to scale from a single machine to a computing cluster with ease.

A key differentiating factor of our FDTD implementation is that it simulations are driven by a ramped continuous wave (CW) source. This allows us to operate the simulator as a frequency-domain solver which solves for the fields at any desired frequency. This ultimately makes our FDTD simulator compatible with the same frequency-domain adjoint method implementation as the FDFD solvers without modification.

Because we are using a ramped CW source, we do not need to calculate the frequency response of the fields using a discrete Fourier transform as is typically done. Instead, we assume a sinusoidal time-dependence of the fields, i.e.,

$$E(t) = A\sin(\omega t + \varphi) + b \tag{3.12}$$

and solve for the field amplitudes A and phase φ at each point in the grid. In order to find the amplitude and phase, we record the fields in the system at three distinct time points (ideally separated by an odd multiple of a quarter of period of the oscillation):

[3]Technically, we could implement a matrix-free iterative solver. This would eliminate the need to explicitly store matrix elements, thereby reducing the memory footprint as well as speeding up each iteration of the solver.

[4]A uniform grid is necessary in order to ensure that the system of equations is approximately symmetric, which is important when using the FDTD solver to compute gradients with the adjoint method.

$$E(t_1) = A\sin(\omega t_1 + \varphi) + b \tag{3.13}$$
$$E(t_2) = A\sin(\omega t_2 + \varphi) + b \tag{3.14}$$
$$E(t_3) = A\sin(\omega t_3 + \varphi) + b \tag{3.15}$$

These field samples and the assumed sinusoidal time dependence form a system of three equations with three unknowns. Solving for the unknowns of this system is relatively straightforward. First, the phase of the field is given by

$$\tan\varphi = \frac{E_{21}\left[\sin\omega t_3 - \sin\omega t_2\right] - E_{32}\left[\sin\omega t_2 - \omega\sin\omega t_1\right]}{E_{32}\left[\cos\omega t_2 - \cos\omega t_1\right] - E_{21}\left[\cos\omega t_3 - \cos\omega t_2\right]} \tag{3.16}$$

where $E_{21} = E(t_2) - E(t_1)$ and $E_{32} = E(t_3) - E(t_2)$. Once the phase is known, we can find the amplitude of the field, which is given by

$$A = \frac{E_{32}}{\cos(\varphi)\left[\sin\omega t_2 - \sin\omega t_1\right] + \sin(\varphi)\left[\cos\omega t_2 - \cos\omega t_1\right]} \tag{3.17}$$

Given this amplitude and phase, the complex field at each point in the simulation domain is given by

$$\widetilde{E}_{mnp} = A_{mnp}e^{i\varphi_{mnp}} \tag{3.18}$$

where m, n, and p are the spatial indices of the grid. This method of calculating the frequency domain fields has the benefit that it does not impact the runtime of the simulator significantly since no additional computation has to be performed during the time stepping process. Furthermore, this process is easily extended to multiple wavelength sources, as is described in appendix C.

3.1.3 Mode Solvers

In addition to Maxwell solvers, EMopt implements both 1D and 2D waveguide mode solvers. Waveguide modes are useful both as simulation source and as reference fields in figures of merits used in optimizations.

The process EMopt follows to solve for waveguide modes is as follows. First, the waveguide mode is assumed to propagate in the $+z$ direction and have the spatial dependence:

$$\mathbf{E}(x,y,z) = \mathbf{E}_m(x,y)e^{ik_z z} \tag{3.19}$$
$$\mathbf{H}(x,y,z) = \mathbf{H}_m(x,y)e^{ik_z z} \tag{3.20}$$

where $\mathbf{E}_m(x,y)$ and $\mathbf{H}_m(x,y)$ are the transverse mode profiles and k_z is the effective wavenumber. Next, we plug this result into Equation (3.3) while simultaneously setting the current sources to zero and dropping the m subscript for convenience. This yields a new eigenvalue problem given by

$$\begin{aligned}
\frac{\partial H_z}{\partial y} + i\varepsilon_r^x E_x &= ik_z H_y & \frac{\partial E_z}{\partial y} - i\mu_r^x E_x &= ik_z E_y \\
-\frac{\partial H_z}{\partial x} + i\varepsilon_r^y E_y &= -ik_z H_x & -\frac{\partial E_z}{\partial x} - i\mu_r^y E_y &= -ik_z E_x \\
\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} + i\varepsilon_r^z E_z &= 0 & \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} - i\mu_r^z E_z &= 0
\end{aligned} \tag{3.21}$$

Just as with the FDFD method, we can discretize these equations on a staggered grid and assemble the resulting equations into a matrix. The resulting matrix equations takes the form $A\vec{x} = k_z B\vec{x}$ which is a generalized eigenvalue problem. Here, the matrix B is diagonal and filled with either i, $-i$, or 0 (in the diagonal blocks corresponding to E_z and H_z). In order to solve this sparse generalized eigenvalue problem with parallelism, EMopt uses the SLEPc library [46] and MUMPS.

EMopt also implements 1D TE and TM mode solvers for 2D waveguiding systems. These mode solvers are implemented in an analogous way to the 2D mode solver.

3.2 General Adjoint Method Implementation

EMopt's primary goal is to provide us with a way to efficiently calculate accurate sensitivities of electromagnetic structures (i.e. gradients of a figure of merit with respect to some set of design parameters which define the device's shape). In order to calculate these sensitivities, EMopt implements a general form of the adjoint method which can be applied to almost any shape optimization problem in electromagnetics.

The process EMopt follows is depicted in Figure 3.3. After the structure has been defined and the simulation setup, a forward simulation is run using one of EMopt's solvers. Next, EMopt calculates the user-defined figure of merit. The derivative of that figure of merit with respect to the fields is then computed based on the user-defined derivative. This derivative serves as the source current density distribution for an adjoint simulation which is performed using one of EMopt's aforementioned solvers.

With the forward and adjoint fields computed, EMopt moves on to calculating the material gradients, which are approximated using finite differences. One at a time, the design variables of the system are perturbed. For each perturbation, the permittivity and permeability distributions are updated based on a user-defined

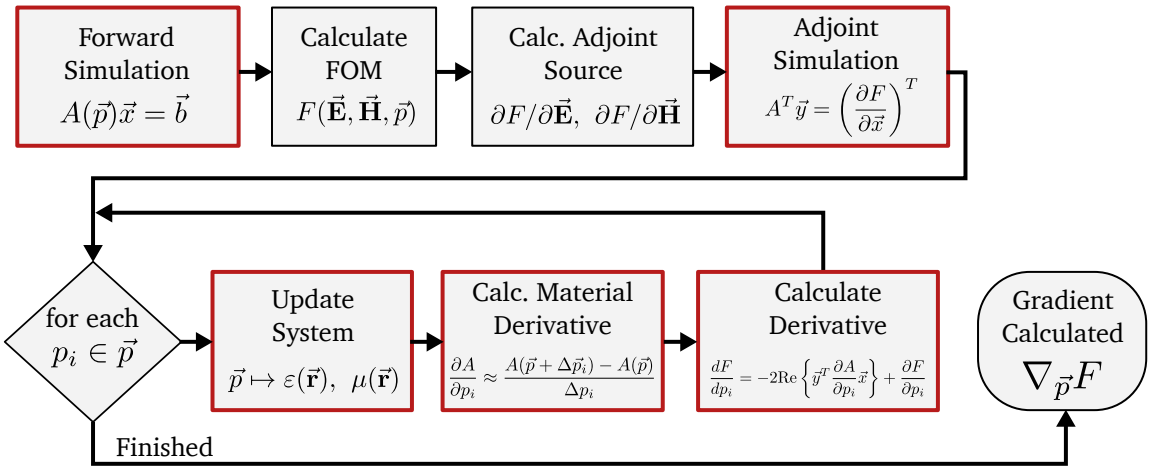

Figure 3.3: Flowchart showing a general adjoint method implementation for shape optimization. By approximating the material gradient using finite differences, the gradient of the figure of merit is easily computed for arbitrary parameterizations. The bold red outlined boxes denote steps in the process that leverage parallel computing for improved performance.

mapping (i.e., the user defines how the geometry of the system should change when the design variables change). Based on the updated material distributions and the original unperturbed material distributions which were computed prior to running the forward simulation, the material gradients are computed. These material gradients are then used to calculate the derivative of the figure of merit with respect to the i'th design variable using Equation (2.15).

By computing the material derivatives using finite differences, it is straightforward to calculate gradients with respect to arbitrary parameterizations[5]. Furthermore, because the unperturbed material distributions must be calculated in order to simulate the structure, computing derivatives using finite differences incurs no additional cost compared to an analytic approach. In order to further speed up this calculation, for each perturbed parameter, only a small part of the full material distribution is updated. This is possible because in most parameterizations, each design parameter has a localized effect on the distribution of permittivity in the system.

Throughout this flexible implementation of the adjoint method, EMopt employs parallelism in order speed up the calculations. In particular, the red boxes in Figure 3.3 indicate parts of the process which are parallelized. Not only are the Maxwell

[5] This is compared to an analytic implementation which would require that the user provide analytic derivatives of the underlying shape properties with respect to the parameterization. When working with polygons as we primarily do in this work, this means calculating the derivatives of the x and y coordinates of the polygon vertices with respect to the design parameters. In many cases, this is a tedious process which can be accomplished much more easily using a finite difference

solvers parallelized, but so are the material derivative and gradient calculations.

3.3 Grid Smoothing Implementation

The primary focus of EMopt at this time is shape optimization. EMopt must therefore ensure that gradients computed with respect to changes in shape boundaries are accurate. To accomplish this, EMopt implements the boundary smoothing methods discussed in Section 2.2 based on high numerical precision polygons.

The specific approach that EMopt implements in two dimensions[6] is depicted in Figure 3.4 and detailed in Algorithm 1. First, the finite difference grid is divided up into grid cells G_{ij} which are centered at the field components and have width Δx and height Δy (equal to the grid spacing along x and y). Next, the geometry of the system is defined using a set of polygons S which are filled with different material values and ordered according to a layer number which defines which shapes appear on top of others. For each grid cell G_{ij}, the intersection between the grid cell and the top-most overlapping polygon is computed. The permittivity of the current grid cell is then updated based on the overlapping area and the total area of the grid cell. This process is then repeated, replacing the grid cell G_{ij} with the difference between the grid cell and the overlapping polygon. The calculation of the effective permittivity of each grid cell is complete when the area of the remaining portion of the grid cell is zero or if no overlapping polygons remain. The end result of this process resembles the examples presented in Section 2.2.

The performance of this method depends on a number of factors. Because the primary operation in computing smoothed boundaries is the polygon intersection and difference operations, the specific implementations of these operations has a strong impact on the overall speed of the algorithm. Fortunately, a number of high quality computational geometry libraries (notably CGAL and Boost.Geometry) exist which provide high performance implementations of the required geometric operations. Furthermore, because the calculation of the effective permittivity in each grid cell is independent of the other grid cells, the algorithm is easily parallelized, leading to a near linear improvement in performance when run on multiple processors.

[6] In 3D, EMopt defines the geometry as slabs which are stacked in the z direction. In each slab, the 2D grid smoothing process is applied.

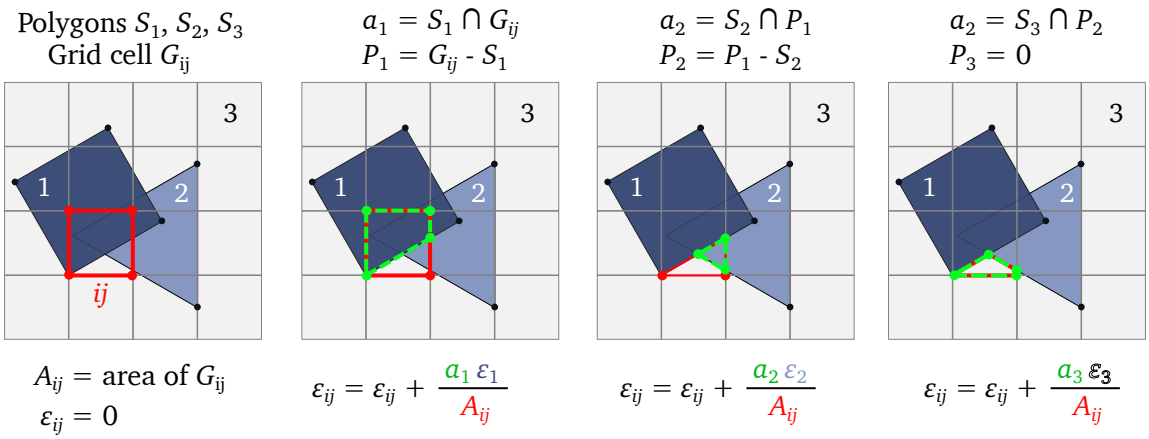

Figure 3.4: Graphical explanation of the grid smoothing algorithm based on high precision polygons. All shapes and geometry is represented using layered polygons. The intersection between a grid cell and the top-level shape is computed and then subtracted from the grid cell polygon. This process is repeated for each overlapping shape in the system until the grid cell polygon is entirely subtracted or no more shapes exist. The areas of the overlap polygons computed in each step serve as the weights in the calculation of the average permittivity.

Algorithm 1: Polygon-based boundary smoothing

Data: $S \leftarrow$ Ordered set of polygons,
$G \leftarrow$ Set of grid cells which define FDFD/FDTD grid
Result: $\varepsilon \leftarrow$ Permittivities associated with the set of grid cells$G$
$\varepsilon_0 \leftarrow 1$;
for *each individual grid cell* G_{ij} *in* G *and* ε_{ij} *in* ε **do**
  $A_{ij} \leftarrow$ Area of G_{ij};
  $\varepsilon_{ij} \leftarrow 0$;
  $P \leftarrow G_{ij}$;
  for *each polygon* S_k *in* S **do**
    $e_i \leftarrow$ Permittivity inside polygonS_k;
    $a_k \leftarrow$ Area of intersection $S_k \cap P$;
    $\varepsilon_{ij} \leftarrow \varepsilon_{ij} + a_k e_i / A_{ij}$;
    $P \leftarrow P - S_1$;
    if *Area of* $P = 0$ **then**
      break;
    end if
  end for
  $\varepsilon_{ij} \leftarrow P\varepsilon_0 / A_{ij}$;
end for

Chapter 4

Hierarchical Design and Optimization Approach

Shape and topology optimization methods, although demonstrated to be very powerful tools, often suffer a glaring flaw: they are inherently *local optimization* techniques. Consequently, it is difficult to make any guarantee that a design "discovered" by the optimization method will end up meeting a desired specification. This behavior becomes particularly evident when designing high performance devices whose insertion losses drop below $0.1\,\mathrm{dB}$. In such cases, it becomes increasingly important that we choose a starting point for the optimization which will lead to a final design with the desired high performance.

In spite of its importance, work on shape and topology optimization in nanophotonics has largely overlooked this process. In many works [10, 15, 16, 47–50] starting structures with seemingly (or even intentionally) arbitrary size and shape are chosen. These choices, while successful in seeding optimizations which yield drastic improvements in the figure of merit, often fail in producing final devices which significantly out perform their hand-designed counterparts (if at all).

In this chapter, we will introduce a hierarchical design methodology which largely eliminates guesswork from the process of designing and optimizing photonic devices. In our approach, which is depicted in Figure 4.1, we first specify an initial device topology based on a simple physical analysis of the problem. Next, using this starting structure we run a *coarse* optimization with only small set of design parameters which allows us to rapidly improve the general shape of the device. Finally, with the coarse optimization result as a starting point, we run a final *refinement* optimization with a large number of degrees of freedoms and constraints. Following this systematic approach to inverse design, one can design photonic components whose performance often far exceeds anything we could design by hand.

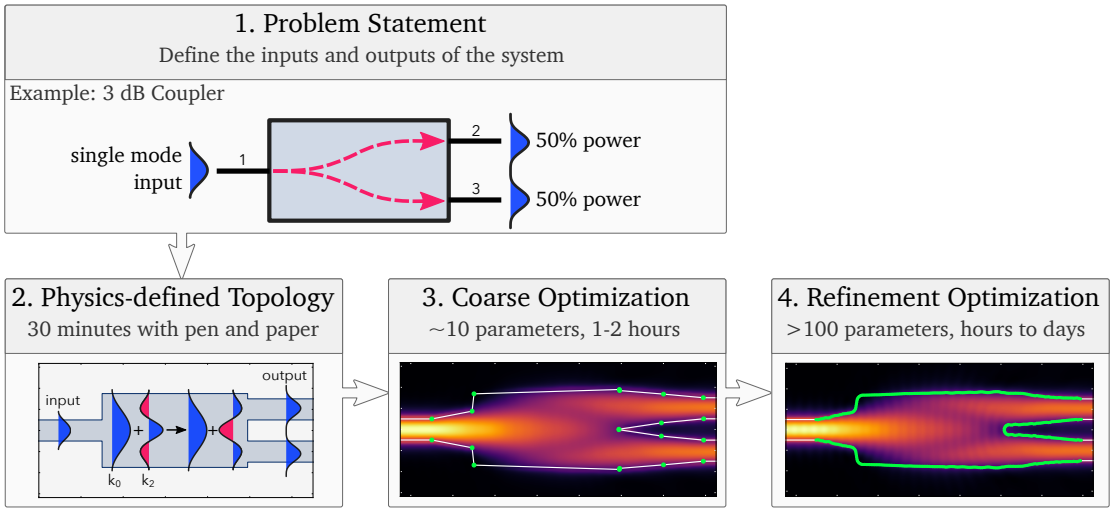

Figure 4.1: Graphical overview of our hierarchical approach to design and optimization of electromagnetic structures. Our systematic approach involves defining the problem we are solving, developing a device topology based on physical intuition, coarsely optimizing this starting topology to improve its performance, and then refining the coarsely optimized structure with a second optimization that incorporates constraints. The end result is a device with high performance that can be fabricated.

4.1 Physics-Defined Topology

The starting structure that we use to initialize an optimization is important for a variety of reasons. In addition to influencing which local optimum we fall into, it can place important constraints on the quality of the available local minima and even restrict the quality of the global optimum. The reason for this lies in the fact that we typically choose and fix the inputs and output of our device at the beginning of the design process. If we choose a starting structure with a certain size, then the final optimized structure will fit roughly within the same footprint. As a result, choosing a starting structure that is too small or too large to achieve the desired functionality can lead to a final result that is worse than we might desire.

The impact of the starting structure on the final optimized device is made particularly apparent in the work by Su et al. [50]. Using randomized initial structures as a starting point, the authors optimize hundreds of grating couplers. Of the optimized results, only a very small number of devices achieve a high coupling efficiency; the rest of the optimized structures exhibit lower performance which indicates that the corresponding optimizations fell into lower quality local optima. This observed sensitivity to initial conditions may give the impression that the task of optimizing a device is imprecise and riddle with guesswork. Fortunately, we have recourse: our strong physical intuition and understanding of Maxwell's equations. By leveraging what knowledge we have of wave mechanics, in many (if not all)

cases, we can devise physically-motivated starting structures which lead to efficient optimized devices. In some sense, this is not unlike choosing a circuit topology when designing electrical circuits: when designing a circuit, we choose a topology that we know will accomplish our goal and then we fine tune our circuit by selecting appropriate values for its constituent components.

We refer to this physically-motivated starting structure as the *physics-defined topology*. For many components in silicon photonics, a usable topology can be devised relatively easily by applying numerous techniques such as modal analysis.

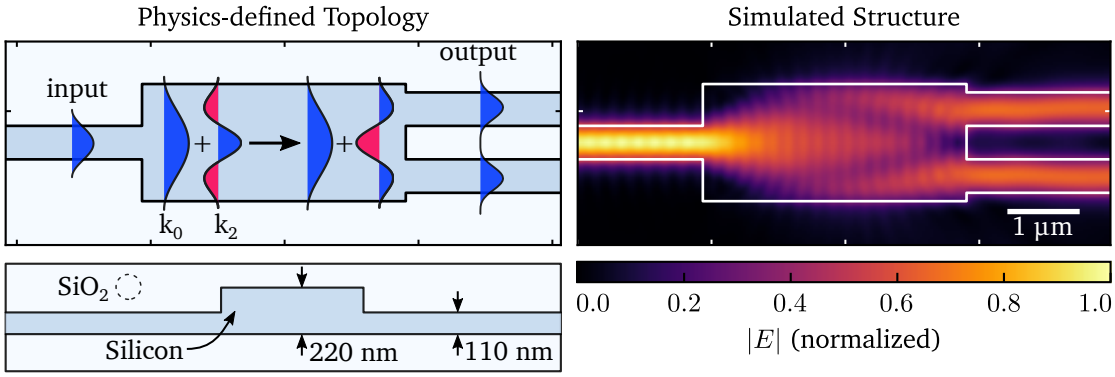

Figure 4.2: (left) Diagram of the physical principles which govern a multimode interference-based $3\,\mathrm{dB}$ coupler. (right) Plot of a slice of the simulated electric field of a $3\,\mathrm{dB}$ coupler based on the topology depicted on the left. The simulation of this structure whose cross section is depicted in the bottom left is performed for a free space wavelength of $1310\,\mathrm{nm}$.

As an example, let us consider the canonical problem of designing a three port $3\,\mathrm{dB}$ coupler. The goal of a $3\,\mathrm{dB}$ coupler is to split power from an input waveguide equally between two output waveguides. One way of achieving this functionality is to take advantage of modal dispersion in a multimode waveguide [51] which is depicted on the left hand side of Figure 4.2. A narrower waveguide connected to a wider multimode waveguide will excite the symmetric modes of that wider waveguide. If the multimode waveguide is not too wide, the majority of the power will reside in the first two even modes of the waveguide. These two modes, with wavenumbers k_0 and k_2 respectively, co-propagate and their relative phase difference,

$$\Delta\varphi = (k_0 - k_2)\Delta z \tag{4.1}$$

increases with distance Δz. Notice that if these two modes propagate over a sufficient distance, the relative phase difference of the two modes will become π, leading to a cancellation of the field in the middle of the multimode waveguide and a concentration of the field at the outer edges of the waveguide. This effect occurs over a distance

$$\begin{aligned} \Delta\varphi = \pi \Rightarrow \Delta z &= \frac{\pi}{k_0 - k_2} \\ \Rightarrow \Delta z &= \frac{\lambda}{2\left(n_0 - n_2\right)} \end{aligned} \tag{4.2}$$

where λ is the free-space wavelength and n_0 and n_2 are the effective indices of the two modes. With the field evenly split and concentrated near the outer edges of the waveguide, we can terminate the multimode waveguide and place two narrower output waveguides to capture the divided optical power as depicted in Figure 4.2.

This simple approach translates into a relatively well performing initial structure. To demonstrate this, we simulate an O-band (center wavelength of $1310\,\mathrm{nm}$) $3\,\mathrm{dB}$ coupler. The $3\,\mathrm{dB}$ coupler is defined in a $220\,\mathrm{nm}$ thick top silicon layer with a $110\,\mathrm{nm}$ deep partial etch as depicted in the bottom left of Figure 4.2. Based on our previously devised topology, the coupler consists of a wide rectangular section which is fed by a narrower input waveguide and connected to two narrower output waveguides. The structure is symmetric to ensure even coupling to the output waveguides. The width of the multimode coupling region is chosen to be $1.75\,\mu\mathrm{m}$ wide (somewhat arbitrarily to be large enough to span the two $500\,\mathrm{nm}$ output waveguides which are separated by $500\,\mathrm{nm}$). The effective indices for the first two even modes of this waveguide are 2.965 and 2.800, respectively. Based on these effective indices, the desired coupler length is $\sim 3.94\,\mu\mathrm{m}$.

The field simulated for this structure is shown on the right side of Figure 4.2. Despite the simplicity of this approach, the chosen topology yields device with a reasonably high coupling efficiency of $93.3\,\%$ ($-0.301\,\mathrm{dB}$). This gives us some confidence that the overall size of the device and the chosen positions of the inputs and outputs are sufficient for obtaining a high efficiency device.

This example demonstrates the general thought process one can follow to develop a topology for many silicon photonic components. In fact, the same logic and calculation can be applied to many other devices like waveguide crossings, 4 port $3\,\mathrm{dB}$ couplers, polarization splitters, etc which we will explore in the next chapter. Other devices, however, may require a different approach to discovering an effective topology; one such example is grating couplers which we also demonstrate in the next chapter. While developing the topology for an electromagnetic device may not always be as simple as this $3\,\mathrm{dB}$ coupler, we should nonetheless be able to apply our intuition of electromagnetics to come up with *something* which is superior to random guesses. To this end, the large amount of literature on electromagnetic device design is tremendously useful.

4.2 Coarse Optimization

One of the primary goals of defining a physically-inspired device topology is to ensure that the inputs and outputs of the system are situated such that high efficiency designs exist within the design space. In many cases, the analysis we follow to develop this topology neglects many of the finer details of these inputs and outputs. For example, in our $3\,\mathrm{dB}$ coupler example, the input and output waveguide widths were chosen to be $500\,\mathrm{nm}$ to match a fairly standard waveguide width. Furthermore, the separation of the output waveguides was chosen to approximately match the intensity distribution formed by the beating modes within the larger multimode waveguide section of the device. In reality, these abrupt transitions from input waveguide to multimode waveguide to output waveguides are unlikely to be optimal. We thus find ourselves with an opportunity to apply the optimization techniques we developed in earlier chapters.

While we may be tempted to throw the full power of inverse design with many degrees of freedom at the problem, it is important that we proceed in a strategic manner. There are a few reasons for this. Typically when we run shape (or topology) optimizations with large numbers of degrees of freedom, we impose constraints to prevent features from developing which we are unable to fabricate. The combination of large numbers of degrees of freedom and constraints significantly increases the complexity of the design problem leading to a larger number of local optima and a slower evolution towards the final optimized structure. This in turn makes it more difficult to make large modifications to the structure like displacing a boundary of the structure over $\sim$micron scale distances.

In order to mitigate these issues, it is instead desirable to run an initial *coarse optimization* without constraints and a smaller number of design variables ($\sim$10). By running an optimization on a coarsely parameterized structure, we are able to rapidly improve the the general shape (and in particular the positions and sizes of the inputs and outputs) of the device. Convergence is typically significantly faster for simpler design spaces, and thus the coarse optimization provides us with the means to quickly verify the quality of our topology and also evaluate the likelihood that the design process will lead to an efficient device. Furthermore, the result of the coarse optimization serves as an excellent starting point for a final optimization which includes more degrees of freedom and fabrication constraints.

In order to demonstrate the coarse optimization, let us continue with our example of designing a $3\,\mathrm{dB}$ coupler. In this example, we choose the coarse parameterization to be the coordinates of the points which define the splitter as depicted by the green dots in Figure 4.3. In total there are seven points and hence fourteen design parameters. This choice of parameterization allows the optimization to manipulate the input and output waveguides and also to modify the size of the multimode waveguide section.

The figure of merit used for this optimization is the coupling efficiency into the

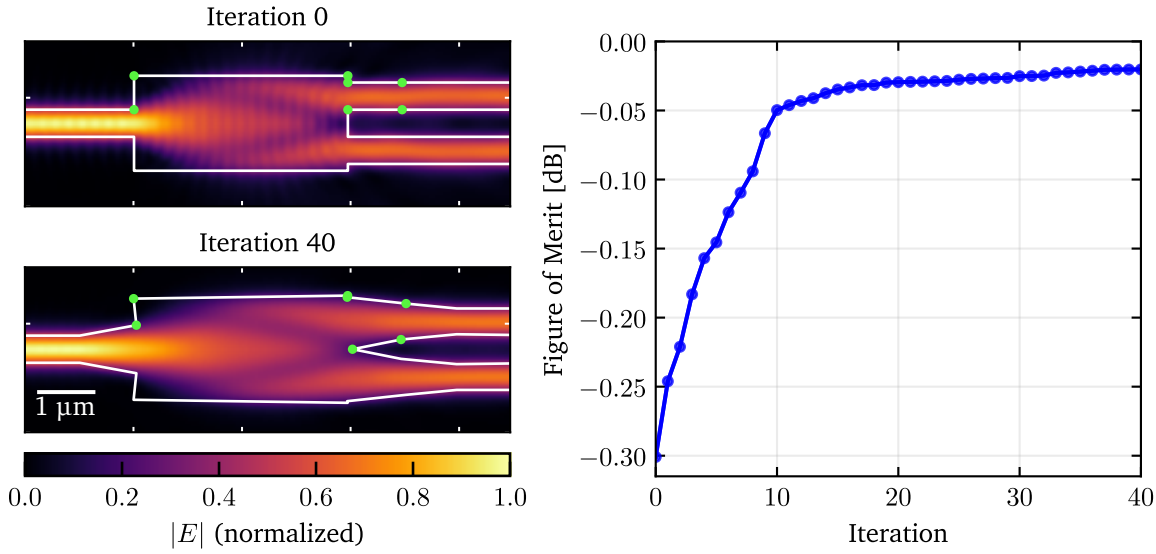

Figure 4.3: (left) Plots of the magnitude of the electric field of the initial structure and the final optimized structure are overlaid with the boundaries of the corresponding structures. The green vertices correspond to the designable points of the optimization. (right) The plot of the figure of merit as a function of iteration of the optimization. The figure of merit in the coarse optimization is simply the coupling efficiency to the desired output mode.

fundamental supermode of the two output waveguides (given by Equation (A.25)) which corresponds to equal splitting of optical power. All simulations are performed for a wavelength of $1310\,\mathrm{nm}$ with O-band operation in mind. The grid used has a step size of $30\,\mathrm{nm}$ which provides a good trade-off between simulation speed and moderate simulation accuracy. The optimization is terminated after 40 iterations at which point the figure of merit does not change appreciably.

In this optimization, we use the L-BFGS-B minimization algorithm which we have found to produce the most rapid convergence. The progression of the figure of merit during the coarse optimization is shown on the right side of Figure 4.3. Beginning with a value of $0.30\,\mathrm{dB}$, the insertion loss is quickly improved during optimization, reaching a very low value of only $0.020\,\mathrm{dB}$ by the end of the optimization. This whole process took just over 30 minutes on a small cluster consisting of 16 Intel Xeon E5-2670 CPUs[1] (totaling 128 cores).

4.3 Refinement Optimization

Although the coarse optimization is successful in producing a structure with high coupling efficiency, the structure itself is not entirely practical as it contains sharp

[1]At the time of authoring this manuscript, these processors were 7 years old. Newer processors with increased clock speed, cache, and bus speed will likely significantly outperform these numbers.

corners that are difficult to fabricate. Simply rounding off these sharp corners will inevitably result in a sub-optimal design. In order to circumvent this issue, we can run a second *refinement optimization* which includes more degrees of freedom and fabrication constraints. By adding more degrees of freedom, we are able to smooth out the corners in the structure. This smoothness is enforced by fabrication constraints which prevent features which are too small or too sharp from forming. The addition of these degrees of freedom has the added benefit that it affords us more flexibility in the shape of the device, allowing us to further improve its performance.

The goal of the refinement optimization is to produce a structure which can be fabricated and which meets a desired level of performance. Typically, this means that in addition to boosting the number of designable degrees of freedom in the device, we also impose a range of constraints. For waveguiding devices, the most common constraint is a radius of curvature constraint which limits the formation of small features. Depending on the device, a minimum gap size or minimum line width constraint may also be imposed. Choosing such constraints is an extremely important part of the refinement optimization and can have a strong impact on the final performance of the optimized device, the manufacturability of the device, and the time it takes to run the optimization.

To demonstrate this, let us continuing with our $3\,\mathrm{dB}$ coupler example. For the refinement optimization, we choose the figure of merit to be the coupling efficiency into the fundamental supermode of the two output waveguides plus two additional penalty functions which impose a radius of curvature constraint and a minimum gap size constraint for the output waveguides. This figure of merit is given by

$$F(\mathbf{E}, \mathbf{H}, \vec{p}) = \eta(\mathbf{E}, \mathbf{H}) - \mathcal{P}_{\mathrm{roc}}(\vec{p}) - \mathcal{P}_{\mathrm{gap}}(\vec{p}) \tag{4.3}$$

where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic fields and $\vec{p}$ is the set of design parameters. The individual functions η, $\mathcal{P}_{\mathrm{roc}}$, and $\mathcal{P}_{\mathrm{gap}}$ are discussed in detail in Appendix A. In this example, we choose a minimum radius of curvature of $120\,\mathrm{nm}$ and a minimum gap size of $200\,\mathrm{nm}$ to make the device compatible with readily available deep UV lithography. By imposing the radius of curvature and gap constraints using a penalty function, the constraints become part of the optimization. This ensures that the structure we end up with is optimal given these constraints.

It is important to note that if we used the structure generated by the coarse optimization without modification, the sharp corners would significantly violate the radius of curvature constraint. This would cause the corresponding penalty function to dominate the figure of merit. As a result, the refinement optimization would attempt to reduce the impact of the penalty function by rounding out the structure, a process which can lead to very slow convergence and is susceptible to falling into undesirable local optima. To avoid this, we can manually round off the corners of the coarse optimized structure (using a fillet operation) and then use this modified structure as the starting point for the refinement optimization. By doing so, we ensure that the radius of curvature (and gap) constraint is initially satisfied. In

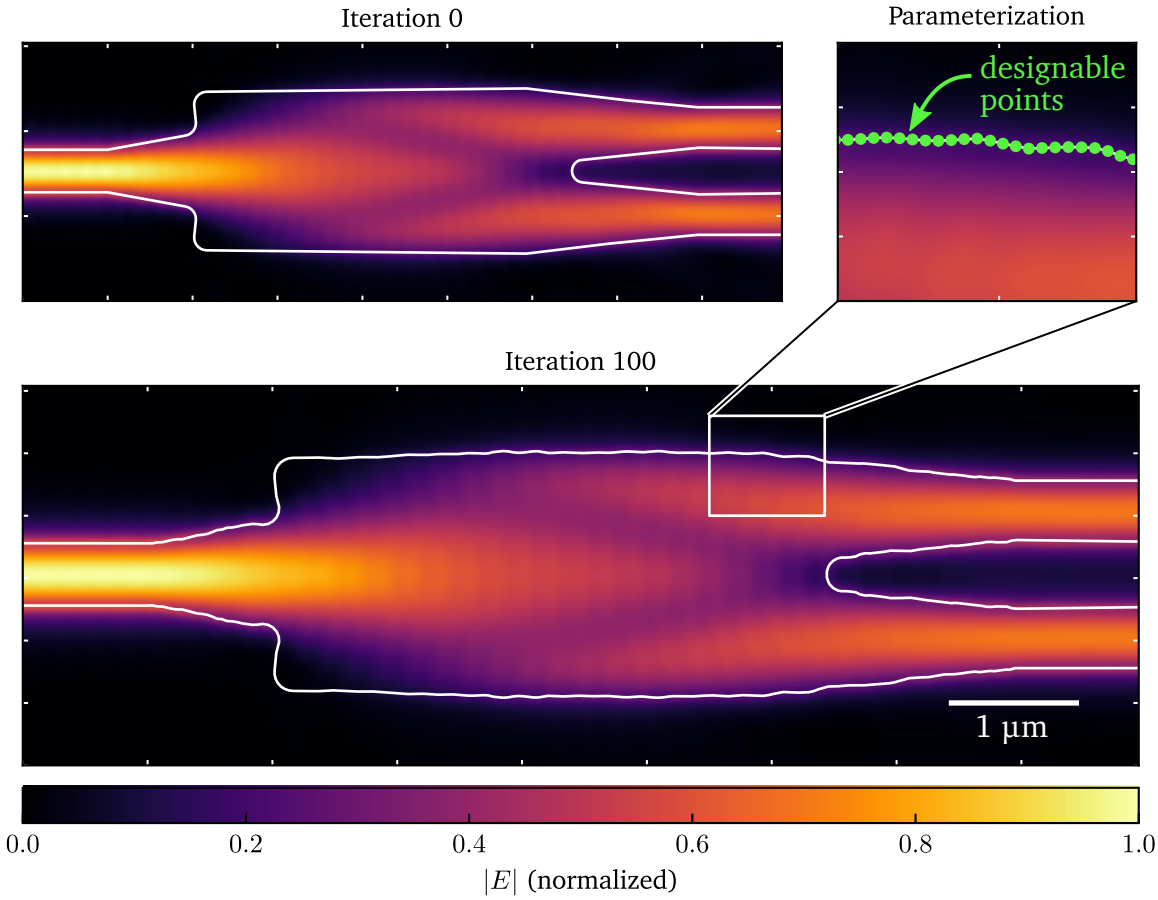

Figure 4.4: Plots of the initial rounded structure and final optimized structure for the refinement phase of the design process. The inset in the top right shows a section of the polygon which defines the boundary of the structure. In all three plots, the electric field amplitude is overlaid with an outline of the device boundaries.

this case, the penalty function acts as a barrier preventing the optimization from entering regions of the design space which cannot be fabricated. Although this process may lead to an initial drop in device performance, the subsequent refinement optimization can make up for the lost performance.

The rounded $3\,\mathrm{dB}$ coupler used as the starting point for the refinement optimization is depicted in the top left of Figure 4.4. In order to represent boundaries with rounded features, we increase the number of points in the polygon which defines the coupler geometry as shown in the inset in Figure 4.4. These additional points serve as the (significantly augmented) design space for the refinement optimization and enable the coupler to take on a more sophisticated shape than was possible with the coarse optimization. This is reflected in the optimized structure depicted in the bottom of Figure 4.4. This optimization is allowed to run for a

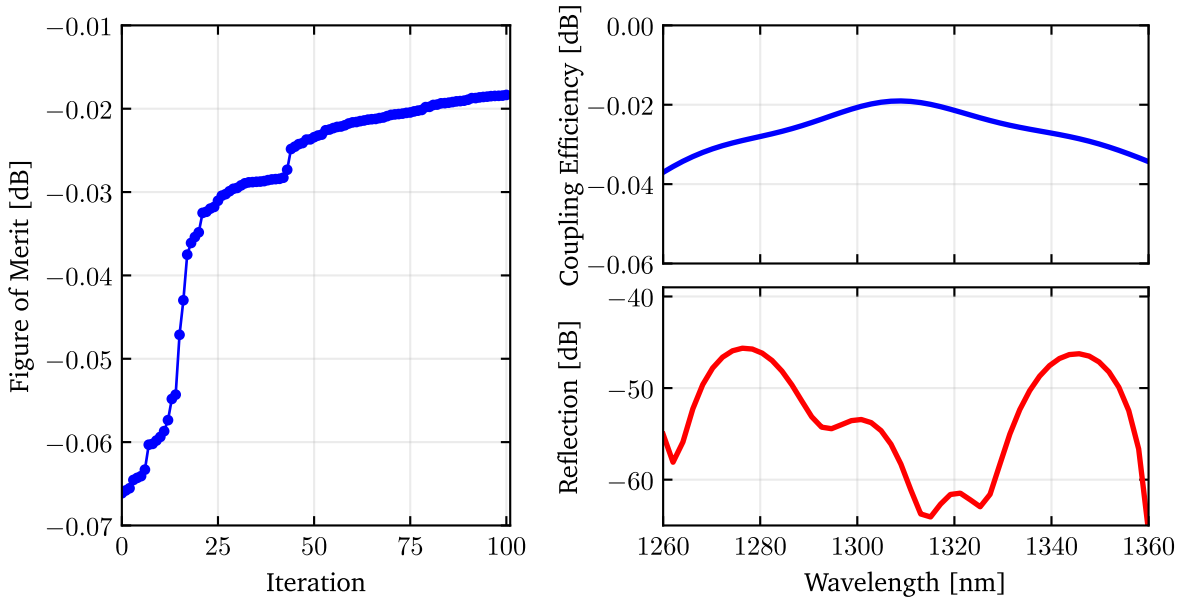

Figure 4.5: (left) The figure of merit is plotted after each iteration showing its improvement during the refinement optimization. (right) Plots of the O-band coupling efficiency and reflection for the final optimized device.

total 100 iterations which yields a satisfactory improvement in performance in a modest amount of time (a little under 9 hours when run on 128 cores of our Intel Xeon-based cluster). Over the course of the the refinement optimization, the figure of merit improves from $-0.065\,\mathrm{dB}$ to under $-0.02\,\mathrm{dB}$ as shown in the left plot of Figure 4.5. Although a few hundredths of a dB improvement may seem small, three of the final optimized devices could be cascaded together and still maintain lower loss than the starting structure.

Not only is the transmission at the design wavelength high, the coupling efficiency is maintained over a broad range of wavelengths as shown in Figure 4.5. The optimized coupler achieves an insertion loss lower than $0.04\,\mathrm{dB}$ over the entire O-band ($1260\,\mathrm{nm}$ to $1360\,\mathrm{nm}$). This performance exceeds other published results for both hand designed and numerically optimized three port $3\,\mathrm{dB}$ couplers [10, 52–57]. Furthermore, the back reflection into the input waveguide is lower than $-40\,\mathrm{dB}$ over the full wavelength as shown in 4.5. Although not included explicitly in the figure of merit, any reflected power detracts from the fraction of power leaving in the desired output mode. As a result, the optimization will naturally try to eliminate reflections.

It is interesting to note that in addition to larger modifications to the general shape of the device (which is particularly visible at the joining of the input waveguide to the multimode section), the refinement optimization introduced some smaller amplitude ripples in the structure. These ripples are direct consequence of

the wave nature of the device. From the formulation of the adjoint method, we know that the gradient of the figure of merit depends explicitly on a product involving two sets of (approximately) electromagnetic fields. In the presence of reflections, the amplitude of these fields will oscillate with a spatial frequency which is approximately equal to the wavelength of light in the material which makes up the device. In the case of this silicon $3\,\mathrm{dB}$ coupler, these amplitude variations have an expected periodicity of $\sim 350\,\mathrm{nm}$. This very closely matches the periodicity of the ripples we see in the optimized structure. While we might be concerned that these ripples could increase the sensitivity of the structure to any variations which occur during fabrication, it is important to note that these ripples still satisfy the fabrication constraints that we impose. Furthermore, the amplitude of these ripples are very small (on the order of $30\,\mathrm{nm}$ or smaller) and we have observed that they tend to form later in the optimization when changes to the figure of merit are very small. As such, we do not expect the ripples to significantly increase the sensitivity of the optimized structure to fabrication variations. To a certain degree, this idea is reinforced by the coupler's very broadband performance.

The final structure produced by the refinement optimization not only achieves high efficiency, but it does so while satisfying the radius of curvature constraint and gap size constraint that we imposed on the structure. As a result, we expect the simulated device to translate well to experiment with minimal additional loss. Notice that even if we modified our constraints (for example, if we improved our fabrication process and could hit smaller feature sizes), we only need to rerun our refinement optimization. To a certain degree, the device topology and coarse structure is independent of the particular constraints we impose. This presents an additional advantage of our hierarchical approach to nanophotonic design in that it allows us to more easily design component libraries which are compatible with a wider range of lithography requirements.

4.4 Putting it all together: Hierarchical Design and Optimization

The superior performance of the optimized three port $3\,\mathrm{dB}$ coupler is a direct consequence of our systematic approach to the design process. First, by establishing a physically-motivated starting structure, we ensured that the device was sufficiently large and the inputs and outputs of the system are appropriately placed such that a high coupling efficiency exists within the design space. Next, by running an optimization with a coarse parameterization and without constraints, we are able to rapidly improve the device performance by changing its general shape. Because the set of design parameters is smaller, the corresponding design space is less complex and as a result we find falling into undesirable local optima less likely. Finally, using the result of the coarse optimization as a starting point, we run a third refinement optimization in which we impose various constraints (e.g. fabrication constraints). During this final optimization, we leverage a larger set of design variables in order to allow the structure to take on a more sophisticated shape. We refer to this whole process as *hierarchical design and optimization*.

Chapter 5

Application to Silicon Photonics

In recent years, there has been a trend towards designing relatively large-scale silicon photonic systems. For example, integrated beam steering for LiDAR [58–60], optical switches [61–64], and optical computing platforms [65–67] require thousands of silicon photonic components. When cascading large numbers of components together as is common in these applications, the efficiency of each individual component becomes increasingly important. Typical silicon photonic components are often too lossy or, in the case of devices which rely on adiabatic transitions, are too large for the given application. This situation is further exacerbated when the platform demands a larger operational bandwidth or a greater insensitivity to fabrication variations.

The hierarchical approach to photonic design that we introduced in Chapter 4 presents a promising strategy for designing silicon photonic components which demand such a high level of performance. In this chapter, we will apply these methods to a variety of devices which form the backbone for many silicon photonic systems. In particular, we will apply our methods in order to optimize broadband four port $3\,\mathrm{dB}$ couplers, fabrication tolerant waveguide crossings, and efficient grating couplers. In each case, we will develop a physically-inspired topology for the device, parameterize it, and run coarse and refinement optimizations. The end results of these optimizations are high performance components which can be fabricated with deep UV lithography.

In the majority of these examples, we employ the silicon photonic platform depicted in Figure 4.2 which consists of a $220\,\mathrm{nm}$ top silicon layer and is patterned with a $110\,\mathrm{nm}$ shallow etch and a $220\,\mathrm{nm}$ deep etch. The devices are clad top and bottom with silicon dioxide. This partially-etched platform has a few advantages. First, it is compatible with lower-loss ridge waveguides. Next, the thicknesses are commonly employed in the nanophotonics community. Finally, the lower index contrast of this partially-etch platform makes for a harder optimization problem, which

highlights the power of our tools and methodology[1].

5.1 Four Port 3 dB Coupler

A key component in many integrated photonic systems is the four port $3\,\mathrm{dB}$ coupler (or 2×2 splitter). This four port splitter takes light from either of its two input waveguides and splits the optical power evenly between its two ouput waveguides. Because the device has an inherent asymmetric operation, it is more difficult to design compared to the three port optical splitter that we used as a guiding example in chapter 4.

Just as with its three port counterpart, multimode interference effects provide us with a good topology for the four port $3\,\mathrm{dB}$ coupler. Specifically, our device topology consists of a larger multimode waveguide with two symmetrically situated narrower waveguides connected at its input and two symmetrically situated waveguides connected at its output. Light incident on the multimode waveguide from one of the two input waveguides will excite higher order modes which copropagate; based on the relative phases of these higher order modes, the field can be effectively split between two output waveguides [51].

For the purpose of choosing the approximate dimensions of our initial design, it is convenient to consider the first three modes of the multimode waveguide section of the device (so long as the device is not too wide, the majority of the input power will reside in these first three modes). Assuming the device lies along the z-axis, the electric field inside the multimode waveguide is given approximately by

$$\begin{aligned}\mathbf{E}(x,y,z) &= \mathbf{E}_0(x,y)e^{-ik_0 z} + \mathbf{E}_1(x,y)e^{-ik_1 z} + \mathbf{E}_2(x,y)e^{-ik_2 z} + \dots \\ &= e^{-ik_0 z}\left[\mathbf{E}_0(x,y) + \mathbf{E}_1(x,y)e^{i(k_0-k_1)z} + \mathbf{E}_2(x,y)e^{i(k_0-k_2)z} + \dots\right] \end{aligned} \tag{5.1}$$

where $\mathbf{E}_0(x,y)$, $\mathbf{E}_1(x,y)$, and $\mathbf{E}_2(x,y)$ are the lateral field profiles of the first three modes depicted in Figure 5.1 and k_0, k_1, and k_2 are their effective wavenumbers. It is important to note that the fundamental and second order modes have even symmetry while the first order mode has odd symmetry. If we add the fundamental mode and second order mode with the correct phase, we produce a symmetric field with two positive lobes. If we add this resulting field to the first order (odd) mode of the multimode waveguide, we can concentrate the optical power in the top half of the multimode waveguide. This field combination roughly corresponds to the input of the multimode waveguide which is excited by one of the input waveguides as depicted in the left half of Figure 5.1.

[1]This is in contrast to some past work on optimizing nanophotonics has focused on designing devices made from fully-etched silicon with air cladding [12, 47] which maximizes the refractive index contrast.

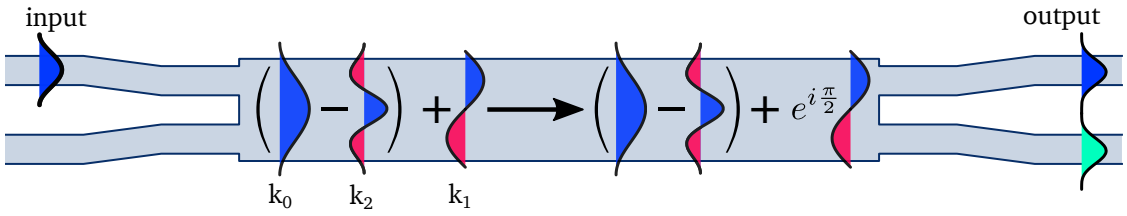

Figure 5.1: Graphical explanation of the basic operating principle of the four port $3\,\mathrm{dB}$ coupler. Light from either of the two input waveguides primarily excites the first three modes of a multimode waveguide. These modes copropagate and their relative phases change. After traveling a certain distance, the relative phases of these modes change such that the field distribution in the waveguide splits evenly towards the outer edges of the multimode waveguide. Truncating the multimode waveguide at this length and connecting it to two output waveguides results in splitting of the input light between the two output waveguides.

As the three modes copropagate, the relative phases between the modes will evolve. Based on Equation 5.1, the relative phase between the fundamental and second order mode is $\varphi_{20} = (k_0 - k_2)\, z$ and the phase between the fundamental and first order mode is $\varphi_{10} = (k_0 - k_1)\, z$. Notice that if we were to add the even symmetry field profile generated by adding the fundamental and second order mode to the odd-symmetry first order mode with a $\pi/2$ phase difference, the resulting distribution of power inside the multimode waveguide will split evenly away from its center. This condition occurs when the waveguide has a length such that $\varphi_{20} = 2\pi p$ and $\varphi_{10} = (2q+1)\pi/2$ where p and q are integers. These relative phase shifts correspond to waveguide lengths of

$$\begin{aligned} 2\pi p = (k_0 - k_2)\, L_{2\pi} \Rightarrow L_{2\pi} &= \frac{2\pi p}{k_0 - k_2} \\ \Rightarrow L_{2\pi} &= \frac{p\lambda}{n_0 - n_2} \end{aligned} \tag{5.2}$$

and

$$\begin{aligned} (2q+1)\frac{\pi}{2} = (k_0 - k_1)\, L_{\pi/2} \Rightarrow L_{\pi/2} &= \frac{(2q+1)\frac{\pi}{2}}{k_0 - k_1} \\ \Rightarrow L_{\pi/2} &= \frac{(2q+1)\lambda}{4(n_0 - n_1)} \end{aligned} \tag{5.3}$$

where n_0, n_1, and n_2 are the effective indices of the three modes and λ is the free-space wavelength. Ideally, we would like to find values for p and q such that

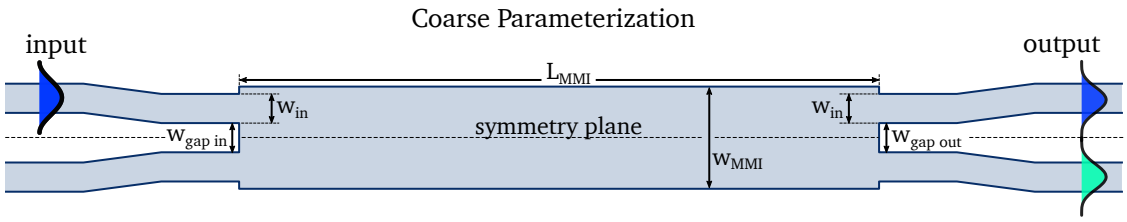

Figure 5.2: Coarse parameterization of the four port $3\,\mathrm{dB}$ coupler. In total, the splitter structure is defined using six design parameters during the coarse optimization. These paramters control the input and output waveguide parameters as well as the width and length of the multimode waveguide section.

$L_{2\pi} = L_{\pi/2}$. Because the effective indices are typically irrational numbers, however, this condition is not strictly possible to satisfy. Fortunately, we can choose p and q such that $L_{2\pi} \approx L_{\pi/2}$. The result approximate length is the length of the multimode waveguide which makes up our four port splitter. At the output of the multimode waveguide, we place two symmetrically situated single mode waveguides. The evenly split fields within the multimode waveguide couple into these two output waveguides.

Our ultimate goal is to design a four port splitter which maintains high coupling efficiency and even splitting over the whole O-band. As such, we choose initial dimensions which are suited to operation at $1310\,\mathrm{nm}$. For the purpose of this optimization, the splitter is defined in a $220\,\mathrm{nm}$ thick silicon with a $110\,\mathrm{nm}$ etch. The input and output waveguides are chosen to be $500\,\mathrm{nm}$ wide and are separated by $500\,\mathrm{nm}$. The multimode waveguide section of the device is chosen to be $1.75\,\mu\mathrm{m}$ which yields a device with a modest footprint which can be accurately simulated in a reasonable amount of time. Based on these parameters and the expressions given in Equations (5.2) and (5.3), we find that $L_{2\pi} \approx L_{\pi/2} \approx 15.8\,\mu\mathrm{m}$.

The starting structure produced by this simple analysis and its simulated electric field is plotted in the top of Figure 5.5. Although the fields are fairly effectively coupled into the output waveguides, there is visible imbalance in the fraction of power exiting in the top and bottom waveguides. This is likely a result of a few different factors. First, our analysis considered only the first three modes of the multimode waveguide section. In reality, higher order modes exist and will contribute to the fields in the device in a way that we did not account for. Furthermore, our analysis largely neglected the exact size and position of the inputs and outputs, both of which have a large impact on the overall performance on the device. Despite these non-idealities which hinder the performance of the starting structure, we nonetheless can be reasonably confident that the device is of an appropriate size and shape that will enable efficient $3\,\mathrm{dB}$ splitting.

In order to improve this initial design, adherent to our hierarchical approach, we run a coarse optimization of the structure. In this coarse optimization, we de-

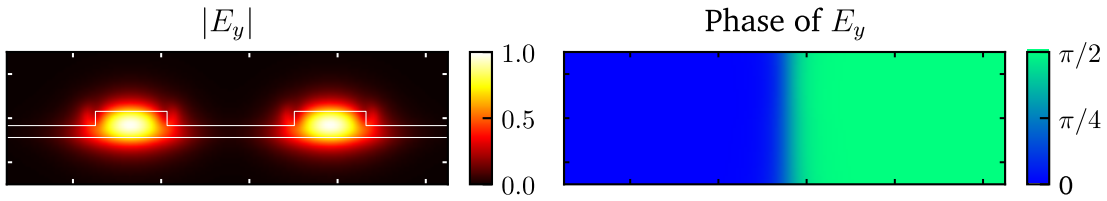

Figure 5.3: Plots of magnitude and phase of the reference E_y used in the calculation of the figure of merit for the four port $3\,\mathrm{dB}$ coupler.

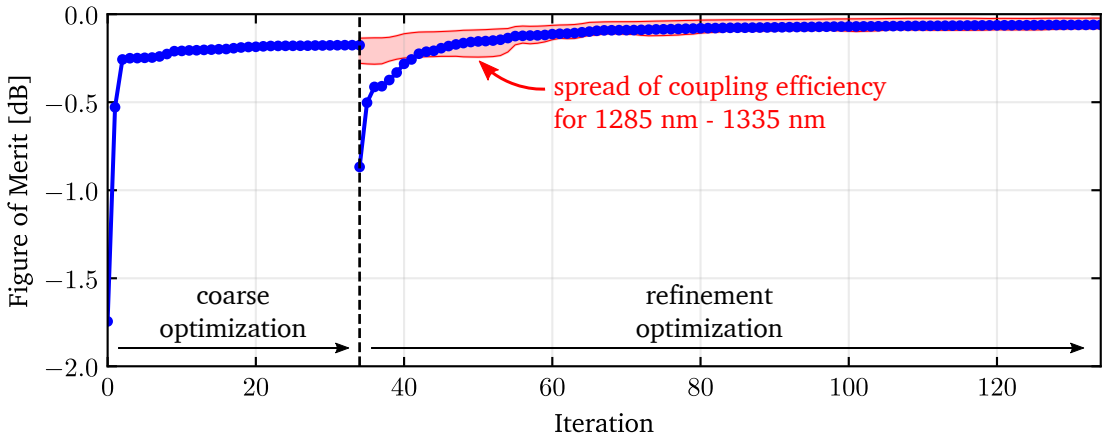

Figure 5.4: Plot of the figure of merit vs iteration for the four port $3\,\mathrm{dB}$ coupler. The first 35 iterations plotted correspond to a coarse optimization of the structure while the latter 100 iterations correspond to a multiwavelength refinement optimization. The red shaded region represents the range of coupling efficiencies computed over the $1285\,\mathrm{nm}$ to $1335\,\mathrm{nm}$ wavelength range. As desired, the optimization both improves the average coupling efficiency over this band, but also narrows the range of efficiencies, indicating that the device has been made more broadband.

fine the structure using the six degrees of freedom shown in Figure 5.2. In order to ensure that the device works for both input ports identically, we force it to be symmetric about the xz plane. The figure of merit for this optimization is the coupling efficiency into the desired output field at $1310\,\mathrm{nm}$ which consists of the sum of the first two super modes of the two-output-waveguide system with a relative $\pi/2$ phase shift applied between the modes (depicted in Figure 5.3). This coupling efficiency is approximated[2] using the mode match expression given by Equation (A.25) derived in Appendix A. Finally, all simulations during the coarse optimization make

[2]It is important to note that the mode match expression derived in chapter A is not an exact measure of coupling efficiency when the reference fields consist of more than a single mode (i.e., one with multiple wave vectors). Nonetheless, the mode match is still a good approximation of the overlap of the simulated and desired fields.

use of a $40\,\mathrm{nm}$ grid spacing which speeds up the process significantly at the expense of a minor reduction in simulation accuracy. As with our three port splitter, the structure is optimized using the L-BFGS-B minimization algorithm.

The coarse optimization is allowed to run until the figure of merit has converged (defined in this case as when the figure of merit changes by less than 1×10^{-5}) which takes 35 iterations. The evolution of the figure of merit during the coarse is plotted in Figure 5.4. Due to the relatively small number of parameters, the coupling efficiency of the structure increases rapidly from it's initial value of $-1.75\,\mathrm{dB}$ to nearly $-0.25\,\mathrm{dB}$ in only two iterations. During the remainder of the optimization, the device is fine tuned until it achieves final coupling efficiency of only $-0.176\,\mathrm{dB}$.

The result of this optimized structure is shown in the middle of Figure 5.5. Visually, it is evident that the majority of the optical power leaves the device in the output waveguides and the field is evenly split between the two output waveguides.

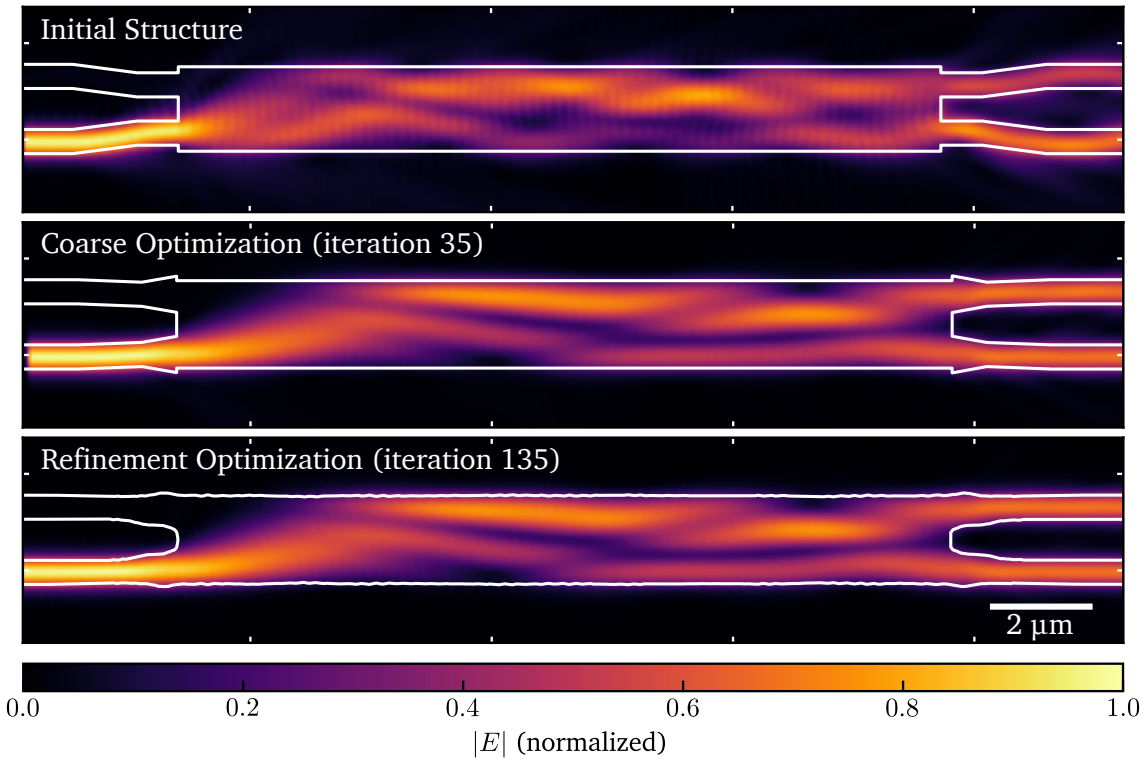

Figure 5.5: Plot of the electric fields of the initial, coarse optimized, and refinement optimized four port $3\,\mathrm{dB}$ coupler. The outline of the optimized structure is overlaid on top of the field plots, showing the evolution of the device as a result of the optimization process.

With the coarse optimization complete, we next move on to the refinement optimization. This refinement optimization has two goals. First, it should produce an optimal structure that does not contain any sharp corners or small features that cannot be resolved using deep UV lithography. Second, it should ensure the the

power leaving the two output ports is as balanced (close to 50/50) as possible over the full O-band.

This latter goal, in particular, requires some additional attention. Naively, we might expect that simply maximizing the coupling efficiency at a few wavelengths spread over the O-band would be enough to ensure even splitting. Unfortunately, this turns out to not be the case. The reason for this lies in the fact that we will inherently be operating near a maximum in the mode match equation which occurs when the power is exactly equally split between the two output waveguides. Because we are near a maximum, modest changes in the relative power in the two waveguides lead to only small changes in the overlap integral. In practice, the relative power in the two outputs needs to change by more than a percent or two before the mode match figure of merit starts to decrease significantly.

One way to get around this issue is to apply a penalty to the figure of merit which increase when the outputs become imbalanced. A simple form of this penalty function is to calculate the overlap with the set of fields which are orthogonal to the desired set of output fields. This function will have a minimum at the point in the design space where the overlap with the desired fields is maximized and the outputs are balanced. Deviating from this point causes the function to increase and thus more heavily penalize the figure of merit. With this in mind, our complete figure of merit for the refinement optimization is

$$F(\mathbf{E},\mathbf{H},\vec{p}) = \frac{1}{N_\lambda}\sum_\lambda \left[\eta(\mathbf{E}_\lambda,\mathbf{H}_\lambda) - \alpha\eta_{\text{ortho}}(\mathbf{E}_\lambda,\mathbf{H}_\lambda) - \mathcal{P}_{\text{roc}}(\vec{p})\right] \tag{5.4}$$

where η is the mode match calculated with respect to the desired and orthogonal fields, $\mathcal{P}_{\text{roc}}(\vec{p})$ is the radius of curvature penalty function, λ denotes the simulation wavelength, N_λ is the total number of wavelengths, and α is sets the weight of the penalty function. In this case, we use $\alpha = 5\times10^3$ (in general the exact value does not significantly impact the final result). It is interesting to note that this average figure of merit has similar behavior to a minimax figure of merit when working high efficiency devices since the maximum coupling efficiency possible is equal to one.

In this optimization, we will cooptimize our splitter at three different wavelengths spanning a $50\,\text{nm}$ range: $1285\,\text{nm}$, $1310\,\text{nm}$, and $1335\,\text{nm}$ (we could optimize over a larger bandwidth, however this can lead to difficulties in maintaining a balanced 50/50 output). For all three wavelengths, the structure is identical and a maximum $120\,\text{nm}$ radius of curvature is enforced. The forward and adjoint simulations of the structure are carried out using our 3D FDTD solver with a simulation domain size of $23.8\,\mu\text{m}$ by $5.25\,\mu\text{m}$ by $2\,\mu\text{m}$ and a uniform grid spacing of $30\,\text{nm}$. Given the size of the simulation domain, this grid spacing gives a good trade off between speed and accuracy. The final results presented later in this section are produced using higher resolution simulations which have a $20\,\text{nm}$ grid spacing to ensure higher accuracy.

The optimization is run for a total of 100 iterations (which took a little over

2 days on 256 cores of our Intel Xeon based cluster) at which point the figure of merit ceases to increase appreciably. The evolution of the figure of merit during the optimization is shown in the right portion of Figure 5.4. Initially, the figure of merit is significantly lower than the final coarse optimization result. This is primarily a consequence of the orthogonal field mode match constraint which reduces the figure of merit when the power in the two output waveguides is not balanced. The red shaded region shows the range of coupling efficiencies computed for the three wavelengths which make up the figure of merit, indicating that the actual coupling efficiency still starts out reasonably high. As desired, the figure of merit rapidly improves during the optimization. This improvement in the figure of merit directly corresponds to an improvement in the broadband performance of the device which is reflected by red shaded region which shifts towards $0\,\mathrm{dB}$ and narrows.

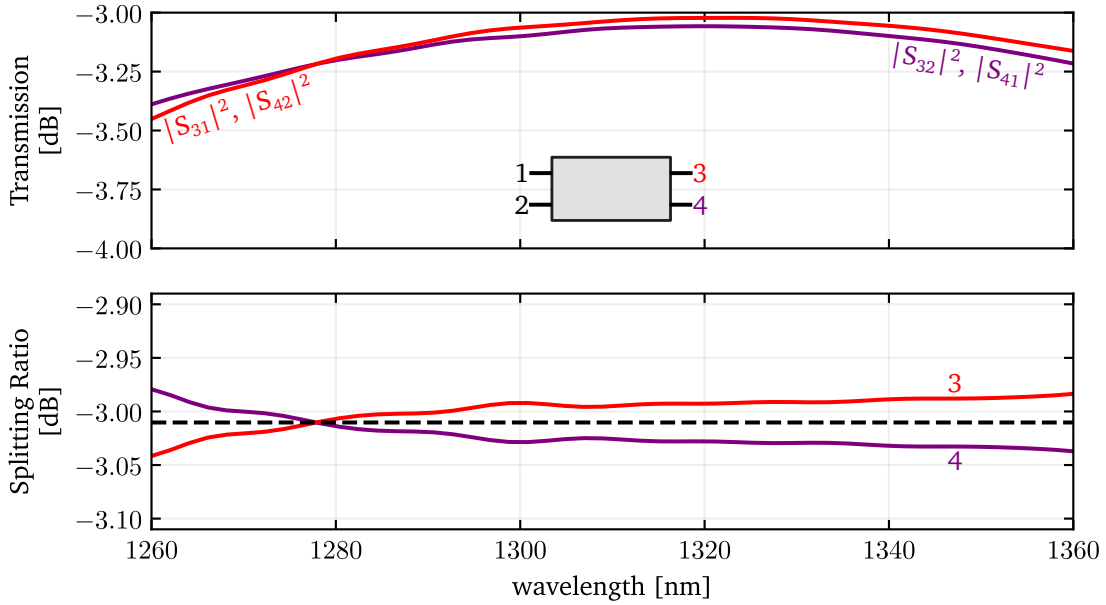

Figure 5.6: Plot of the optimized four port $3\,\mathrm{dB}$ coupler performance. In the top, the transmission into the two output is plotted versus wavelength. In the bottom, the splitting ratio (defined as the fraction of output power in each output waveguide) is plotted versus wavelength. The black dashed line corresponds to 50%, which is the desired splitting ratio.

The performance of the final optimized structure is shown in Figure 5.6. As desired, the total excess loss remains low over the full O-band, reaching a maximum value of $0.41\,\mathrm{dB}$ at $1260\,\mathrm{nm}$. This is further reflected in a $0.1\,\mathrm{dB}$ bandwidth of $61\,\mathrm{nm}$. In addition to low excess loss, a splitting ratio very close to 50% is achieved over a broad range of wavelengths as shown in the bottom half of Figure 5.6. Specifically, the relative fraction of output power in either of the two waveguides deviates from 50% by no more than a few hundredths of a decibel over the full O-band.

These results are superior to previously reported on multimode interference and directional coupler designs [16,68–71] and are consistent with adabatic-transition-based designs [72–76] which are typically an order of magnitude (or more) larger than our design. Such high performance makes these optimized 2×2 splitters ideal for large scale integrated photonics applications which demand splitters that are low loss without sacrificing large amounts of die area.

5.2 Fabrication-Tolerant Waveguide Crossing

One of the great advantages of optical waveguides compared to conventional electrical interconnects is that they can intersect each other with minimal crosstalk. This functionality turns out to be essential to many silicon photonic integrated circuits which are only able to leverage a single layer of silicon. In large systems in particular, a single waveguide may have to pass through many waveguide crossings. In such cases, it is essential that the waveguide crossings be very low loss over the relevant bandwidth, even in the presence of any fabrication errors.

In this section, we will design a waveguide crossing which is tolerant to fabrication variations. In particular, we will optimize a crossing which maintains exceptionally low insertion loss over the full O-band even when subject to $\pm10\%$ thickness variations in the top silicon layer. Our starting point for this optimization is a standard waveguide crossing design based on multimode interference effects. The topology we choose, depicted in Figure 5.7, consists of a narrow input waveguide which abruptly tapers into a large multimode waveguide. Light in the fundamental mode of the input waveguide is coupled primarily into the first two even modes of the multimode waveguide. These two modes co-propagate and their relative phase evolves according to Equation 4.2. Over a distance

$$\Delta z = \frac{\lambda}{n_0 - n_2} \tag{5.5}$$

the relative phase between the two modes becomes equal to 2π. At this distance, the modes interfere such that the field focuses towards the center of the multimode waveguide (effectively imaging the input waveguide mode). If we make the total length of the crossing waveguide (including tapers) twice this length, i.e.

$$L_{\text{crossing}} = \frac{2\lambda}{n_0 - n_2} \tag{5.6}$$

and place a second narrow waveguide at its output, the field will focus a second time and couple effectively into the output waveguide. At the first focal point, we can place an identical perpendicular intersecting waveguide which will minimally influence the propagation of light in the horizontal multimode waveguide as a result of the focusing effect (so long as the tapers are larger than half of the multimode

waveguide width). With a bit of fine tuning, this configuration produces a very effective waveguide crossing with minimal loss and crosstalk.

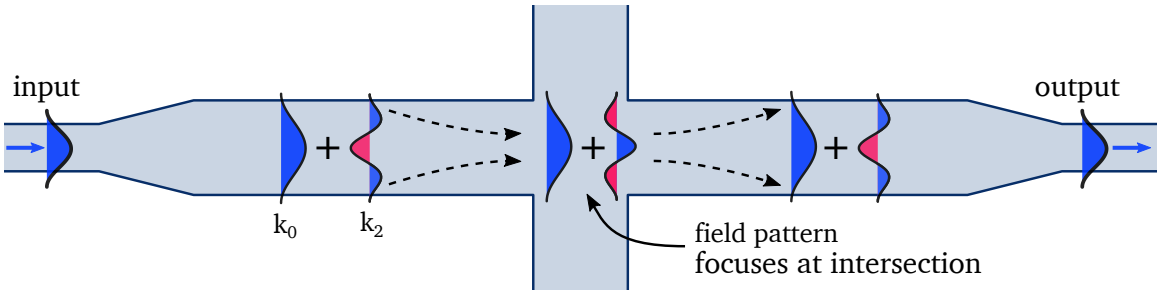

Figure 5.7: Graphical explanation of the basic operating principle of a waveguide crossing. Light from the fundamental mode of an input waveguide is coupled primarily into the first two even modes of a wider multimode waveguide. These two modes coopropagate and interfere to form a focusing intensity pattern near the center of the crossing waveguide. This focusing effect allows the light to diffract through an intersecting waveguide with minimal disruption.

As with the previous examples, we will be working with a nominal silicon thickness of $220\,\mathrm{nm}$ and the crossing will be defined using a $110\,\mathrm{nm}$ partial etch. For our starting structure, we use a $1.7\,\mu\mathrm{m}$ wide multimode waveguide which, based on Equation (5.6), corresponds to a crossing length of $15\,\mu\mathrm{m}$. In order to ensure that the focusing distance within the crossing waveguide is sufficiently long, we make the input and output taper lengths equal to the width of the multimode waveguide ($1.7\,\mu\mathrm{m}$). These tapers connect the multimode crossing waveguide to the input and output waveguides which are $500\,\mathrm{nm}$ in order to remain consistent with the previous examples. The simulated fields of this starting structure are shown in the top of Figure 5.11. As desired, the power is effectively coupled from the input to output waveguide with little visible light coupling into the perpendicular waveguide.

Unfortunately, it turns out that this standard waveguide crossing design is neither very broadband nor insensitive to fabrication variations. Our goal in running the coarse and refinement optimizations is thus to improve its performance when subject to such variations, and in particular, significant silicon thickness variations. In order to do so, our figure of merit for the coarse optimization will be the average (mode matched) transmission into the output waveguide at $1310\,\mathrm{nm}$ for three different top silicon thicknesses corresponding to the desired thickness, a 10% increase in thickness, and a 10% decrease in thickness:

$$F(\mathbf{E},\mathbf{H}) = \frac{1}{3}\left[\eta_{0\%}(\mathbf{E},\mathbf{H}) + \eta_{+10\%}(\mathbf{E},\mathbf{H}) + \eta_{-10\%}(\mathbf{E},\mathbf{H})\right] \tag{5.7}$$

These thickness variations are depicted in Figure 5.8. For each thickness, we define

the crossing structure using the same $110\,\mathrm{nm}$ partial etch depth [3].

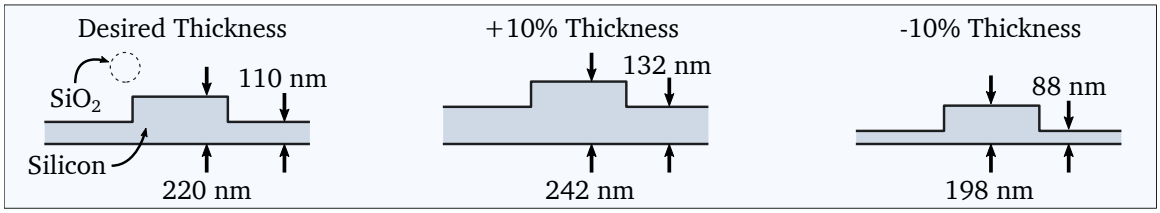

Figure 5.8: Diagrams of the thickness variations included in the optimization of a waveguide crossing. The three cases depicted correspond to 0%, +10%, and -10% thickness variations in the top silicon thickness. In all three cases, the etch depth is kept fixed at $110\,\mathrm{nm}$.

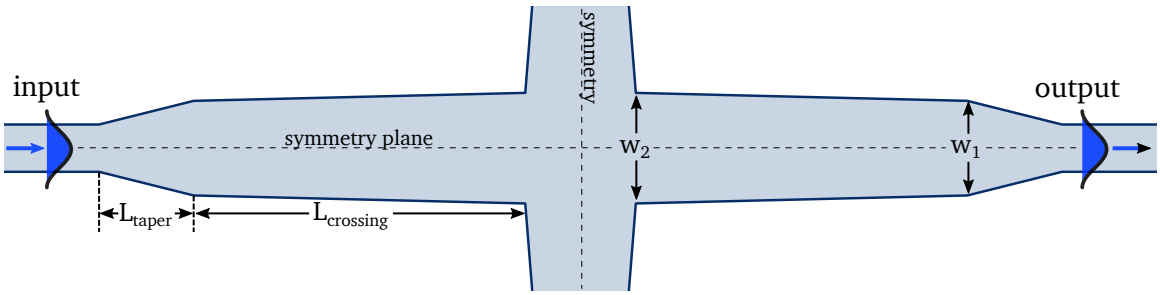

Figure 5.9: Parameterization of the waveguide crossing for the coarse optimization. The design parameters are the taper length, crossing length, final widths of the taper, and width of the crossing waveguide at the waveguide intersection. The structure is forced to be symmetric about both the x and y axes.

For the purpose of the coarse optimization, we parameterize the structure using only four degrees of freedom. These design parameters, as shown in Figure 5.9, are the length of the input and output tapers, the length of the crossing waveguide, and width of the crossing waveguide at the end of the taper, and the width of the crossing waveguide at the waveguide intersection.

The structure itself is simulated on a uniform $40\,\mathrm{nm}$ grid for the sake of speed. The optimization is performed using the L-BFGS-B minimization algorithm and is allowed to iterate until the figure of merit changes by less than 10^{-6}. The progression of the figure of merit is plotted in Figure 5.10. The blue trace shows the value of the figure of merit while the red shaded region shows the range of transmissions calculated for the $\pm 10\%$ thickness variations. Despite having only four degrees of freedom to manipulate, the optimization is able to rapidly improve the average transmission to $-0.0269\,\mathrm{dB}$ and the total variation is drastically reduced to only $0.0102\,\mathrm{dB}$ at $1310\,\mathrm{nm}$.

[3] in practice, we have found that the performance is much less sensitive to the etch depth than the total silicon thickness

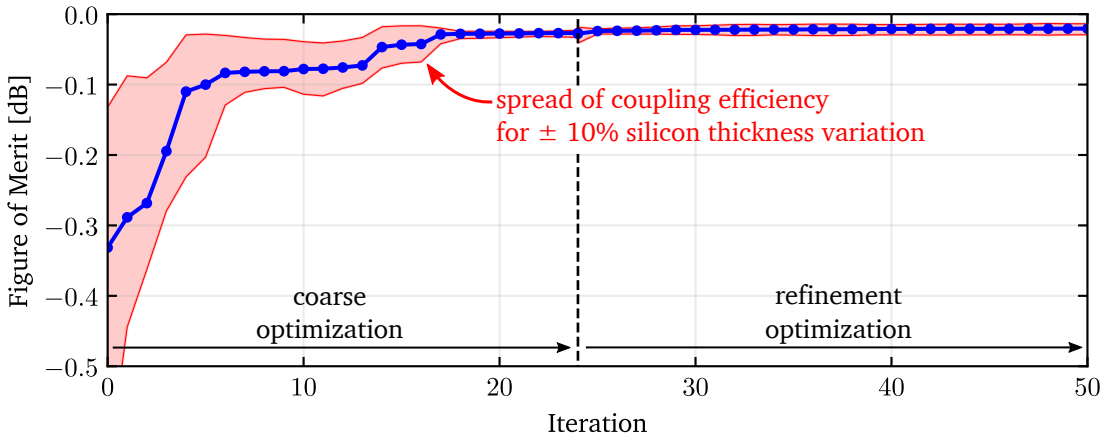

Figure 5.10: Plot of the figure of merit vs iteration for the fabrication tolerant waveguide crossing. The first 24 iterations plotted correspond to a coarse optimization of the structure while the latter 26 iterations correspond to the refinement optimization. The red shaded region represents the range of coupling efficiencies computed for the device when subject to $\pm 10\%$ thickness variations. As desired, the optimization both improves the average coupling efficiency over this band, but also narrows the range of efficiencies, indicating that the device has been made more tolerant to fabrication variations.

The structure produced by this coarse optimization and its simulated electric field is shown in the middle of Figure 5.11. Interestingly, the optimized structure has a bow tie shape with a much longer input taper. From the standpoint of reducing sensitivity to fabrication variations, a structure with more gradual transitions is intuitive.

Although this coarsely optimized structure already achieves exceptional performance, there is nevertheless room for improvement. Furthermore, the coarse result contains sharp corners that need to be corrected. As with our previous examples, this is easily accomplished using a refinement optimization. To initialize this refinement optimization, we use the coarse optimization result with rounded corners and an increased number of vertices (one point every $40\,\mathrm{nm}$). The positions of these vertices are the design parameters of the optimization. The figure of merit of the optimization is the same as the coarse optimization with the addition of a radius of curvature penalty term, i.e.,

$$F(\mathbf{E},\mathbf{H},\vec{p}) = \frac{1}{3}\left[\eta_{0\%}(\mathbf{E},\mathbf{H}) + \eta_{+10\%}(\mathbf{E},\mathbf{H}) + \eta_{-10\%}(\mathbf{E},\mathbf{H})\right] - \mathcal{P}_{\mathrm{roc}}(\vec{p}) \tag{5.8}$$

where $\mathcal{P}_{\mathrm{roc}}(\vec{p})$ is derived in Appendix A.2.2. Finally, all simulations in this final optimization use a $22\,\mathrm{nm}$ grid spacing to ensure that the results are accurate. This optimization is run until the value of this figure of merit decreases by less than 10^{-6}. The change in the figure of merit during this phase of the optimization is plotted in

the right half Figure 5.10 and the final result is shown at the bottom of Figure 5.11. Despite the fact that we modified the coarse optimized structure by rounding off its corners, the initial drop in performance at the start of the refinement optimization is very small. This is explained by the fact that our coarse optimized design is designed to be very insensitive to thickness variations, and we expect this insensitivity to apply to other small non-thickness-related changes to the structure as well. Furthermore, the figure of merit appears to change minimally during the coarse of the refinement optimization. Although seemingly small, the minimum coupling efficiency improves from $-0.0410\,\mathrm{dB}$ to $-0.0292\,\mathrm{dB}$ which is approximately a 30% improvement. This means that 30% more crossings can be used along a path for the same amount of loss which is not insignificant.

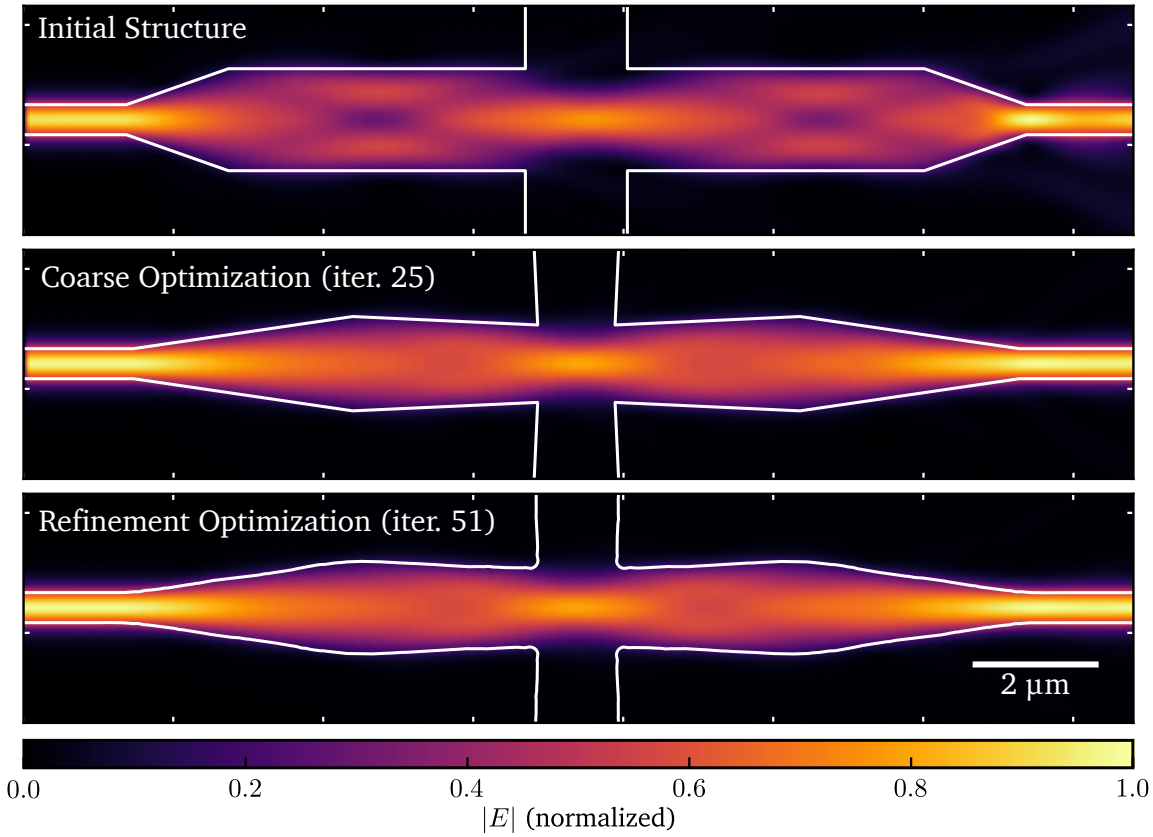

Figure 5.11: Plots of the (top) starting waveguide crossing structure, (middle) result of the coarse optimization, and (bottom) result of the refinement optimization. In all three cases, the fields are shown for $1310\,\mathrm{nm}$. In all three plots, the magnitude of the electric field taken from a slice running through the center of the device is overlayed with an outline of the structure.

Compared to the initial maximum loss of more than $0.5\,\mathrm{dB}$, the final optimized result is both highly efficient and exceptionally insensitive to thickness variations.

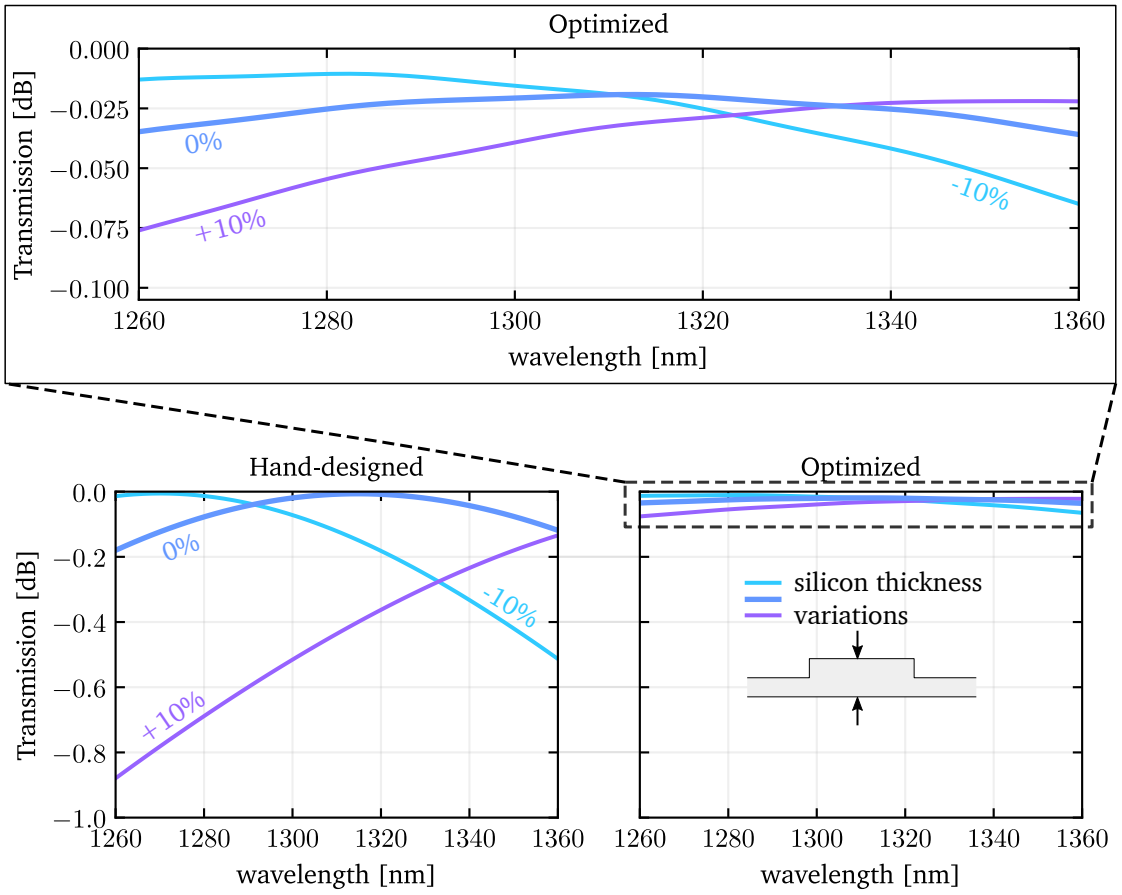

Figure 5.12: Plots of the performance of the optimized waveguide crossing (right) compared to a "typical" waveguide crossing (left). When subject to $\pm 10\%$ SOI thickness variations, the optimized design maintains a maximum loss across the O-band which is nearly 10 times smaller than a typical waveguide crossing.

This improved tolerance to variations is a direct consequence of the improved bandwidth of the optimized structure which is made apparent in Figure 5.12. In the bottom left, the transmission of a typical hand-designed waveguide crossing is plotted versus wavelength. A crossing with the desired silicon thickness achieves a high peak coupling efficiency above $-0.01\,\mathrm{dB}$, however this performance quickly drops to nearly $-0.2\,\mathrm{dB}$ at the edges of the O-band. This limited bandwidth translates to significant reductions in efficiency when the silicon thickness varies; when subject to a ten percent increase in thickness, the spectrum of the crossing shifts significantly causing the transmission to drop below $-0.8\,\mathrm{dB}$ at $1260\,\mathrm{nm}$. Even at the center wavelength of $1310\,\mathrm{nm}$, the transmission drops below $-0.4\,\mathrm{dB}$ when subject to a ten percent thickness increase, eliminating its usefulness.

Our optimized design, has significantly improved performance across the O-

band, both with and without thickness variations. The transmission of our optimized design is plotted in the bottom right of Figure 5.12 and shares the same y axis as the hand-designed crossing plotted on the left. Immediately apparent, the sensitivity to wavelength and silicon thickness variations is reduced by roughly an order of magnitude. The plot in the top of Figure 5.12 shows a zoomed in version of the transmission of the optimized structure. Across the entire O-band when subject to plus or minus ten percent variations in the silicon thickness, our optimized design achieves better than $-0.075\,\mathrm{dB}$. At the ideal thickness, the optimized design achieves better $-0.036\,\mathrm{dB}$ over the entire O-band.

This broadband behavior and insensitivity to fabrication variations comes only at the expense of a small reduction in peak coupling efficiency ($-0.0065\,\mathrm{dB}$ to $-0.019\,\mathrm{dB}$). Compared to previously published results, our optimized crossing has either comparable or higher peak coupling efficiency and significantly higher bandwidth [77–83]. To our knowledge, our optimized crossing has achieves better performance when subject to variations than any previously reported design.

It is worth noting that this record result is effectively achieved using only 4 design parameters. This is at odds with much of the work on shape and topology optimization in nanophotonics which often associate larger numbers of design variables with an improved ability to design devices. We attribute this discrepancy to the fact that we carefully chose a physically-motivated topology. By doing so, the general shape of the device was already very conducive to achieving the desired waveguide crossing behavior. Only a small set of design parameters which preserve this general topology was therefore needed to fine tune the device. This highlights the strength of our hierarchical approach to design and optimization.

5.3 Single Polarization Grating Coupler

One of the most important components in silicon photonics is the optical coupler which enables us to couple light into and out of silicon photonic chips. Optical couplers used in silicon photonics typically come in one of two varieties: edge couplers and vertical grating couplers. Grating couplers in particular present a number of advantages over alternative coupling methods [84] by providing a flexible means of interfacing high-index-contrast integrated optical devices with the outside world.

In general, a grating coupler consists of a high index waveguiding slab with either partially- or fully-etched periodic corrugations which cause light to scatter out of the slab at a desired angle. By correctly choosing the positions and sizes of these corrugations, we can shape the beam of light generated by the grating such that it is well matched to an optical fiber (or some other desired output mode). Due to the inherently large number of degrees of freedom in a grating coupler and

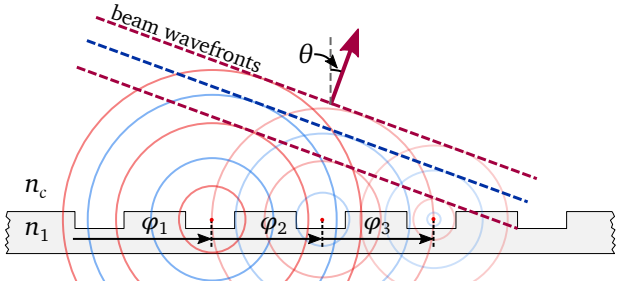

Figure 5.13: Phased array interpretation of grating couplers. A grating coupler can be thought of as a collection of independent scatterers which give rise to waves propagating away from the grating. The relative phases of these scatterers is given by the phase of a guided mode propagating within the grating. The scattered waves constructively interfere to form a beam propagating at an angle.

the degree to which they are coupled together[4], designing grating couplers using more traditional methods has proven difficult. Fortunately, hierarchical design and optimization presents us with a new and highly effective strategy for designing efficient grating couplers.

5.3.1 The Foundations of a Good Grating Coupler

Before diving into grating optimization, it is important that we understand how grating couplers work and what drives their overall performance. Typically, the goal of a grating coupler is to couple as much light as possible into a desired output mode. In most cases, this output mode will be an optical fiber, although in some applications, we may wish to couple into some other mode or field. In either case, the *coupling efficiency* of the grating coupler depends on four contributing factors: mode match, directionality, back reflection, and incomplete scattering. Mathematically, this can be summarized as

$$\eta_{\text{grating}} = \eta_{\text{MM}} D(1-R)(1-T) \tag{5.9}$$

where η_{MM} is the mode match efficiency, D is the directionality, R is the back reflection of the grating coupler, and T is the fraction of power not scattered by the grating. In this section, we will review these contributions to grating coupler efficiency and discuss how we can control them. This knowledge serves as the foundation of our hierarchical approach to grating coupler design.

Grating Scattering Angle

Perhaps the most important characteristic of a grating coupler is the angle at which it scatters light. If this angle is not matched with the desired output mode, then coupling will be severely diminished.

To understand how grating couplers produce coherent beams which propagate at a well defined angle, we can think of the grating as a phased array of independent scatterers as depicted in Figure 5.13. As a guided wave within the grating encounters an etched slot, the discontinuity gives rise to a scattered wave which we can model as emanating from the center of the slot. The phase of this scattered wave is equal to the phase of the guided wave at the time it encounters the slot. Similarly, the amplitude of that wave will be proportional to the amplitude of the guided wave. Multiple scattered waves produced by a series of periodic etched slots propagate through the cladding medium and interfere to form a coherent beam with wavefronts that move at a well-defined angle (indicated by the dashed lines in Figure 5.13).

The angle of the generated beam depends on the relative phase that a guided wave accumulates when moving between two slots in the grating. An expression for this phase can be derived based on simple geometric arguments (see Appendix D.1). Assuming the slots are periodic, then the phase and emission angle are related by

$$\varphi_{\text{in}}(\Lambda) - 2\pi m = k_c \Lambda \sin\theta \tag{5.10}$$

where φ_{in} is the phase of the wave as it propagates through one period of the grating, k_c is the wavenumber of the cladding medium, Λ is the periodicity of the grating, θ is the scattering angle, and m is the integer diffraction order (which we typically choose to be ± 1). The phase φ_{in} is a function of the period of the grating as well as the effective indices of the slabs which make up the grating. Based on the exact topology of the grating coupler, we can use this expression to select a period of the grating which will produce a beam at the desired angle. This is a key step in grating coupler design.

It is important to note that grating couplers are not strictly characterized by a single angle. Because grating couplers couple light into finite width beams, they actually have an angular *spectrum* which is centered around the scattering angle of the grating coupler. In most cases, we attempt to couple light into single mode optical fibers which are relatively wide compared to the wavelength of light. As a result, corresponding angular spectrum is quite narrow (and as a result coupling efficiency will be very sensitive to fiber angle). In contrast, for some applications, it may be desirable to couple light into a wider angular spectrum. In such cases, it is important that we design our grating coupler to couple light into a narrower beam.

[4] In other words, we cannot easily break up the design problem into many smaller design problems with only one or two variables

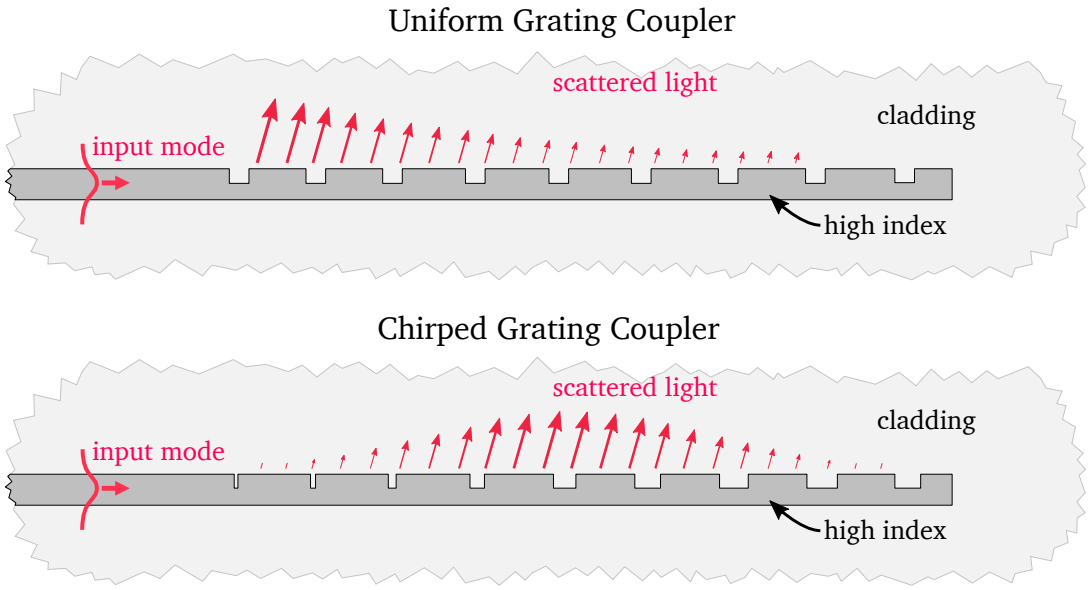

Figure 5.14: Diagrams of uniform and chirped partially-etched grating couplers. In both cases, a waveguide has slots etched into its top surface which scatter light. By engineering the periodicity of these slots, the scattered light constructively interferes to form a beam propagating at a desired angle. If the size of these etched slots is uniform (top), then the intensity of the scattered light will decay along the length of the grating coupler. If, on the other hand, the slot width and period is chirped, the scattered field's intensity profile can be controlled.

Doing so can be accomplished by either making the grating coupler itself smaller, or by properly engineering the periodicity of the grating. This latter technique is akin to mode matching which we discuss next.

Mode Matching

Choosing the scattering angle is the first step towards accomplishing the broader goal of *mode matching*. In order to effectively couple light into a desired output mode, the grating coupler must generate a field—both amplitude and phase—that matches that mode. Based on the scattering angle analysis, we know that the phase of the generated field is directly related to the phase of the guided mode propagating within the grating coupler. Based on this, we can deduce that the exact phase of the scattered field can be manipulated by manipulating the per-period phase of the grating.

The amplitude or intensity profile of the scattered field, on the otherhand, depends on more than the periodicity of the grating. Recall that in our simple model, we posited that the discontinuity introduced by the etched slots gave rise to the scattered waves. If these etched slots have a uniform width over the length of the

grating, then each slot will scatter a fixed fraction of optical power out of the grating. In other words, the power scattered, P_i, by the ith slot of the grating is given by

$$P_i = \alpha P_g \tag{5.11}$$

where P_g is the guided power in the grating. The constant of proportionality α is determined primarily by the size (width and depth) of the slot[5]. Notice that the power in the grating decays according to the same constant of proportionality. After the $i-1$ th slot, the power in the grating is $P_g = (1-\alpha)^{i-1}P_0$ where P_0 is the total power entering the grating coupler. The power scattered by the ith slot is therefore

$$P_i = \alpha(1-\alpha)^{i-1}P_0 \tag{5.12}$$

Based on this, the scattered light will have an exponentially decaying intensity profile [85] along the length of the grating as depicted in the top of Figure 5.14. In most cases in silicon photonics, grating couplers are used to couple light into and out of single-mode optical fibers whose guided mode is not well matched this exponential field profile. Instead, it is thus desirable to design the grating corrugations to scatter light into a Gaussian mode, which approximately matches the mode of a single mode fiber.

In order to do this, we must chirp the widths and periodicity of the etched slots which form the grating coupler as depicted in the bottom of Figure 5.14. With the right grating *apodization* (as this chirp is often called), we can generate beams which which are well matched to the mode of optical fibers. This process of *mode matching* is essential to designing grating couplers with high efficiencies and is a prime candidate for optimization.

Grating Directionality

In addition to mode matching, grating coupler efficiency is additionally limited by its *directionality*. The directionality refers to the fraction of light which is scattered in the desired direction (any light not going into the desired mode is loss). In the case of the partially-etched grating couplers depicted in Figure 5.14, this directionality is largely defined by the thickness of the waveguiding slab and the depth of the etched grooves [86, 87].

[5]The dependence of the scattering strength of the slot width is not monotonic. Instead, the scattering strength peaks when it is equal to 50% of the grating period. This makes sense if we think about the grating as a series of discontinuities which scatter light. When the slot becomes larger than one half of the period, then the unetched portion of the grating acts as the primary discontinuity. In the limit that the etch width approaches the period of the grating, the scattering will cease altogether. For this reason, we often define the slot widths using *duty factor* which is the fraction of the total grating period which is unetched. The scattering is therefore minimized when the duty factor is equal to 0% or 100% and maximized when the duty factor equals 50%

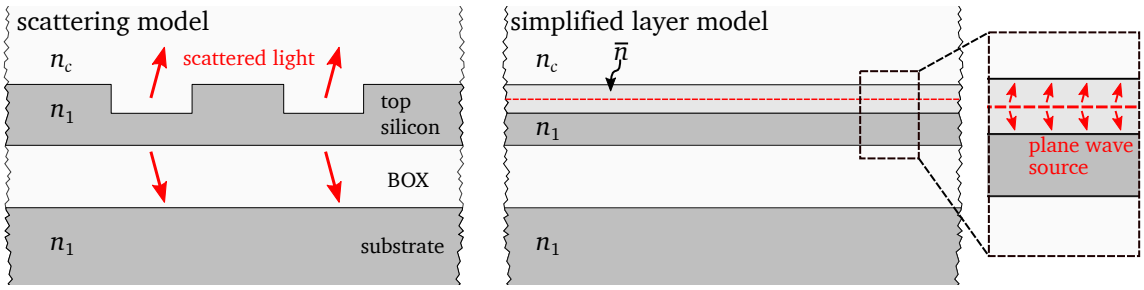

Figure 5.15: Depiction of the simplified layered medium model for grating couplers. In order to estimate the directionality of a grating coupler, we approximate it as a stack of homogenous slabs. The slab corresponding to the etched portion of the grating has an average refractive index based on the duty factor and contains a plane wave source which injects optical power into the stack. The directionality is given by the fraction of total injected power which leaves in the top cladding.

One way of understanding this relationship is to model the grating as a stack of homogenous slabs as depicted in Figure 5.15. In this model, we divide the cross section of the grating into multiple homogenous layers. The layer corresponding to the etched portion of the grating is assigned a weighted average refractive index based on the duty factor of the grating (e.g., in the case of a 50% duty factor, the refractive index would be the average of high index and cladding materials). In the center of this layer, we place a plane wave source which uniformly emits plane waves in the upward and downward directions. Based on the refractive indices and thicknesses of the layers, these plane waves will interfere within the stack, leading to a certain amount of power leaving the top and bottom of the stack. The fraction of power leaving out of the top of the stack, $P_{\text{up}}/(P_{\text{up}} + P_{\text{down}})$, is the directionality of the grating. This model, which is a simple extension of matrix methods[6], is an indispensable tool for understanding the limits in efficiency that can be achieved by typical grating couplers.

Beyond thickness and etch depth, the directionality can be further improved in a variety of ways as depicted in Figure 5.16. In particular, employing multiple etch depths, multiple silicon layers, or adding anti-reflection layers above the grating or reflectors below the grating have been proven to be effective. The former methods rely on constructing two arrays of scatterers which produce waves that destructively interfere below the grating in order to maximize the light coming out of the top of the grating. The latter two methods take advantage of interference within the

[6]The process involves “solving" the wave equation in each layer of the stack, which admits plane waves as a solution. Based on these general plane wave solutions, the boundary conditions for the electric and magnetic field are then enforced at each interface in the stack. This leads to a series of 2×2 matrices which can be multiplied together and inverted in order to find the amplitudes of the fields leaving the top and bottom of the stack. The application of matrix methods to grating couplers is a bit unique in that the source of the fields is located inside the stack instead of the source being a plane wave incident on the stack as is typically done.

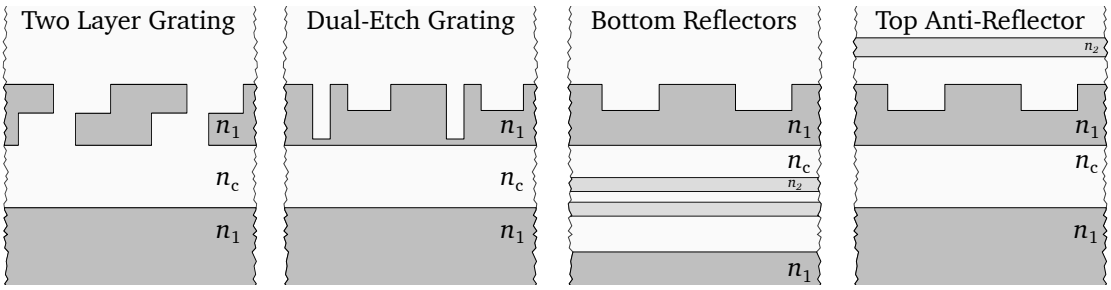

Figure 5.16: Ways of enhancing grating coupler directionality. In the left half, two layer gratings and dual-etch gratings rely on destructive interference of downward scattered waves to enhance the directionality. In the right half, reflectors or anti-reflectors leverage interference within the multilayer stack that defines the grating coupler to improve directionality.

layered stack that composes the grating to maximize light scattered out the top of the grating. In this work, we will take a detailed look at two layer gratings as they are a particularly promising solution for high efficiency vertical optical couplers.

Grating Back Reflection

A side effect of using discontinuities to scatter light is that they also produce back reflections into the input waveguide. In addition to reducing the power coming out of the grating coupler, these back reflections also pose an issue to system stability[7].

The fraction of power reflected is strongly tied to the scattering angle. To see why, consider Equation (5.10) when $\theta = 0$. In this case, the grating equation becomes

$$\varphi_{\text{in}}(\Lambda) = 2\pi m \quad . \tag{5.13}$$

In other words, in order to scatter light at zero degrees (perfectly vertically), the phase accumulated by a wave guided by the grating between two adjacent slots is a multiple of 2π! This means that all waves *reflected* by the grating slots will constructively interfere, which will inevitably lead to a very strong back reflection coming out of the grating.

In order to get around this issue, we often design gratings to scatter at a small angle ($\sim 10^\circ$). This results in a per-period phase which is a little larger than 2π, limiting constructing interference of the large ensemble of reflected waves. From the standpoint of packaging, this is not necessarily the most convenient constraint. Fortunately, as we will find out later in this chapter, we can use sophisticated structures like two layer gratings to simultaneously emit vertically and cancel out the reflected waves.

[7]For example, if a grating is directly coupled to a laser cavity, any reflection back into the laser can affect its operation.

5.3.2 General Design Approach

As with the devices treated previously in this work, we will tackle grating coupler design using our hierarchical approach. Our starting point is therefore to devise a physics-defined topology. Based on our previous discussion, the bare minimum for a good starting structure is one that scatters light at the desired angle. Furthermore, if we have control over the thicknesses of the grating, our starting structure should also maximize directionality. This ensures that we are starting with a “good" grating coupler.

In addition to starting with the correct angle, it turns out to be imperative that our initial topology have a large duty factor (small slots), and hence scatter light very gradually. This is a direct consequence of the fact that we typically want to couple light into a Gaussian beam which has a smooth and gradually changing intensity profile. Large slots would lead to a large discontinuity at the beginning of the grating coupler which would cause a strong spike in scattered light. This does not match the gradually increasing tails of a Gaussian beam well and can make the optimization susceptible to falling into local optima. Conversely, by starting with a large duty factor, we introduce a minimal discontinuity between the input waveguide and grating. This is well matched to the gradual tails of the Gaussian and yields consistently good optimized results.

After defining the initial topology, which consists of a high duty factor grating with the right scattering angle, we run a coarse optimization of the structure. In most cases, the silicon platform we use constrains us to a specific set of silicon thicknesses and etch depths. In this case, the goal of the optimizations will primarily be to define the grating chirp which maximizes the mode match with respect to a Gaussian mode and which minimizes reflections. Once again, the gradual nature of the Gaussian mode into which we couple light gives us clues about the type of parameterization to choose. In particular, it is desirable to choose a parameterization which guarantees that the grating period and etch width evolves smoothly and gradually over the length of the grating, since any sudden changes are unlikely to produce a field profile which matches the desired beam shape. For this purpose, a truncated Fourier series is well suited as it is guaranteed to be smooth and does not blow up at the ends as is the case with other expansions (like Taylor series). In particular, we define the grating period as

$$\Lambda(n) = a_0 + \sum_{m=1}^{M} a_m \sin\left(\frac{\pi}{2}\frac{m}{N}n\right) + \sum_{m=1}^{M} b_m \cos\left(\frac{\pi}{2}\frac{m}{N}n\right) \tag{5.14}$$

where a_m and b_m serve as design parameters and n is an index assigned to each period of the grating (the first period is $n = 0$, the second is $n = 1$, etc). The etched slot widths (or duty factor) can be parameterized using an identical function. Notice that by tuning the number of Fourier terms, we can directly control how quickly the

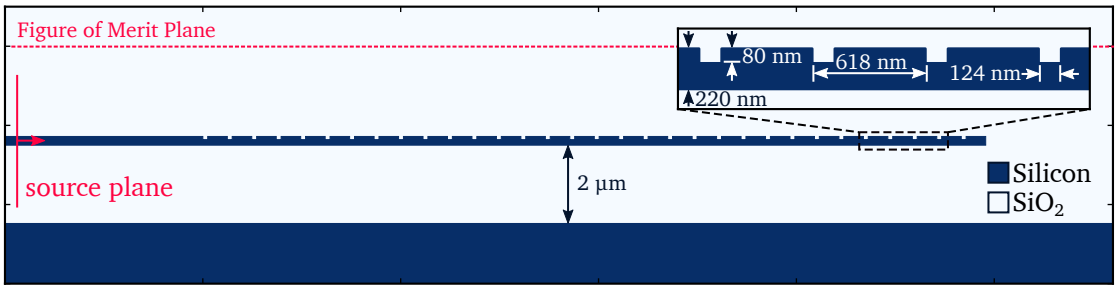

Figure 5.17: Starting geometry and problem setup for partially-etch grating coupler optimization. The starting structure consists of a uniform grating coupler with large duty factor and a grating period chosen to scatter light at the desired angle.

periods and slot widths can change along the length of the grating.

A natural shortcoming of this parameterization and the fact that we want to couple into a Gaussian beam which decays to zero in its tails is the inevitable formation of slots with near zero widths at the beginning of the grating. These small widths cannot typically be resolved using deep UV lithography, and therefore it is important that we introduce a linewidth constraint into our optimization. This is the purpose of the refinement optimization.

A good way to accomplish this is to start with the coarse optimized design and re-parameterize the grating in terms of the individual periods and slot widths. We can then introduce a penalty term in our figure of merit based on Equation (A.35) which discourages slots with widths between 0 and some minimum line width (negative widths are acceptable as they correspond to adjacent grating teeth merging). Typically, we tune the strength and sharpness of this penalty gradually over two or three separate short refinement optimizations in order to avoid getting stuck in a local optima which does not satisfy the constraint.

By following this general process, we have had great success optimizing a wide range of grating couplers including more common partially-etched grating couplers, dual etch grating couplers, hybrid silicon-silicon nitride grating couplers, two layer grating couplers, etc. In the remaining sections of this chapter, we will demonstrate this methodology by optimizing a two different grating couplers.

5.3.3 Partial Etch Grating Coupler

The most common type of single polarization grating coupler consists of a waveguide slab with partially etched grooves. In this section, we will optimize a grating coupler which is defined in $220\,\mathrm{nm}$ thick SOI, with an $80\,\mathrm{nm}$ etch depth and a $2\,\mu\mathrm{m}$ thick buried oxide (BOX) layer. The grating will be optimized for C-band operation, targeting a center wavelength of $1550\,\mathrm{nm}$.

The starting topology for our grating optimization is shown in Figure 5.17. The

structure consists of 32 identical etched slots whose widths correspond to a duty factor of 80%[8] In order to find the period and slot width, we apply Equation (5.10). The period accumulated by a guided wave in the grating over one period is

$$\varphi_{\mathrm{in}} = k_{\mathrm{ft}} D\Lambda + k_{\mathrm{pe}}(1-D)\Lambda \tag{5.15}$$

where $k_{\mathrm{ft}} = 2\pi n_{\mathrm{ft}}/\lambda$ is the effective wavenumber of the full thickness silicon, $k_{\mathrm{pe}} = 2\pi n_{\mathrm{pe}}/\lambda$ is the effective wavenumber of the partially etched silicon, and D is the duty factor (in this case, $D = 0.8$). Substituting this expression into Equation (5.10) and solving for the grating period yields

$$\Lambda = \frac{m\lambda}{n_{\mathrm{ft}} D + n_{\mathrm{pe}}(1-D) - n_c \sin\theta} \tag{5.16}$$

where n_{ft} is the effective index for full thickness silicon, n_{pe} is the effective index for partial etch silicon, and n_c is the refractive of the cladding medium. Based on our chosen top silicon thickness, etch depth, and duty factor, the period of our starting structure is $618\,\mathrm{nm}$, and the slot width is $124\,\mathrm{nm}$.

For the sake of time, we simulate the grating in two dimensions, which is a reasonably good approximation since grating couplers are typically much wider than they are thick. In order to ensure sufficient simulation accuracy, we use a grid spacing of $20\,\mathrm{nm}$[9].

In total, we perform three optimization in order to optimize this grating coupler. In the first optimization, we parameterize the period and slot width using Equation (5.14) with 6 $\sin$ and $\cos$ coefficients and also allow the starting position of the grating to change. This means that our first (coarse) optimization has a total of 25 design parameters. The figure of merit for this optimization is simply the coupling efficiency computed using Equation (A.25) with respect to a $10.4\,\mu\mathrm{m}$ mode field diameter Gaussian *mode* which is tilted at eight degrees, which approximates the mode of an SMF-28 fiber.

The optimization is performed using the BFGS minimization algorithm and is allowed to run until the figure of merit changes by less than 10^{-5}. This process takes 39 iterations (or about 17 minutes on 14 cores of our Intel Xeon based server) as shown in the left portion of Figure 5.18, improving the initial low coupling efficiency of nearly $-6\,\mathrm{dB}$ to $-1.89\,\mathrm{dB}$. The result of this coarse optimization is shown in Figure 5.19 which plots the field overlayed with the optimized geometry. As desired, the generated field closely matches the desired Gaussian beam, deviating only near the end of the grating. It is important to note that the expected directionality for this grating is around 67% ($-1.74\,\mathrm{dB}$) based on our homogeneous slab model, implying that our optimized design is likely at the limit of performance given the

[8] A good rule of thumb is to choose an initial duty factor between 65% and 90%. Duty factors outside of this range tend to drive the optimization into local optimas which have low efficiencies.

[9] In many cases, we could work with a coarses grid without significantly impacting the calculated coupling efficiency. The back reflection, however, tends to be very sensitive to the grid spacing.

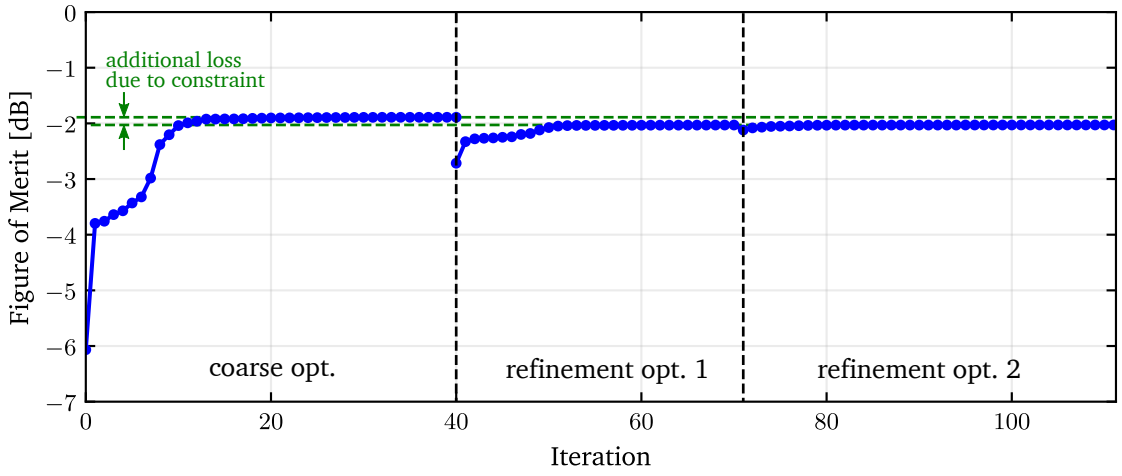

Figure 5.18: Evolution of the figure of merit during the coarse and refinement optimizations of a partially-etched grating coupler. The final performance is slightly lower than that of the coarse optimized design as a result of the minimum feature size constraint.

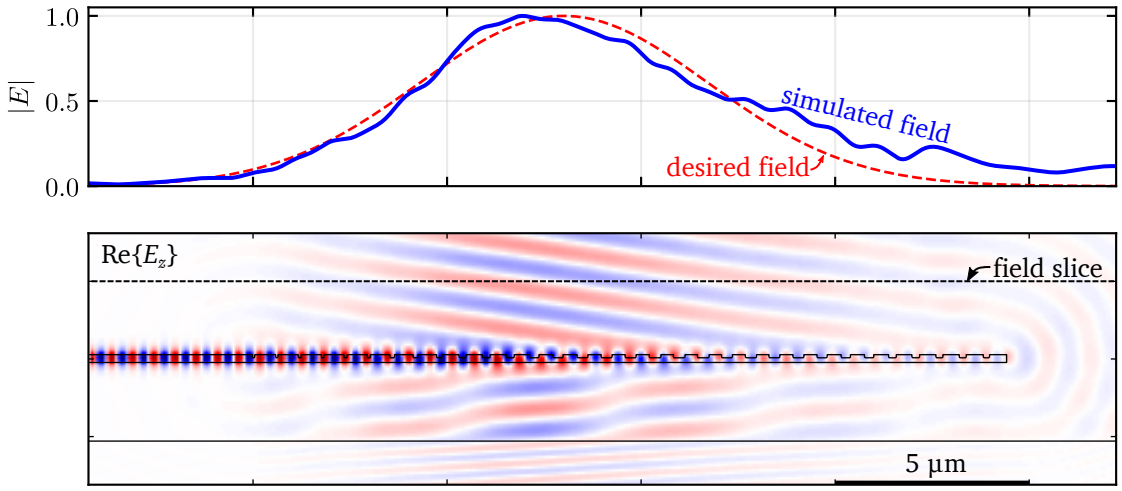

Figure 5.19: Simulated field of the coarse optimized partially-etched grating coupler. In the bottom, Re$\{E_z\}$ is overlayed with the outline of the optimized grating geometry. At the top, a slice of the electric field taken from the dotted line is overlayed with the desired electric field profile. As intended, the simulated field of the optimized grating closely matches the desired field profile.

constrained silicon thickness and etch depth. This demonstrates just how powerful our optimization method is.

The period and duty factor of this optimized design is plotted in Figure 5.20. As desired, the parameterization has forced the grating apodization to remain smooth.

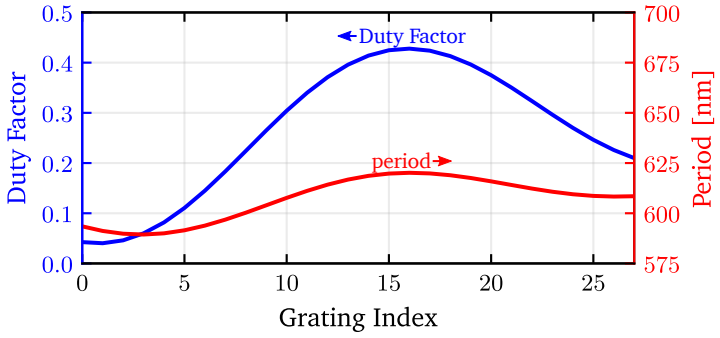

Figure 5.20: Final grating chirp (period and duty factor vs position along the grating) of the coarse optimized partially-etch grating coupler. The evolution of the period and duty factor is smooth as a result of the underlying Fourier series parameterization.

It is interesting to note that the optimized duty factor approaches zero at the beginning of the grating, and the first slot in the grating is only $15\,\mathrm{nm}$ wide. This is not achievable using deep UV lithography.

In order to account for this, we run additional optimizations which incorporate line width constraints. For these constrained optimizations, we make each slot width and pitch a separate design variable and add a penalty term to the figure of merit. In particular, the new penalized figure of merit has the form

$$F(\mathbf{E}, \mathbf{H}, \vec{p}) = \eta(\mathbf{E}, \mathbf{H}) - \alpha \mathcal{P}_{\mathrm{LW}}(\vec{p}) \tag{5.17}$$

where $\eta(\mathbf{E}, \mathbf{H})$ is the coupling efficiency with respect to the same $10.4\,\mu\mathrm{m}$ Gaussian mode, $\mathcal{P}_{\mathrm{LW}}(\vec{p})$ is a penalty function based on Equation (A.35), and α is the weighting of the penalty function. In this case, $\alpha = 0.02$ and two optimizations with this figure of merit are performed in succession[10] for a minimum line width of $90\,\mathrm{nm}$. In the first optimization, the penalty function has very gradual transitions from low to high (the transition occurs nominally over a distance equal to $40\,\mathrm{nm}$). In the second optimization, the transition is made very steep to ensure that the penalty function is more strictly enforced.

It is interesting to note that the optimization maintains a high performance by eliminating the first five grating slots and by modulating the remaining grating slot widths and periods in a slightly more complicated way as shown in Figure 5.22. The period chirp of the optimized result in particular, is unintuitive. We believe the oscillation in the period towards the beginning of the grating is in part an “attempt" to reduce reflection back into the input waveguide. This back reflection is plotted versus wavelength for both the unconstrained coarse optimization result and line-width-constrained refinement optimization result in Figure 5.23. Although not

[10] In many cases, three optimizations with a more gradual turn on of the penalty function may work better. This will depend on the specific grating you are trying to optimize.

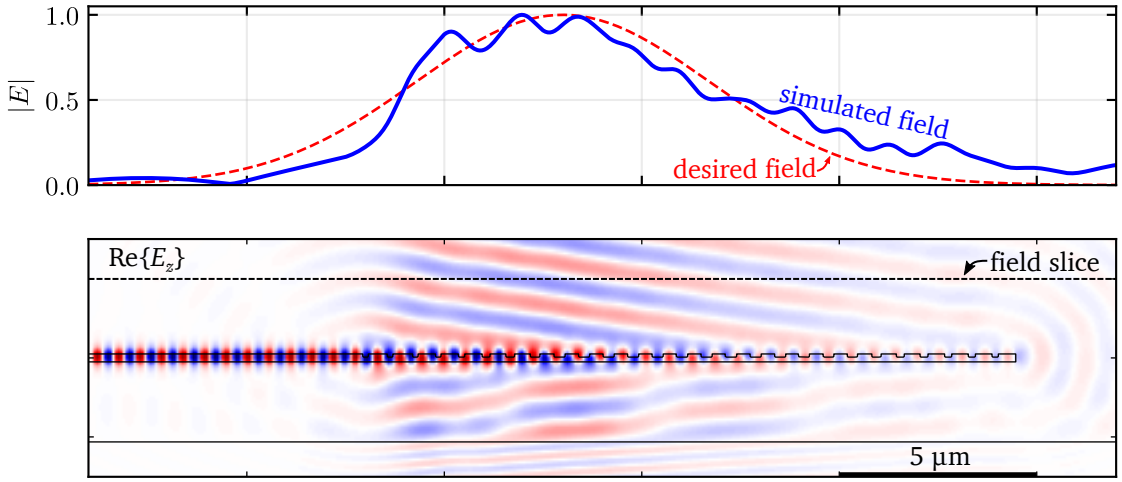

Figure 5.21: Simulated field of the refinement optimized partially-etched grating coupler. In the bottom, Re$\{E_z\}$ is overlaid with the outline of the optimized grating geometry. At the top, a slice of the electric field taken from the dotted line in the bottom plot is overlaid with the desired electric field profile. In spite of imposing a minimum feature size constraint, the scattered field still closely matches the desired field.

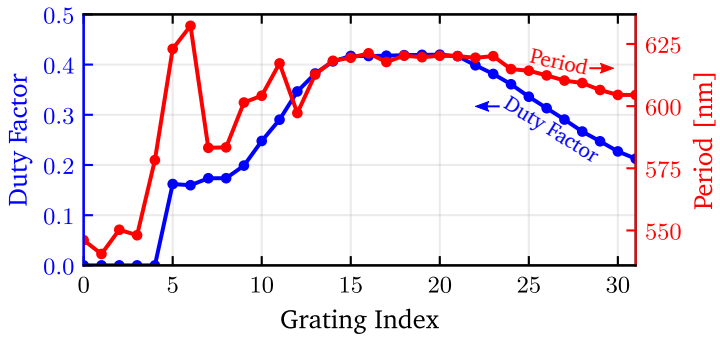

Figure 5.22: Final grating chirp (period and duty factor vs position along the grating) of the refinement optimized partially-etch grating coupler. The evolution of the period and duty factor is no longer smooth due to the tooth-wise parameterization.

explicitly included in the figure of merit, the optimization is successful in producing designs with unprecedentedly low back reflections. This makes intuitive sense, however, as any reflection will reduce the coupling efficiency. These results achieve back reflections comparable to so called “reflectionless" grating couplers [88] over the C-band without sacrificing significant coupling efficiency as is the case in those designs.

In addition to achieving low back reflection, the optimized design achieves bandwidths that are on par with other published grating couplers as shown by the C-

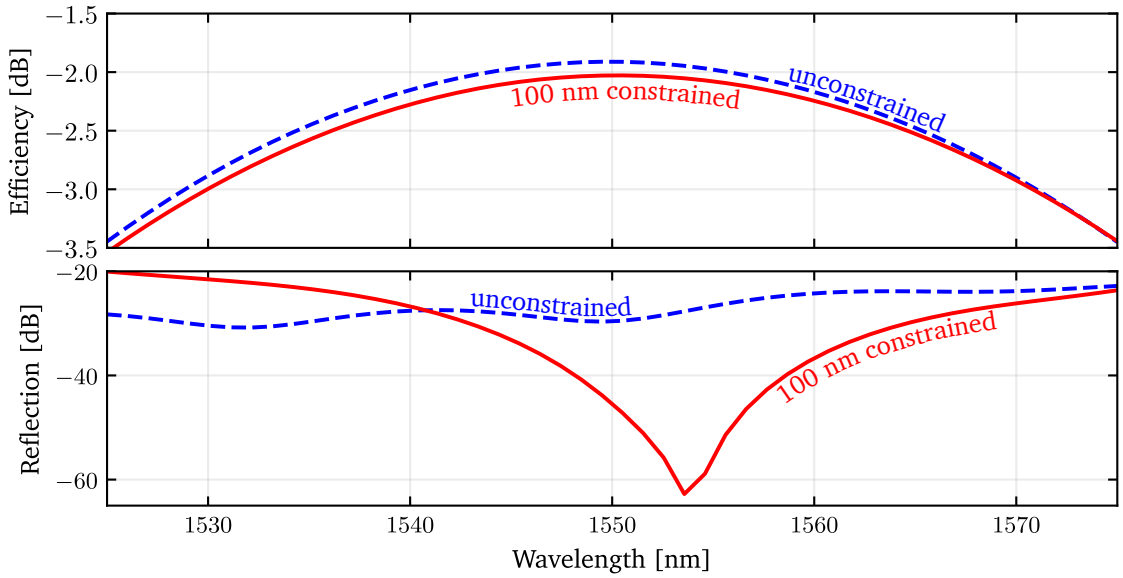

Figure 5.23: Plot of the (top) coupling efficiency and (bottom) back reflection of an optimized partially-etched grating coupler. The optimized grating with constrained features sizes predictably achieves a slightly lower peak performance compared to the unconstrained design. Both constrained and unconstrained designs achieve exceptionally low back reflections, which is a hallmark of our approach to design and optimization.

band coupling efficiency plotted in the top of Figure 5.23. In particular the optimized design achieves a $1\,\mathrm{dB}$ bandwidth of approximately $40\,\mathrm{nm}$.

These results demonstrate how effectively inverse design techniques can be applied to grating couplers which have historically been among the most difficult silicon photonics components to design. In particular, we see here that grating couplers compatible with deep UV lithography can be designed without sacrificing significant performance. This proves that inverse design is poised to become an essential tool in the silicon photonic designer's toolbox.

5.3.4 Dual Layer Grating Couplers

In the previous grating coupler optimization, the final performance of the grating coupler is largely limited by its inherent directionality. In order to get around these limitations, two layer grating couplers [89,90] have been proposed. When designed properly, these two-layer gratings act as a phased array of scatterers which couple light out of the waveguide with a high directionality as depicted in the left half of Fig. 5.24.

A guided wave propagating from left to right along the grating will encounter grooves in the top and bottom surfaces of the grating which couple some of the propagating power into free-space modes. The grating achieves high directionality

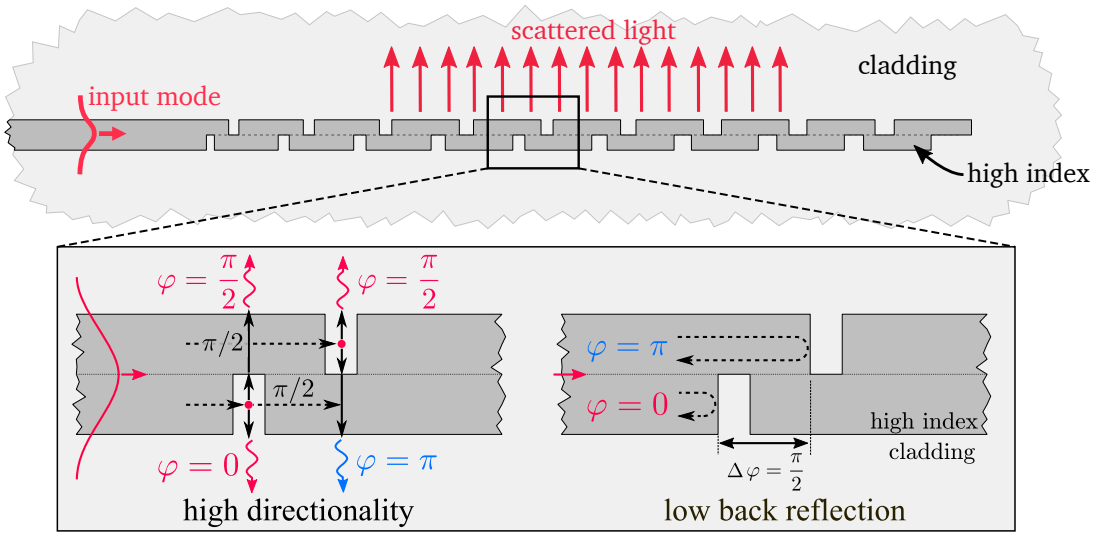

Figure 5.24: Description of the operation of two-layer grating couplers. The two-layer grating resembles two fully-etched grating couplers which have been stacked one on top of the other with the top layer shifted forward by a small amount relative to the bottom layer. When each grating layer is one quarter of a wavelength thick and the spacing between grooves in each layer is a quarter wavelength, then constructive interference in the upwards direction and destructive interference in the downwards direction can be achieved. Furthermore, under these conditions, waves reflected back into the input of the grating destructively interfere leading to an inherently low back reflection.

if three conditions are satisfied. First, the horizontal separation between grooves in the top and bottom layer is chosen such that the guided mode accumulates a phase shift of $\pi/2$ as it propagates from a groove in the bottom layer to its adjacent top layer groove. Second, the thicknesses of the top and bottom layers and etch depths of the grating are chosen such that a wave scattered out of a bottom groove acquires an additional $\pi/2$ phase shift relative to a wave scattered out of a top groove. Third, the scattering strength of the two layers should be equal. The resulting structure will produce scattered waves which interfere constructively in the upwards direction and destructively in the downwards direction. In addition to enabling a directionality of nearly 100%, the relative shift between the top and bottom layers produces two sets of reflected guided waves within the grating which are π out of phase as depicted in the right half of Fig. 5.24. These waves destructively interfere, resulting in an exceptionally low back reflection, even for perfectly vertical coupling.

Two layer gratings provide the additional degrees of freedom needed to simultaneously control the upward scattered waves, downward scattered waves, and back reflected waves. The addition of these degrees of freedom, however, adds additional complexity to the design problem, almost necessitating that we apply our

inverse design. As with the previous partially-etched grating coupler example, we will apply our hierarchical approach to design two layer grating couplers which couple light *perfectly vertically* with unmatched coupling efficiency. The process is very similar to the optimization of the partially-etched grating with the exception that the two separate layers lead to twice as many degrees of freedom.

The starting structure for the optimization is shown in Fig. 5.25 and consists of a uniform two-layer silicon grating clad both on top and underneath with silicon dioxide. The grating structure is excited by the fundamental TE mode of the waveguide shown at the left edge of Fig. 5.25 at a wavelength of $1550\,\mathrm{nm}$. In the following optimizations, we only consider this single wavelength (although the same method could be applied to a broadband figure of merit if desired). Given this wavelength, we choose the layer thicknesses to be $110\,\mathrm{nm}$ to produce the desired $\pi/2$ phase shift between the top and bottom layer, resulting in a total film thickness of $220\,\mathrm{nm}$. Based on our previous reasoning, we choose a starting duty factor of 80%, a grating period of $586\,\mathrm{nm}$, and a shift between the top and bottom layers of $160\,\mathrm{nm}$ to ensure that the starting efficiency is reasonably high, which in this case was 48%. Finally, a total of 26 grating periods are used, which we found sufficient for scattering all of the light out of the grating.

This initial structure is modified using our gradient-based optimization methods in order to maximize the efficiency with which the grating couples light into a $10.4\,\mu\mathrm{m}$ mode field diameter Gaussian beam corresponding to the approximate mode of a single mode fiber which is situated $2\,\mu\mathrm{m}$ above the grating and oriented *normally*[11] to its top surface. The figure of merit is therefore the the mode match with respect to the desired Gaussian beam.

The optimization process is executed until the figure of merit changes by less than 10^{-5}, which takes 67 iterations of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) minimization algorithm (requiring 166 simulations in total and about 30 minutes on 14 processors). The corresponding optimized structure and simulated electric field is displayed in Fig. 5.26. As desired, light is coupled vertically and downwards coupling is almost entirely suppressed. The wavefronts, furthermore, are very flat and well-behaved; the quality of this generated beam is reflected by a record chip-to-fiber coupling efficiency of 99.2% ($-0.035\,\mathrm{dB}$). This coupling efficiency is significantly higher than previously reported values [90] and does not rely on coupling to narrow mode field diameter fibers or steep emission angles as in previous related work [89].

The mode matching capabilities of the optimized grating are more quantitatively demonstrated by the slices of E_z shown in Fig. 5.27. As is readily apparent, the magnitude of the simulated electric field very closely matches the desired Gaussian field profile, with the exception of some weak rippling. We attribute this rippling to

[11] Scattering light perfectly vertically is typically very difficult due to the increased back reflections. The anti-reflection nature of these two layer gratings enable us to couple perfectly vertically with high efficiency.

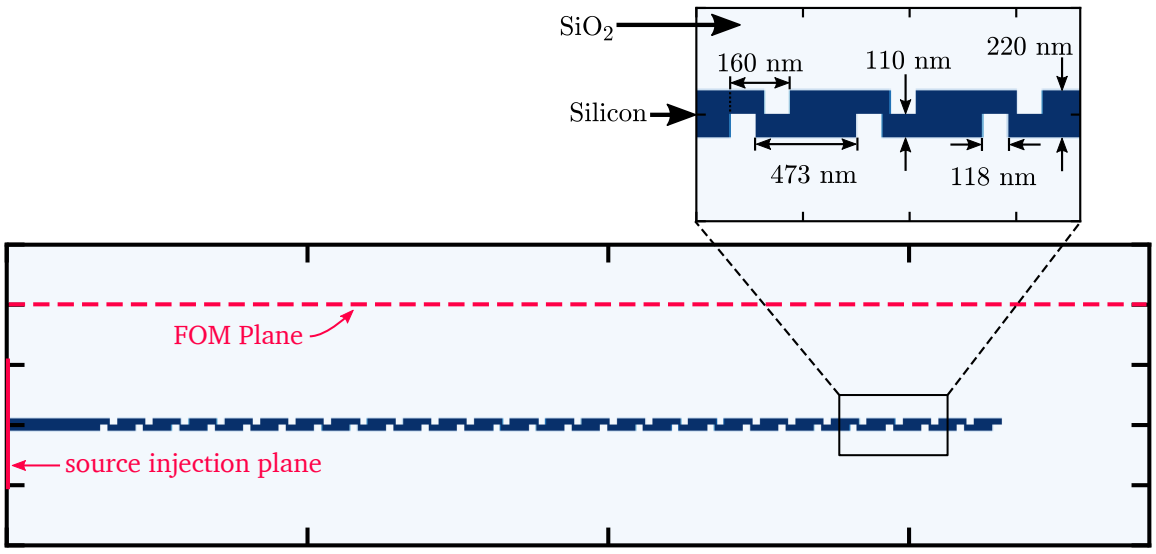

Figure 5.25: Plot of the initial starting geometry used in the optimization of a two layer SiO_2-clad silicon grating coupler. A uniform grating with a duty factor of 80% and a period of $586\,\mathrm{nm}$ is chosen for both the top and bottom layers. This choice of period and duty factor results in a nearly vertical beam with a high directionality.

the inherently discrete rectangular etches of the grating. This has a minimal impact on the overall efficiency, which is demonstrated by the calculated mode match of over 99%. As with the electric field amplitude, the simulated phase is very flat over the majority of the beam's width, deviating only near the edges of the beam. This deviation does not lead to an appreciable decrease in efficiency since the field amplitude is nearly zero far away from the beam's center where the phase begins to fluctuate. In addition to mode matching, the back reflection of waveguide mode incident on the grating coupler is suppressed by $-43\,\mathrm{dB}$ for our optimized design.

The unprecedented efficiency of this optimized grating is a direct consequence of the chirping of the grating period and duty factor visible in the insets of Fig. 5.26 which is plotted as a function of grating index (position along the grating) in Fig. 5.28. The optimal chirp function for the duty factor (upon which the scattering strength of the grating depends most strongly) of the top layer roughly matches the behavior of the theoretical scattering parameter for an ideal grating coupler matched to a Gaussian field profile [91]. The chirp function of the bottom layer's duty factor, meanwhile, deviates from this ideal behavior, highlighting the strength of our optimization methods: its ability to design structures that would otherwise be difficult or impossible to design "by hand."

Figure 5.28 also highlights the primary shortcoming of these optimization results: the optimal design contains duty factors that approach 100%. This means that the optimal structure contains features that are as small as a few nanometers wide and hence cannot be fabricated using available deep UV lithography. We

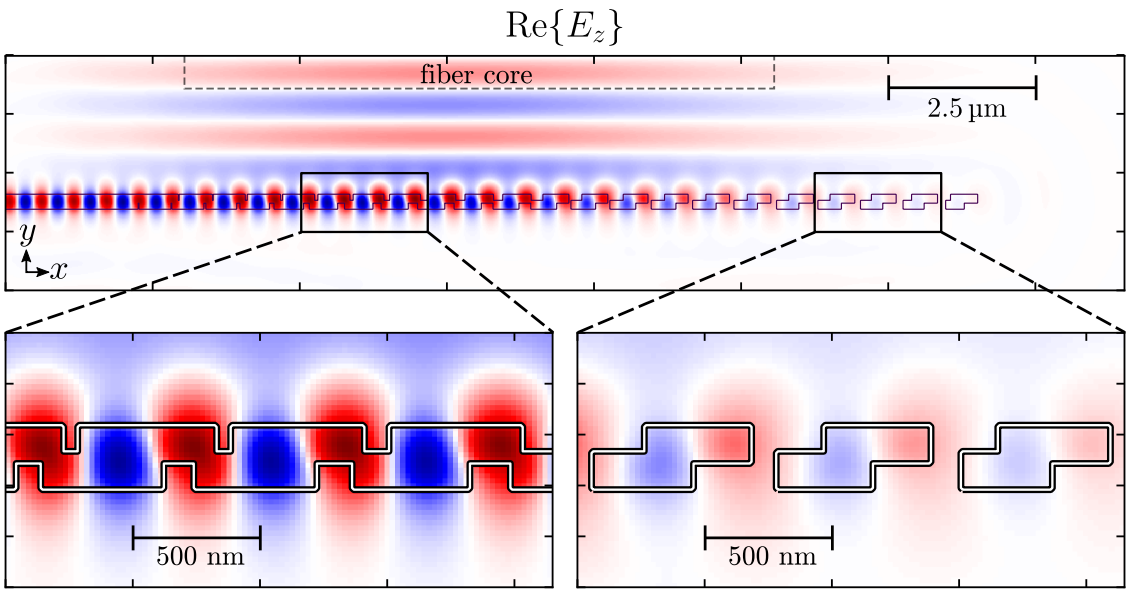

Figure 5.26: Optimization results for our perfectly vertical two layer grating coupler. The real part of E_z, which has been overlaid with an outline of the optimized refractive index, shows perfectly vertical emission with an extremely high directionality and flat wavefronts. This optimized structure is very well mode-matched to the mode of a $10\,\mu\text{m}$ mode field diameter single mode fiber located $2\,\mu\text{m}$ above the grating surface which is reflected by a chip-to-fiber efficiency of 99.2% ($-0.035\,\text{dB}$).

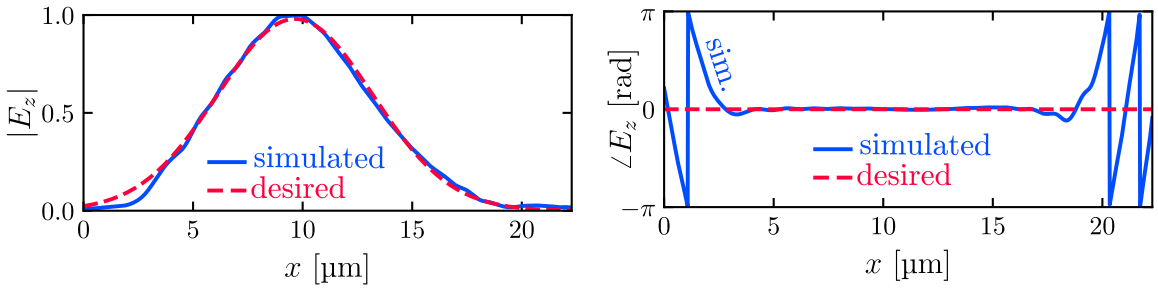

Figure 5.27: Plots comparing the simulated electric field amplitude (top) and phase (bottom) of the optimal grating coupler design to the desired amplitude and phase. The visible deviation in the simulated phase is inconsequential as it occurs only when the field amplitude is very small.

overcome this by running additional refinement optimizations which introduce a minimum feature size constraint directly into the optimization and use our unconstrained design as a starting point for those subsequent optimizations.

In order to accomplish this, we modify our original figure of merit such that the efficiency of the grating is penalized when small features form. The new figure of merit is given by

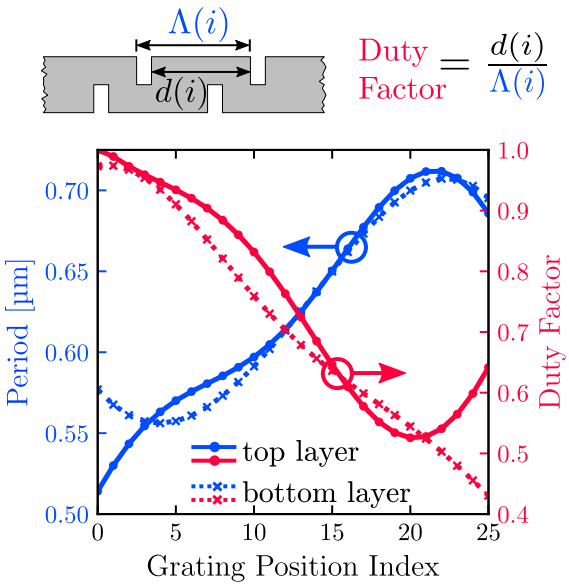

Figure 5.28: Plot of the chirp functions for the optimized two-layer grating. The blue curves show the the period as a function of the position (index) along the grating. The red curves, meanwhile, show the duty factor along the grating. In both sets of curves, the solid trace corresponds to the top layer while the dotted curve corresponds to the bottom layer of the grating.

$$F(\mathbf{E}, \mathbf{H}, \mathbf{p}) = \eta(\mathbf{E}, \mathbf{H}) - f_{\text{penalty}}(\mathbf{p}) \tag{5.18}$$

where $F(\mathbf{E}, \mathbf{H}, \vec{p})$ is our new constrained figure of merit, $\vec{p}$ is the set of design variables (i.e. grating dimensions), $\eta(\mathbf{E}, \mathbf{H})$ is the mode match efficiency given by Equation (A.25), and $f_{\text{penalty}}(\vec{p})$ is a function of the design variables based on Equation (A.35) which penalizes the efficiency when the feature sizes of the grating are too small. Because this penalization process operates on each individual gap and tooth in the grating, we no longer use a Fourier series parameterization in subsequent optimizations, but instead opt for parameterization in which each gap and tooth dimension is a separate independent design variable. Since we begin with an optimized design with a smooth chirp function, there is less danger of falling into a non-robust local optimum. Finally, in order to control the influence of this penalty function, we adjust its maximum value and the steepness of its edges.

For the purpose of exploring the impact of minimum feature size on grating coupler efficiency, we ran a series of refinement optimizations for a specific set of minimum feature sizes. Beginning with the previous optimized structure shown in Fig. 5.26, we introduce a penalty function which is very weakly weighted and then gradually increase its strength in a series of separate optimizations. In total we perform three optimizations per minimum feature size, strictly enforcing the

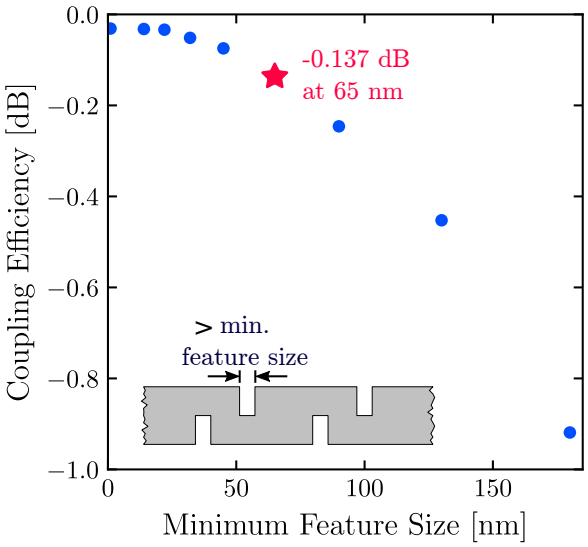

Figure 5.29: Plot of the coupling efficiency for perfectly vertical grating couplers optimized with a minimum feature size constraint. Using the "ideal" optimized result as a starting point, additional constrained optimizations are performed in order to design gratings that can be fabricated with lithography that has a limited resolution.

minimum feature size before running the third optimization.

The result of this process, which we performed for a set of minimum feature sizes between $0\,\mathrm{nm}$ and $180\,\mathrm{nm}$, is shown in Fig. 5.29. Using constrained optimization, we are able to maintain exceptionally high efficiencies out to more practical feature sizes. Of particular interest is the $65\,\mathrm{nm}$ constraint which, due to the maturity of the $65\,\mathrm{nm}$ CMOS platform, shows future promise for silicon photonics. For a minimum feature size of $65\,\mathrm{nm}$, we have achieved an optimized efficiency of 96.9% ($-0.137\,\mathrm{dB}$). Furthermore, better than $-0.5\,\mathrm{dB}$ is achievable out to a minimum feature size of over $130\,\mathrm{nm}$, roughly corresponding to lithography technologies that have already been used in commercial nanophotonic settings [84]. As the minimum feature size is further increased, the efficiency of the optimized design begins to fall off quickly. This is a direct consequence of our inability to match to the gradually changing tails of the Gaussian mode. These results are very promising for both current and future applications, especially considering that such structures have already been fabricated using an existing CMOS process [89] which is capable of resolving the feature sizes we have considered.

In addition to maintaining a high efficiency at the design wavelength, our feature-size-constrained designs also maintain a reasonable bandwidth. Figure 5.30 shows the coupling efficiency plotted as a function of wavelength for the "ideal" unconstrained design and the $65\,\mathrm{nm}$ design. In both cases, the $1\,\mathrm{dB}$ bandwidth is about $24\,\mathrm{nm}$. This is partially due to the number of periods in the grating: in general,

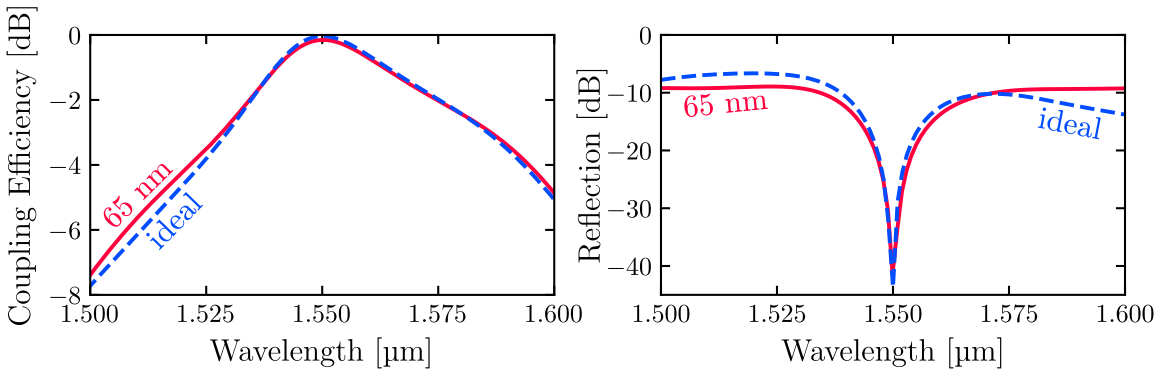

Figure 5.30: Plots of the coupling efficiency (top) and back reflection (bottom) as a function of wavelength for the ideal optimized result (blue dashed line) and the $65\,\mathrm{nm}$ constrained minimum feature size result (red solid line). Both cases achieve a high peak coupling efficiency at $1550\,\mathrm{nm}$ as well as a modest $1\,\mathrm{dB}$ bandwidth of $24\,\mathrm{nm}$. The reflection in both cases is exceptionally low, reaching below $-40\,\mathrm{dB}$ at $1550\,\mathrm{nm}$.

the more periods a grating has, the narrower its bandwidth will be. The number of grating periods used was kept the same, independent of the minimum feature size. If necessary for a given application, this bandwidth could be increased by coupling into a fiber with a smaller mode field diameter which would allow us to reduce the number of periods within the grating. In addition to the number of grating periods, the insertion loss bandwidth is also reduced some as a result of the narrower wavelength response of the back reflection plotted in the bottom of Fig. 5.30.

Depending on the application, the reflective properties of the grating may be very important. In the case of both our unconstrained and constrained designs, the back reflection into the input waveguide is exceptionally low at the design wavelength. At $1550\,\mathrm{nm}$, the unconstrained and constrained designs achieve reflections of $-43\,\mathrm{dB}$ and $-41\,\mathrm{dB}$, respectively. Although not explicitly part of the figure of merit used during optimization, reducing reflection is important to improving the overall insertion loss of the device. It is thus not unexpected that the optimal solution has a very low reflectivity. It is worth noting that while the reflection is so low at the design wavelength, it increases very quickly as the wavelength deviates from $1550\,\mathrm{nm}$, as indicated by the $-20\,\mathrm{dB}$ bandwidth of only $8\,\mathrm{nm}$. In situations in which low reflection is required, the usable bandwidth of the device may be significantly smaller than the $1\,\mathrm{dB}$ insertion loss bandwidth. This is in part due to our choice to emit vertically, which likely increases the sensitivity of the back reflection to variations in wavelength. If a more broadband back reflection and coupling efficiency is required, gratings designed to couple light at a small angle could be used. Designing such gratings is a minor modification to the process we have presented here.

In our previous discussion, we considered only the chip-to-fiber coupler case. Depending on the application, the fiber-to-chip coupling case may be equally im-

portant. Unfortunately, it is not immediately apparent that the two cases should act in a reciprocal way. In order to get a sense of how well our designs perform in he fiber-to-chip coupling case, we have performed additional simulations of the structure at the design wavelength of $1550\,\mathrm{nm}$. In these simulations, the grating is excited by a $10.4\,\mu\mathrm{m}$ mode field diameter Gaussian beam and the coupling efficiency into the grating coupler waveguide as well as the back reflection into the input fiber are calculated. In the case of our unconstrained design, we found that the coupling efficiency was identical to the chip-to-fiber case shown previously while the back reflection into the fiber was slightly reduced to $-44\,\mathrm{dB}$. In the case of our $65\,\mathrm{nm}$ constrained design, the coupling efficiency was once again the same as the chip-to-fiber case shown previously, while the back reflection was minimally increased to $-37\,\mathrm{dB}$ (which is still uniquely low).

Chapter 6

Towards an Optimized Silicon Photonic Component Library

Over the last two decades, silicon photonics has rapidly matured, leading to a growing interest in building large complex systems consisting of thousands of silicon photonic components. A direct consequence of this push towards large scale integration is the need for high efficiency silicon photonic building blocks.

In this work, we have presented a concrete path towards realizing those essential silicon photonic building blocks. The foundation of our approach to designing silicon photonic components is gradient-based shape optimization. A key enabling aspect of our formulation of shape optimization is *boundary smoothing* based on high numerical precision polygons. In addition to helping us calculate accurate device sensitivities, this method affords us a great amount of flexibility when representing device geometries and enables us to incorporate design constraints directly into optimizations in a simple and intuitive way.

Our approach to gradient-based optimization shares an important similarity to other forms of shape and topology optimization employed in the nanophotonics community: on its own, it is not a complete solution to designing high performance and robust devices. Due to the inherently non-convex nature of electromagnetic optimization problems, we cannot expect convex optimization to universally yield good devices without outside input. In order to directly tackle this problem, we have systematized the process of providing “outside input" through our hierarchical approach to design and optimization. Using a strategic combination of simple physical analysis to find good starting geometries, optimization with coarse parameterizations, and constrained optimization with fine parameterizations, we have shown that efficient and robust devices can be designed with minimal guesswork.

Using our hierarchical approach, we have demonstrated how a variety of silicon photonic components can be designed with superior performance. These demonstrations represent an important step towards realizing an *optimized silicon photonics component library*. Nonetheless, numerous silicon photonics components remain

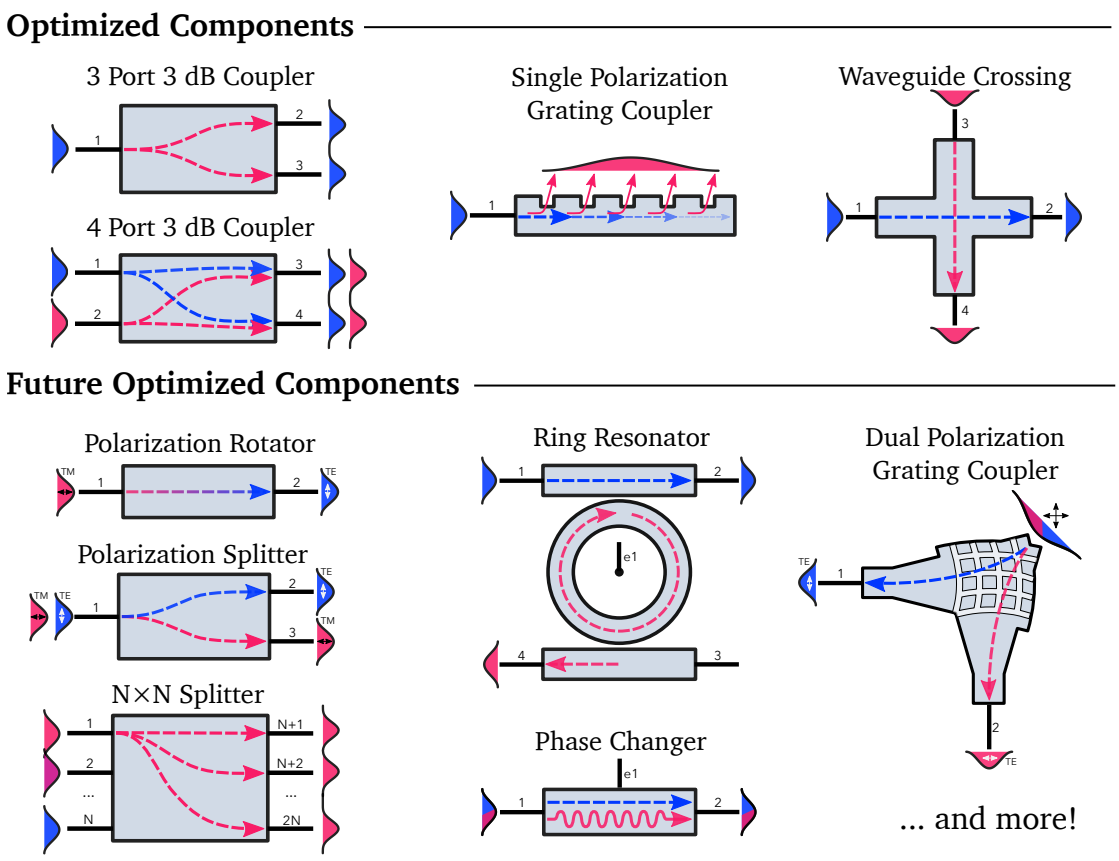

Figure 6.1: Overview of already optimized silicon photonic components and components which could benefit from optimization in the future.

which will benefit from hierarchical design and optimization (see Figure 6.1). Based on the growing demand for efficient silicon photonic components and the ability of shape optimization to meet those demands, we expect to see these methods play an important role in a number of rising applications and industries that leverage silicon photonics.

Appendix A

Figures of Merit

In this chapter, we will review a variety of common figures of merit and derive the derivatives needed for the adjoint method. In most optimizations that we concern ourselves with in this dissertation, the figures of merit are built up from functions which can be separated into two categories: functions of the fields and explicit functions of the design parameters. The former set of functions are typically used to describe the performance of a device while the latter functions are typically used as *penalties* which constrain the device in some way. The way in which we leverage these different types of functions can have a large impact both on the final result we obtain from the optimization as well as the path the optimization takes to get to that final result. As such, carefully constructing a figure of merit is an essential step in setting up an optimization.

A.1 Functions of the Fields

The primary purpose of a figure of merit is to quantify the performance of a device. The performance of an electromagnetic device is typically expressed as a function of the electric and magnetic fields of the device. For example, our figure of merit may measure propagating power or stored energy density, both of which can be expressed as explicit functions of the electric and magnetic fields.

In this section, we will introduce a variety of different functions of the electric and magnetic fields that we frequently use in figures of merit when optimizing electromagnetic devices. In each case, we will derive the relevant physical equations and then follow the process described in Section 2.1.2 to derive the corresponding derivatives, $\partial F/\partial \vec{x}$, that we need in order to apply the adjoint method. The end results of these derivations can be used with the adjoint method to calculate gradients efficiently and accurately.

A.1.1 Power Transmission

In some optimizations, it is desirable to optimize the total power flowing through a plane. In most cases, we work with the frequency domain fields and thus we are concerned with the time-averaged power. In this case, the power is defined by the flux of the time-averaged Poynting vector and has the functional form

$$P_T = \iint_A dA\ \hat{n} \cdot \frac{1}{2}\mathrm{Re}\{\mathbf{E} \times \mathbf{H}^*\} \tag{A.1}$$

Here the integral is taken over a planar surface A whose normal direction is $\hat{n}$. Numerically, the simplest way to approximate this integral is using a Riemann sum:

$$P_T = \sum_i \sum_j \Delta A\ \hat{n} \cdot \frac{1}{2}\mathrm{Re}\{\mathbf{E}_{ij} \times \mathbf{H}^*_{ij}\} \tag{A.2}$$

where ΔA is the area occupied by a single field "sample" (which is related to how the fields are discretized). In order to differentiate this function with respect to the fields, it is useful to first explicitly write out the cross product and also expand the $\mathrm{Re}\{\ldots\}$ in terms of the Poynting vector and its complex conjugate:

$$\begin{aligned} P_T = \sum_i \sum_j \Delta A \Bigg[& \frac{1}{4}\, n_x \left(E_y H_z^* - E_z H_y^* + E_y^* H_z - E_z^* H_y\right) \\ & - n_y \left(E_x H_z^* - E_z H_x^* + E_x^* H_z - E_z^* H_x\right) \\ & + n_z \left(E_x H_y^* - E_y H_x^* + E_x^* H_y - E_y^* H_x\right) \Bigg]_{ij} \end{aligned} \tag{A.3}$$

where we the ij on the square bracket means that everything within the square brackets is evaluated at spatial index i, j. By explicitly writing the power flux as a function of the individual field components, we have made the process of calculating its derivative with respect to each field component relatively straightforward. In this case, the derivatives are

$$\begin{aligned} \frac{\partial P_T}{\partial E_{x,ij}} &= \frac{\Delta A}{4}(-n_y H^*_{z,ij} + n_z H^*_{y,ij}) & \frac{\partial P_T}{\partial H_{x,ij}} &= \frac{\Delta A}{4}(n_y E^*_{z,ij} - n_z E^*_{y,ij}) \\ \frac{\partial P_T}{\partial E_{y,ij}} &= \frac{\Delta A}{4}(n_x H^*_{z,ij} + n_z H^*_{x,ij}) & \frac{\partial P_T}{\partial H_{y,ij}} &= \frac{\Delta A}{4}(-n_x E^*_{z,ij} + n_z E^*_{x,ij}) \\ \frac{\partial P_T}{\partial E_{z,ij}} &= \frac{\Delta A}{4}(n_x H^*_{y,ij} + n_y H^*_{x,ij}) & \frac{\partial P_T}{\partial H_{z,ij}} &= \frac{\Delta A}{4}(n_x E^*_{y,ij} - n_y E^*_{x,ij}) \end{aligned} \tag{A.4}$$

In most cases of interest, this plane will align with a Cartesian direction; As a result, $\hat{n}$ will equal either $\hat{x}$, $\hat{y}$, or $\hat{z}$ and the integral will be performed over the

corresponding Cartesian plane. This greatly simplifies the integrand derivatives as well. For example, if the normal direction of our integral plane lies along the $\hat{x}$ direction, then the derivatives in Equation (A.4) simplify to

$$\begin{aligned}
\frac{\partial P_T}{\partial E_{x,ij}} &= 0 & \frac{\partial P_T}{\partial H_{x,ij}} &= 0 \\
\frac{\partial P_T}{\partial E_{y,ij}} &= \frac{\Delta A}{4} H^*_{z,ij} & \frac{\partial P_T}{\partial H_{y,ij}} &= -\frac{\Delta A}{4} E^*_{z,ij} \\
\frac{\partial P_T}{\partial E_{z,ij}} &= \frac{\Delta A}{4} H^*_{y,ij} & \frac{\partial P_T}{\partial H_{z,ij}} &= \frac{\Delta A}{4} n_x E^*_{y,ij}
\end{aligned} \tag{A.5}$$

A.1.2 Power Absorbed

Another common function of the electric and magnetic fields is the power absorbed by a lossy medium. From Poynting's theorem, we know that the electromagnetic power absorbed by a *non-dispersive*[1] isotropic lossy material is given by

$$P_{\text{abs}} = \frac{1}{2} \iiint_V dV \mathbf{J} \cdot \mathbf{E}^* = \frac{1}{2} \iiint_V dV \sigma \mathbf{E} \cdot \mathbf{E}^* \tag{A.6}$$

where σ is conductivity of the medium and V is the volume in which the absorbed power is calculated. You may be used to seeing an alternative form of this expression in which the absorbed power per unit volume is given by $\frac{1}{2}\text{Re}\{i\omega\tilde{\epsilon}\mathbf{E}\cdot\mathbf{E}^*\}$ where $\tilde{\epsilon}$ is the complex permittivity. These two expressions are made equivalent by noting that the imaginary part of the permittivity can be written in terms of the conductivity as $\epsilon^{''} = \sigma/\omega$ (see Appendix B for a short derivation). Numerically, we can approximate this continuous expression as

$$P_{\text{abs}} = \frac{1}{2} \sum_i \sum_j \sum_k \Delta V \sigma_{ijk} \mathbf{E}_{ijk} \cdot \mathbf{E}^*_{ijk} \tag{A.7}$$

where the fields are now distributed on a grid with indices ijk. This expression is relatively straightforward to differentiate with respect to the fields:

[1] For dispersive materials, the power stored in the fields and power absorbed from the fields is modified slightly. This modification only superficially modifies the results in this section.

$$\begin{aligned}
\frac{\partial P_{\text{abs}}}{\partial E_{x,ij,k}} &= \Delta V \sigma_{ijk} E^*_{x,ijk} \\
\frac{\partial P_{\text{abs}}}{\partial E_{y,ij,k}} &= \Delta V \sigma_{ijk} E^*_{y,ijk} \\
\frac{\partial P_{\text{abs}}}{\partial E_{z,ij,k}} &= \Delta V \sigma_{ijk} E^*_{z,ijk} \\
\frac{\partial P_{\text{abs}}}{\partial H_{x,ij,k}} &= \frac{\partial P_{\text{abs}}}{\partial H_{y,ij,k}} = \frac{\partial P_{\text{abs}}}{\partial H_{z,ij,k}} = 0
\end{aligned} \tag{A.8}$$

The power absorbed is a bit of a unique function in that it contains a direct dependence not only on the fields, but also on the design parameters. This is because the conductivity may be an explicit function of the design variables. As such, in order to calculate the gradient of a function involving power absorbed using the adjoint method, we will also need to compute the derivative of this function with respect to the design variables, i.e.,

$$\frac{\partial P_{\text{abs}}}{\partial \vec{p}} = \frac{1}{2} \sum_i \sum_j \sum_k \Delta V \frac{\partial \sigma_{ijk}}{\partial \vec{p}} \mathbf{E}_{ijk} \cdot \mathbf{E}^*_{ijk} \quad . \tag{A.9}$$

It is important to account for this derivative when implementing a figure of merit which contains the absorbed power.

A.1.3 Source Power Normalization

In many optimizations in nanophotonics, we seek to maximize or minimize a function which is normalized with respect to the total source power injected into the system. For example, when optimizing a waveguiding device, we typically seek to maximize the efficiency of the device which is defined as the fraction of input power which leaves in the desired output mode. In such devices, we might expect the source power to be independent of changes in the structure and thus assume that it is just a constant multiplicative factor in our figure of merit. In reality, the finite extent of our source current distribution can lead to small changes in source power in response to changes in the permittivity distribution. These small changes in source power can lead to potentially significant inaccuracies in the gradient if we do not properly account for the source power in our adjoint source. We can account for changes in the source power by treating it as function of the fields and taking derivatives as we would any other figure of merit.

Based on Poynting's theorem, the power injected into a system is given by

$$\begin{aligned} P_{\text{src}} &= \frac{1}{2}\iint\limits_{\partial V} dA\ \hat{n}\cdot \text{Re}\{\mathbf{E}\times\mathbf{H}^*\} + \frac{1}{2}\iiint\limits_{V} dV\sigma\mathbf{E}\cdot\mathbf{E}^* \\ &= P_T + P_{\text{abs}} \end{aligned} \tag{A.10}$$

where V is a volume which fully encompasses the source current distributions and ∂V is the surface of that volume. When numerically representing the fields on a rectangular grid, this expression can be rewritten once again in terms of Riemann sums as shown in Equations (A.2) and (A.7). In this case, the volume is a rectangular prism defined by three ranges of indices, $[i_1, i_2]$, $[j_1, j_2]$, and $[k_1, k_2]$. The surface of this volume is defined by six planes with unit normals pointing along the three Cartesian directions (in each case, one in the positive direction and one in the negative direction).

Typically, a power-normalized figure of merit will take the form

$$F(\vec{\mathbf{E}},\vec{\mathbf{H}}) = \frac{f(\vec{\mathbf{E}},\vec{\mathbf{H}})}{P_{\text{src}}} = \frac{f(\vec{\mathbf{E}},\vec{\mathbf{H}})}{P_T + P_{\text{abs}}} \tag{A.11}$$

where $f(\vec{\mathbf{E}},\vec{\mathbf{H}})$ is a function with units of power and where we use the vector symbol to denote that the fields are not only vector fields with $\hat{x}$, $\hat{y}$ and $\hat{z}$ components but also defined numerically as a set of vectors $\vec{E}_x$, $\vec{E}_y$, etc. We take the derivative of this function by applying the quotient rule as we would any other derivative:

$$\frac{\partial F}{\partial E_{x,ijk}} = \frac{1}{(P_T + P_{\text{src}})^2}\left[\frac{\partial f}{\partial E_{x,ijk}}(P_T + P_{\text{abs}}) - (\frac{\partial P_T}{\partial E_{x,ijk}} + \frac{\partial P_{\text{abs}}}{\partial E_{x,ijk}})f\right] \tag{A.12}$$

The derivatives with respect to the remaining field components have an identical form. In this expression, we assume $\partial f/\partial E_{x,ijk}$ is known. The derivatives $\partial P_T/\partial E_{x,ijk}$ and $\partial P_{\text{abs}}/\partial E_{x,ijk}$ are given by Equations (A.4) and (A.8), respectively. In the next section, we will discuss a very common example of power-normalized functions which relies on this derivative.

A.1.4 Mode Match Transmission

Mode overlap, or mode-matched efficiency, refers to the fraction of power in an incident field which resides in a desired mode of the system. Computing the mode overlap is essential to determining the efficiency of many optical devices and is therefore highly relevant to shape optimization. In this section, we will derive the mode match integral and its derivatives.

Our first step determining the mode-matched efficiency is to express our input field as a sum of the allowed propagating modes of the system. Specifically, the basis we will use consists of the electric and magnetic fields of both forward and

backward traveling waves. We begin by writing the electric field as a sum of these basis functions,

$$\mathbf{E} = \mathbf{E}_{\text{fwd}} + \mathbf{E}_{\text{back}} \tag{A.13}$$

$$= \sum_m \left(a_m \mathrm{e}^{ik_m z} + b_m \mathrm{e}^{-ik_m z} \right) \mathbf{E}_m \tag{A.14}$$

where the two terms in parentheses correspond to the forward and backward components and are given separately by

$$\mathbf{E}_{\text{fwd}} = \sum_m a_m \mathrm{e}^{ik_m z} \mathbf{E}_m \tag{A.15}$$

$$\mathbf{E}_{\text{back}} = \sum_m b_m \mathrm{e}^{-ik_m z} \mathbf{E}_m \quad . \tag{A.16}$$

In the expressions above, $\mathbf{E}_m$ is the electric field profile of the mth mode of the system. The magnetic field, meanwhile, can be written in a similar form. Applying Faraday's Law and assuming harmonic time dependence of the electric field, the magnetic field can be written as an expansion of the forward and backward wave

$$\mathbf{H} = \mathbf{H}_{\text{fwd}} + \mathbf{H}_{\text{back}} \tag{A.17}$$

$$= \sum_m \left(a_m \mathrm{e}^{ik_m z} - b_m \mathrm{e}^{-ik_m z} \right) \mathbf{H}_m \tag{A.18}$$

Notice in Eq. (A.18), that the backward propagating term is preceded by a negative sign. This arises out of the requirement that power flow in the negative direction (and hence the Poynting vector, $\mathbf{E}_m \times \mathbf{H}_m$ point in the negative direction). In both Eq. (A.18) and (A.14), we have chosen to write the fields as a sum of the forward and backward traveling waves. In general, given an arbitrary field, we will not know the forward and backward traveling components but only their sum. Our goal now is to develop the machinery needed to separate the different forward and backward traveling modes which compose an arbitrary field.

This is equivalent to finding the coefficients a_m and b_m. To do so, we must take advantage of the orthogonality condition of our electromagnetic basis which arises as a result of Lorentz reciprocity and is given by [92]

$$\frac{\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m \times \mathbf{H}_n^*}{\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m \times \mathbf{H}_m^*} = \delta_{mn} \tag{A.19}$$

where δ_{mn} is the Kronecker delta. We apply this orthogonality condition by computing the surface integral of $\mathbf{E} \times \mathbf{H}_m^*$ and rearranging terms yields an expression relating a_m and b_m to the electric field and mth mode

$$a_m \mathrm{e}^{ik_m z} + b_m \mathrm{e}^{-ik_m z} = \frac{\iint\limits_A d\mathbf{A} \cdot \mathbf{E} \times \mathbf{H}_m^*}{S_m} \tag{A.20}$$

where S_m is related to the power propagating in the mth mode

$$S_m = \iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m \times \mathbf{H_m}^* \quad . \tag{A.21}$$

A second expression for a_m and b_m can be found by computing the surface integral of $\mathbf{E}_m \times \mathbf{H}^*$ which yields

$$a_m \mathrm{e}^{ik_m z} - b_m \mathrm{e}^{-ik_m z} = \frac{\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m^* \times \mathbf{H}}{S_m^*} \quad . \tag{A.22}$$

With two equations and two unknowns, we are now able to solve for the coefficients. Adding the two equations produces an expression for a_m

$$a_m = \frac{1}{2} \mathrm{e}^{-ik_m z} \left(\frac{\iint\limits_A d\mathbf{A} \cdot \mathbf{E} \times \mathbf{H}_m^*}{S_m} + \frac{\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m^* \times \mathbf{H}}{S_m^*} \right) \tag{A.23}$$

while subtracting them yields and expression for b_m. Using these equations, decomposing a field $\mathbf{E}$ and $\mathbf{H}$ into the modes of the system is a straightforward calculation. Mode matching requires that we take this one step further and determine how much *power* is propagating in a desired mode. To find this, we compute the power propagating through a plane in the forward wave, i.e. $P_{\mathrm{fwd}} = \frac{1}{2}\mathrm{Re}\{\iint d\mathbf{A} \cdot \mathbf{E}_{\mathrm{fwd}} \times \mathbf{H}_{\mathrm{fwd}}^*\}$. This calculation results in a sum whose terms correspond to the power propagating in each mode. This power is more conveniently expressed as a fraction of the total power propagating in the field:

$$\eta_m^{\mathrm{fwd}} = \left.\frac{P_m}{P_{\mathrm{in}}}\right|_{\mathrm{fwd}} = |a_m|^2 \frac{\mathrm{Re}\left\{S_m\right\}}{\mathrm{Re}\left\{\iint\limits_A d\mathbf{A} \cdot \mathbf{E} \times \mathbf{H}^*\right\}} \quad . \tag{A.24}$$

Eq. (A.24) describes the fraction of power in an incident field that is propagating in a desired mode of the system. In most cases, it is desirable to know the fraction of *total* injected power which is contained in a desired mode. In order to get this, we simply substitute P_{in} for the source power P_{src},

$$\eta_m^{\text{fwd}} = |a_m|^2 \frac{P_m}{P_{\text{src}}} \quad . \tag{A.25}$$

where $P_m = \frac{1}{2}\text{Re}\{S_m\}$ is the power in the source fields.

This expression can be further simplified by noticing that in most problems, a backward traveling wave is not present at the output of the system. In this case, b_m equals zero and as a result

$$\frac{\iint\limits_A d\mathbf{A} \cdot \mathbf{E} \times \mathbf{H}_m^*}{S_m} = \frac{\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m^* \times \mathbf{H}}{S_m^*} \tag{A.26}$$

Taking this into account allows us to simplify Eq. (A.24) to

$$\eta_m = \frac{\frac{1}{2}\text{Re}\left\{\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m \times \mathbf{H}_m^*\right\}}{\frac{1}{2}\text{Re}\left\{\iint\limits_A d\mathbf{A} \cdot \mathbf{E} \times \mathbf{H}^*\right\}} \frac{\left|\iint\limits_A d\mathbf{A} \cdot \mathbf{E} \times \mathbf{H}_m^*\right|^2}{\left|\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m \times \mathbf{H}_m^*\right|^2} \tag{A.27}$$

where $\mathbf{E}_m$ and $\mathbf{H}_m$ are the field profiles of the desired mode. This equation is the most general form of the mode-matched efficiency. We can further simplify the expression by noting that for a guided or free space mode, the integral $\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m \times \mathbf{H}_m^*$ is real valued. In this case we can cancel the first term in the numerator and we find that the mode overlap is given by

$$\eta_{m,\text{guided}} = \frac{1}{4P_m P_{\text{in}}} \left|\iint\limits_A d\mathbf{A} \cdot \mathbf{E} \times \mathbf{H}_m^*\right|^2 \tag{A.28}$$

where we have chosen to write

$$P_m = \frac{1}{2}\text{Re}\left\{\iint\limits_A d\mathbf{A} \cdot \mathbf{E}_m \times \mathbf{H}_m^*\right\}$$

$$P_{\text{in}} = \frac{1}{2}\text{Re}\left\{\iint\limits_A d\mathbf{A} \cdot \mathbf{E} \times \mathbf{H}^*\right\}$$

which describe the power in the incident field and the power in the desired mode (neither of which are guaranteed to be normalized to unity power).

For the sake of flexibility, we will typically employ the more general Equation (A.25) in our figures of merit. The derivatives with respect to the different field components are a bit messy. To keep things organized, it is convenient to write

the derivatives in terms of derivatives of the quantities a_m and P_{src}. Given a field component $\psi \in \{E_x, E_y, E_z, H_x, H_y, H_z\}$, the derivative of the mode match equation is

$$\frac{\partial \eta_m^{\text{fwd}}}{\partial \psi} = P_m \frac{\frac{\partial a_m}{\partial \psi} a_m^* P_{\text{src}} - \frac{\partial P_{\text{src}}}{\partial \psi} |a_m|^2}{P_{\text{src}}^2} \tag{A.29}$$

where we have used the fact that $\partial a_m^* / \partial \psi = 0$ for all of the field components (which depends only on $\mathbf{E}^*$ and $\mathbf{H}^*$). Before calculating the derivatives of a_m, we first need to express a_m in terms of discrete sums that we can calculate numerically. Once again, for simplicity we use a Riemann sum:

$$a_m = \frac{1}{2} \mathrm{e}^{-ik_m z} \left(\frac{\sum_i \sum_j \Delta \mathbf{A} \cdot \mathbf{E}_{ij} \times \mathbf{H}_{m,ij}^*}{S_m} + \frac{\sum_i \sum_j \Delta \mathbf{A} \cdot \mathbf{E}_m^* \times \mathbf{H}}{S_m^*} \right) \tag{A.30}$$

The derivatives of this function with respect to the different field components are given by

$$\begin{aligned}
\frac{\partial a_m}{\partial E_{x,ij}} &= \frac{1}{2S_m} e^{-ik_m z} (-n_y H_{m,z,ij}^* + n_z H_{m,y,ij}^*) \\
\frac{\partial a_m}{\partial E_{y,ij}} &= \frac{1}{2S_m} e^{-ik_m z} (n_x H_{m,z,ij}^* + n_z H_{m,x,ij}^*) \\
\frac{\partial a_m}{\partial E_{z,ij}} &= \frac{1}{2S_m} e^{-ik_m z} (n_x H_{m,y,ij}^* + n_y H_{m,x,ij}^*) \\
\frac{\partial a_m}{\partial H_{x,ij}} &= \frac{1}{2S_m^*} e^{-ik_m z} (n_y E_{m,z,ij}^* - n_z E_{m,y,ij}^*) \\
\frac{\partial a_m}{\partial H_{y,ij}} &= \frac{1}{2S_m^*} e^{-ik_m z} (-n_x E_{m,z,ij}^* + n_z E_{m,x,ij}^*) \\
\frac{\partial a_m}{\partial H_{z,ij}} &= \frac{1}{2S_m^*} e^{-ik_m z} (n_x E_{m,y,ij}^* - n_y E_{m,x,ij}^*)
\end{aligned} \tag{A.31}$$

where n_x, n_y, and n_z are the components of the unit normal vector $\hat{n}$ of the integration surface. The derivative of the source power, meanwhile, is discussed in detail in the previous section. Plugging these results into Equation (A.29) yields the desired derivatives with respect to the fields.

A.2 Penalty Functions

In this work, we primarily focus on shape optimization, choosing to represent material boundaries using polygons. Polygons are very flexible as they allow us to represent almost arbitrary shapes (as long as we have enough points in that polygon)

and are relatively easy to work with. Furthermore, because polygons are inherently geometric, it is very easy to constrain their shapes and sizes. Such constraints are an essential part of designing devices which can be fabricated.

An easy way to implement these constraints is to introduce a *penalty function* to the figure of merit. Penalty functions are additions to the figure of merit, i.e.,

$$F(\vec{\mathbf{E}}, \vec{\mathbf{H}}, \vec{p}) = f(\vec{\mathbf{E}}, \vec{\mathbf{H}}) + \mathcal{P}(\vec{\mathbf{E}}, \vec{\mathbf{H}}, \vec{p}) \tag{A.32}$$

which either increase or decrease its value when certain criteria are or are not met. For example, if we wish to maximize our figure of merit which is the coupling efficiency to a desired output mode, we could add a penalty function which is negative when a constraint is violated. This would cause the figure of merit to decrease, forcing the optimization to navigate towards a different region of the design space.

This method of imposing constraints on our devices has the great advantage that the constraints are directly incorporated into the optimization. The disadvantage of these penalty functions, on the other hand, is that they do not guarantee that the constraints are exactly satisfied. In practice, however, by weighting the penalty functions sufficiently, we can be reasonably certain that the constraints are not violated.

In this section, we will review common penalty functions which are explicit functions of the design variables. In each case, we will derive the function and, when practical, its derivatives.

A.2.1 Threshold Functions

A key component in most penalty functions is a threshold. This threshold is typically implemented using two different functions: a step function or a rect function. In cases in which a single “less than" or “greater than" constraint is required, we can use a step function which is zero below a threshold value and equal to one above a threshold value. In case where a “greater than X and less than Y" constraint is required, we can use a rect function which is equal to one between two threshold values and zero everywhere else.

In order to use threshold-based penalty function in our figures of merit, we need analytic and smooth approximations of the step and rect functions and their derivatives. These can be accomplished in a variety of ways. In our case, we will draw inspiration from the Fermi-Dirac distribution which is effectively an analytic approximation of a step function. In particular, we can define a step function as

$$H(x) = \frac{1}{1 + e^{-kx}} \tag{A.33}$$

where k determines the steepness of the step. Notice that as $x \to \infty$ $H(x) \to 1$ and as $x \to -\infty$ $H(x) \to 0$ which is the desired behavior. The derivative of this function, meanwhile, is

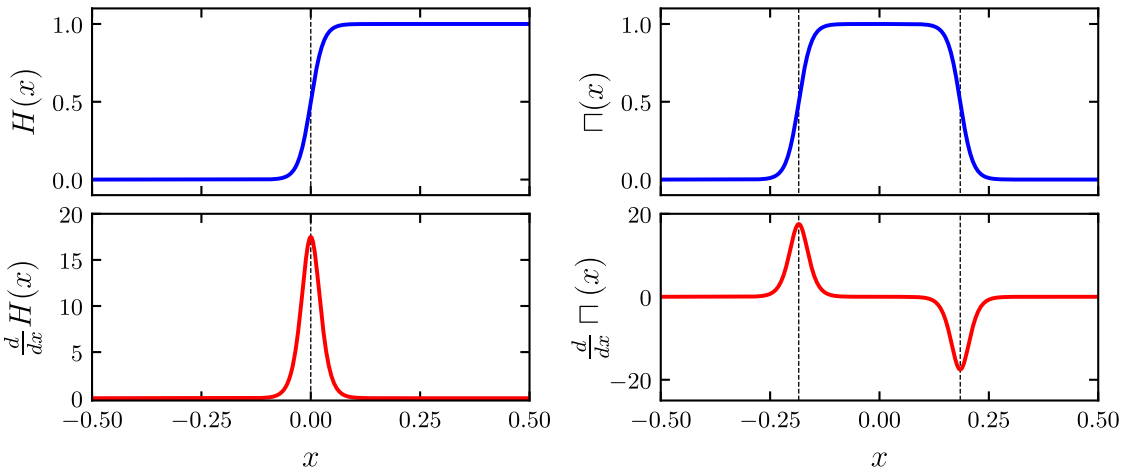

Figure A.1: Plot of examples of analytic approximation of a step and rect function (top) and their first derivatives (bottom).

$$\frac{dH}{dx} = \frac{ke^{-kx}}{(1+e^{-kx})^2} \tag{A.34}$$

By controlling the value of k we can control how steep the step function is and how sharp its first derivative is. Typically, when incorporating thresholds into our figure of merit, we want to avoid making the step too sharp as it can lead to convergence issues, especially if the threshold is initially violated.

In order to implement an analytic approximation of the rect function, we can simply add two step functions,

$$\sqcap(x) = H(x+\frac{1}{2}) - H(x-\frac{1}{2}) \tag{A.35}$$

which defines a rect function with a width of 1. Similarly, we can write the derivative of the rect function in terms of the derivatives of the step function:

$$\frac{d\sqcap}{dx} = \frac{dH}{dx}\bigg|_{x+1/2} - \frac{dH}{dx}\bigg|_{x-1/2} \tag{A.36}$$

Examples of the step and rect functions and their first derivatives are shown in Figure A.1. In both cases, $k = 70$ which leads to a relatively sharp step. As desired, the functions and their derivatives remain smooth, which is essential to the optimization process.

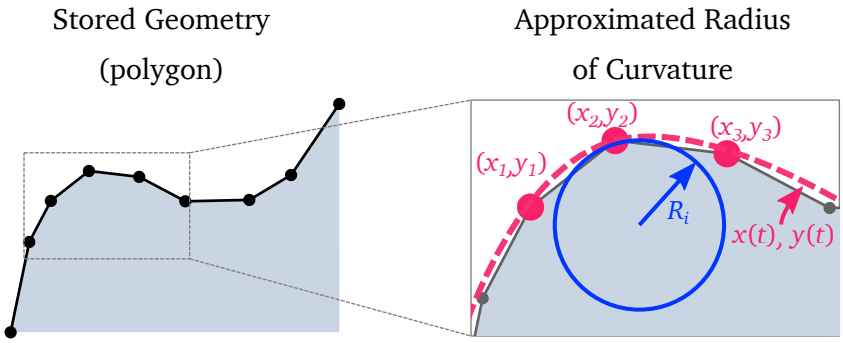

Figure A.2: Diagram showing the process of calculating the approximate radius of curvature of a polygon. A parametric quadratic curve is fit to sets of three points. The curvature of this parametric curve is solved for and used to represent the approximate radius of curvature of the polygon.

A.2.2 Radius of Curvature

One of the most important class of constraints when optimizing nanophotonic devices is feature size constraints. In general, when fabricating nanophotonic components at any scale, we use deep UV lithography which cannot produce infinitesimally small structures. In order to ensure that the devices we design can be fabricated, it is thus essential that we impose constraints which limit the formation of device features that are too small. One common example of this constraint is the *radius of curvature constraint* which limits the curvature of the boundaries of our structure.

It turns out to be relatively straightforward to impose radius of curvature constraints in optimization problems in which boundaries are defined using polygons. We do this by calculating an approximate curvature at each point in the polygon and then applying a smooth analytic thresholding function described in the previous section which reduces the figure of merit when the radius of curvature is below the chosen minimum value. This soft constraint does not guarantee that the approximate radius of curvature at every point in the polygon will be larger than the minimum value, however in practice the penalty can be made significant enough that small curvature will not appear.

In general, the curvature of a polygon is ill-defined since it is made up of straight segments. However, intuitively we know that with a sufficient number of points, a polygon can resemble a smooth curve. We should thus be able to define an effective radius of curvature that reflects the apparent smoothness. We accomplish this by fitting higher order polynomials to sets of points in the polygon and then calculating the curvature of those curves at each point as depicted in Figure A.2. An easy implementation of this idea is to fit quadratic parametric curves to sets of three points and compute the curvature at the middle point. These parametric curves

take the form

$$x(t) = at^2 + bt + c \tag{A.37}$$
$$y(t) = dt^2 + et + f \tag{A.38}$$

where a-f are coefficients that we need to solve for using the coordinates of the points in the polygon. For simplicity, we choose the t to range from 0 at (x_1, y_1), to 0.5 at (x_2, y_2), and terminate at 1.0 at (x_3, y_3). Based on this choice, it is relatively easy to solve for the coefficients:

$$\begin{aligned} a &= \frac{1}{4}\left[x_1 + \frac{1}{2}(x_3 - x_1) - x_2\right] \\ b &= x_3 - x_1 - a \\ c &= x_1 \\ d &= \frac{1}{4}\left[y_1 + \frac{1}{2}(y_3 - y_1) - y_2\right] \\ e &= y_3 - y_1 - d \\ f &= y_1 \end{aligned} \tag{A.39}$$

Given a solution to the parametric curve which fits three points of the polygon, we can approximate the radius of curvature of the polygon of the middle point (x_2, y_2) by evaluating the radius of curvature of the quadratic curve. The radius of curvature of an arbitrary parametric curve is given by

$$R = \frac{(x'^2 + y'^2)^{3/2}}{x'y'' - y'x''} \tag{A.40}$$

where x' and x'' denote the first and second derivatives of x with respect to t. Differentiating the equations in (A.38), we find that the radius of curvature at the middle point is

$$R_2 \approx \frac{[(a+b)^2 + (d+e)^2]^{3/2}}{2d(a+b) - 2a(d+e)} \tag{A.41}$$

We can repeat this same process for every point in the polygon by letting $x_1 = x_{i-1}$, $x_2 = x_i$, and $x_3 = x_{i+1}$ where i is the index of each point in the polygon.

The radius of curvature of a circle calculated using this process is shown in Figure A.3. When the circle is defined using fewer points, there is some visible error in the estimated radius of curvature, which is to be expected. As the number of points increases, the radius converges quickly to the desired value. It is important to note that this process generally underestimates the radius of curvature. From the

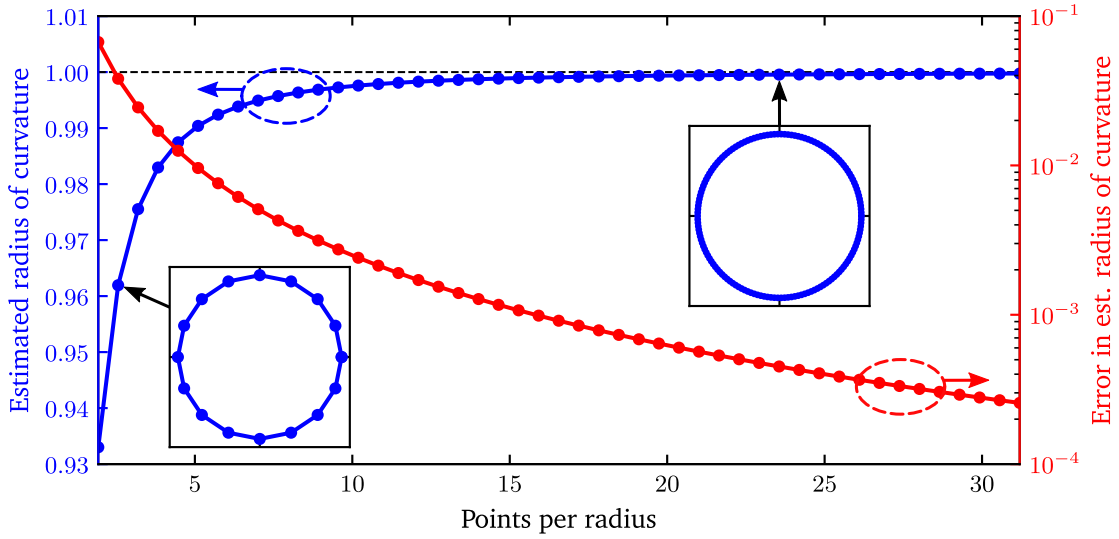

Figure A.3: Plot of the estimated radius of curvature of a circle with a radius of 1 (arb. units) versus the number of points used to represent the circle (normalized with respect to the radius of the circle). Even for smaller numbers of points, the error is reasonably low and in general the radius is underestimated (which is preferred for the purpose of imposing optimization constraints).

standpoint of imposing *minimum* radius of curvature constraints, this is preferable to an overestimate.

With the radius of curvature now known, we can construct a penalty function which reduces the figure of merit when the radius of curvature falls below the minimum value by applying a thresholding function that we discussed in the previous section. Note that the estimated radius of curvature in Equation (A.41) can go negative. As a result, it is best to use a smooth rect function to define a positive and negative threshold. In particular, given a minimum radius of curvature $R_{\min}$, our penalty function will be

$$\mathcal{P}_{\text{roc}}(\vec{p}) = -\alpha_{\text{roc}} \sum_i \sqcap \left(\frac{R_i(\vec{p})}{2R_{\min}} \right) \tag{A.42}$$

where $R_i(\vec{p})$ is the radius of curvature at each point in the polygon and α_{roc} is the minimum radius of curvature. Notice that if the estimated radius of curvature at a point in the polygon is between $-R_{\min}$ and $R_{\min}$, Equation (A.42) will drop to $-\alpha_{\text{roc}}$. Adding this function to a figure of merit will penalize that figure of merit when bumps in the boundary become too sharp, as desired. In many cases, the field-dependent part of the figure of merit will relate to an efficiency and hence its value will range from 0 to 1. With this in mind, we can typically set the value of α_{roc} to around 0.01, which is equivalent to penalizing the figure of merit by 1%

efficiency for each point which violates the radius of curvature constraint. If a figure of merit is not a measure of efficiency, however, choosing this value appropriately may take some experimentation.

In addition to defining the penalty function, we also need its derivative. In this case, the derivative is rather messy and laborious to compute. Fortunately, we can can compute it using symbolic math libraries like sympy[2] which will automatically generate the code needed for the calculation. This reduces the likelihood of human error.

A.2.3 Minimum Bridge and Gap Size

Due to the finite resolution of existing fabrication methods in nanophotonics, it is often desirable to limit the sizes of bridges and gaps in a structure. Similar to the radius of curvature, we can limit these sizes by imposing a penalty function which consists of a threshold applied to a calculated gap width or bridge width at each point in the polygon.

The process of calculating these widths is relatively straightforward for structures defined using polygons. First, we will calculate the approximate normal direction at a point in the polygon. Next, we will calculate the distance from that point to any other boundaries of the polygon (or other polygons) which intersect a line extending from that point along its normal direction. These distances are the bridge and gap widths.

Our first step is to calculate the normal direction at each point in the polygon. Let the vertex of interest be (x_2, y_2) and its adjacent points be (x_1, y_1) and (x_3, y_3) which are arranged in a counter-clockwise manner. The normal direction for this point is approximately equal to

$$\begin{aligned}\hat{n}_2 &= \frac{1}{\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2}} \left[-(y_3 - y_1)\hat{x} + (x_3 - x_1)\hat{y}\right] \\ &= n_x\hat{x} + n_y\hat{y}\end{aligned} \tag{A.43}$$

Having found the normal direction at our vertex of interest, it is straightforward to define the line which extends along the vertex's normal direction and intersects the vertex

$$x(t) = n_x t + x_2 \tag{A.44}$$

$$y(t) = n_y t + y_2 \tag{A.45}$$

which we refer to as the "normal line." Here, we have chosen to write the line parametrically in terms of the parameter t. At $t = 0$, this line intersects the point

[2] https://www.sympy.org/en/index.html

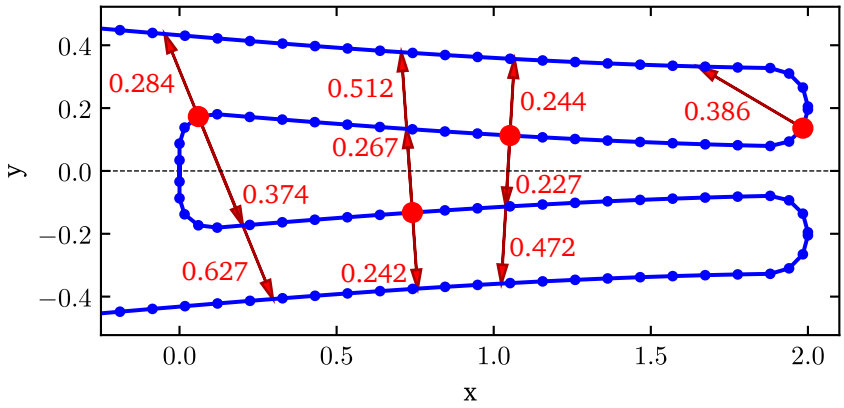

Figure A.4: Demonstration of gap and bridge width calculation for polygons. The gap and bridge widths are calculated for a selected set of points (red dots) in a polygon (blue lines and dots). These widths are marked using the red arrows and the accompanying numerical values.

(x_2, y_2). Our goal now is to determine where else in this polygon the line intersects. To do so, we consider each pair of adjacent points in our polygon and for that pair, we find the parametric line which connects the two points. Given two points (x_j, y_j) and (x_{j+1}, y_{j+1}), the parametric line connecting them is given by

$$x_j(u) = \Delta x_j u + x_{j0} \tag{A.46}$$
$$y_j(u) = \Delta y_j u + y_{j0} \tag{A.47}$$

where $\Delta x_j = x_{j+1} - x_j$ and $\Delta y_j = y_{j+1} - y_j$. Notice that the line extends from point j to point $j+1$ as u ranges from 0 to 1. This makes it very easy to determine if our normal line intersects this line between the two points. In order to find where the normal line intersects the parametric lines connecting each pair of points in the polygon, we simply equate $x(t) = x_j(u)$ and $y(t) = y_j(u)$ and solve for t and u. Do this, we find that

$$t = \frac{\Delta x_j (y_2 - y_{j0}) - \Delta y_j (x_2 - x_{j0})}{\Delta y_j n_x - \Delta x_j n_y} \tag{A.48}$$

and

$$u = \frac{n_x (y_{j0} - y_2) - n_y (x_{j0} - x_2)}{n_y \Delta x_j - n_x \Delta y_j} \quad . \tag{A.49}$$

Calculating u, we can determine if the normal line intersects the line segment connecting the two other points of the polygon. If the two lines do intersect, we can find the distance between our vertex of interest and the intersection:

$$d_j = \sqrt{[x_j(u) - x(t)]^2 + [y_j(u) - y(t)]^2} \tag{A.50}$$

This gives us the width of the gap or bridge in our structure. In order to determine whether the feature is a gap or bridge, we need only look at the sign of t. If t is positive, then the line exits the polygon before intersecting another edge of the polygon. This means that the distance in Equation (A.50) corresponds to a gap. If the sign of t is negative, this means that the normal line crosses through the interior of the polygon, and therefore the distance corresponds to a gap.

A demonstration of this process is shown in Figure A.4. The red dots correspond to the points at which the gap and bridge widths are evaluated. The method is proves very robust and yields consistent results for every point in the polygon.

With the gap and bridge widths calculated, we can define a minimum gap/bridge width penalty function. We want our penalty function to reduce the figure of merit when the gap/bridge width d_j drops below a minimum width $d_{\min}$. To accomplish this, we use our analytic step function approximation,

$$\mathcal{P}_{\mathrm{gb}}(\vec{p}) = -\alpha_{\mathrm{gb}} \sum_j H(d_{\min} - d_j) \tag{A.51}$$

where gb refers to"gap and bridge" and α_{gb} is the weight of the penalty function.

As we have defined the process, the end result is unfortunately not differentiable in some cases, which makes it potentially dangerous to use in an optimization. In particular, if an edge is displaced such that it intersects a new normal line, the function will be discontinuous. This is a result of how we filtered the edge intersections based on the value of u. To get around this, we can instead calculate the distance to every edge line irrespective of whether or not the normal line intersects between the two points that define each edge. Next, we modify the penalty function to effectively "ignore" contributions which have $u < 0$ and $u > 1$:

$$\mathcal{P}_{\mathrm{gb}}(\vec{p}) = -\alpha_{\mathrm{gb}} \sum_j H(d_{\min} - d_j) \sqcap (u_j - 1/2) \tag{A.52}$$

Here u_j is the parameter given in Equation (A.49). Notice that when u_j is not between 0 and 1, the distance does not contribute to the penalty function as desired. This function is differentiable in all cases except when a normal line and edge start out parallel but are modified such that they are no longer parallel. In this case, however, the $H(d_{\min} - d_j)$ value should remain remain near 0 for even modest changes to the edge. As a result, the function should be nearly differentiable even in this edge case.

Appendix B

Nondimensionalized Maxwell's Equations

Normally, the magnitude of $\mathbf{E}$ and $\mathbf{H}$ differ by orders of magnitude. This difference can lead to poor conditioning of the discretized system of equations which are used to numerically solve Maxwell's equations using frequency domain methods. As a result, it is computationally advantageous to work with a dimensionless version of Maxwell's equations whose corresponding electric and magnetic fields are similar in magnitude. Even when solving Maxwell's equations using matrix-free method like the finite difference time domain method, the dimensionless equations can simplify the process of setting up and running simulations. In this section, we will briefly derive a dimensionless form of Maxwell's equations.

To derive the nondimensionalized Maxwell's equations, we begin with the dimensionalized time-dependent Maxwell's Equations,

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{M} \tag{B.1}$$

$$\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} - \sigma \mathbf{E} = \mathbf{J} \tag{B.2}$$

where $\mathbf{M}$ is the magnetic current density which we have included in order to symmetrize the equations. For the sake of simplicity, let us consider the specific case in which the material properties are non-dispersive. In this case, we can rewrite Maxwell's Equations as

$$\nabla \times \mathbf{E} + \mu \frac{\partial \mathbf{H}}{\partial t} = \mathbf{M} \tag{B.3}$$

$$\nabla \times \mathbf{H} - \varepsilon \frac{\partial \mathbf{E}}{\partial t} - \sigma \mathbf{E} = \mathbf{J} \tag{B.4}$$

In order to rewrite these equations in a dimensionless form, we first write nondimensionalized forms of the variables in terms of the original dimensional variables. We do this by separating the dimensional quantities into a constant factor which has units and a second unitless quantity that possesses any spatial or time dependence. These dimensionless quantities are:

$$\begin{array}{llll} \mathbf{r} = r_0\mathbf{r}' & t = \frac{t'}{\omega_0} & \mathbf{E} = E_0\mathbf{E}' & \mathbf{H} = H_0\mathbf{H}' \\ \mathbf{J} = J_0\mathbf{J}' & \mathbf{M} = M_0\mathbf{M}' & \mu = \mu_0\mu_r & \varepsilon = \varepsilon_0\varepsilon_r \end{array} \tag{B.5}$$

Here, E_0, H_0, M_0, J_0, r_0, and ω_0 are the as-of-yet unknown normalization factors. The constants μ_0 and ε_0, meanwhile, are the vacuum permeability and permittivity and are known. Substituting these new dimensionless quantities into Equation (B.4) yields:

$$\frac{E_0}{r_0}\nabla' \times \mathbf{E}' - \mu_0\mu_r H_0\omega_0\frac{\partial \mathbf{H}'}{\partial t'} = M_0\mathbf{M}' \tag{B.6}$$

$$\frac{H_0}{r_0}\nabla' \times \mathbf{H}' + \epsilon_0\epsilon_r E_0\omega_0\frac{\partial \mathbf{E}'}{\partial t'} - \sigma E_0\mathbf{E}' = J_0\mathbf{J}' \tag{B.7}$$

Next, we isolate the curl quantities by multiply across by E_0/r_0 in the top equation and H_0/r_0 in the bottom equation, which produces

$$\nabla' \times \mathbf{E}' - \frac{\mu_0\mu_r H_0 r_0\omega_0}{E_0}\frac{\partial \mathbf{H}'}{\partial t'} = \frac{M_0 r_0}{E_0}\mathbf{M}' \tag{B.8}$$

$$\nabla' \times \mathbf{H}' + \frac{\epsilon_0\epsilon_r E_{0r_0}\omega_0}{H_0}\frac{\partial \mathbf{E}'}{\partial t'} - \frac{\sigma E_0 r_0}{H_0}\mathbf{E}' = \frac{J_0 r_0}{H_0}\mathbf{J}' \tag{B.9}$$

Notice that these isolated curl quantities are now dimensionless since they are made up of our new dimensionless quantities. It follows that the constant coefficients of the remaining terms are also unitless. Furthermore, because we have yet to assign values to any of the normalization factors, we are free to assign an arbitrary value to these coefficients. For convenience, we make them all equal to one, i.e.,

$$\frac{\mu_0 H_0 r_0\omega_0}{E_0} = \frac{\mu_0 E_0 r_0\omega_0}{H_0} = \frac{r_0 M_0}{E_0} = \frac{r_0 J_0}{H_0} = 1 \tag{B.10}$$

This leaves use with four equations for the normalization factors but a total of six unknowns. This tells us that we are free to arbitrarily choose the values of two of the unknowns and the remaining unknowns will fall out of these expressions. For convenience, we will choose to set $\omega_0 = c/k_0$ where c is the speed of light and $k_0 = 2\pi/\lambda_0$ where λ_0 is a characteristic wavelength of the problem we are trying to solve. Furthermore, we will assume E_0 has a pre-defined value which ultimately

will specify the amount of power in the system. Based on these two choices, the remaining quantities are given by

$$r_0 = \frac{1}{\sqrt{\varepsilon_0\mu_0}\omega} = \frac{\lambda_0}{2\pi} \tag{B.11}$$

$$H_0 = \sqrt{\frac{\varepsilon_0}{\mu_0}}E_0 = \frac{E_0}{\eta_0} \tag{B.12}$$

$$M_0 = \sqrt{\varepsilon_0\mu_0}\omega_0 E_0 \tag{B.13}$$

$$J_0 = \varepsilon_0\omega_0 E_0 \tag{B.14}$$

where η_0 is the impedance of free space. Notice that there is one quantity we have not yet dealt with: the conductivity. A convenient way of handling this is to relate it back to the imaginary part of the permittivity using the Maxwell-Ampere law assuming a harmonic time dependence ($\mathbf{E} \sim \exp(-i\omega_0 t)$):

$$\begin{aligned}\nabla \times \mathbf{H} &= \sigma\mathbf{E} - i\omega_0\epsilon\mathbf{E} \\ &= -i\omega_0(\epsilon - i\frac{\sigma}{\omega_0})\mathbf{E} \\ &= -i\omega_0\epsilon_0(\epsilon_r - i\frac{\sigma}{\epsilon_0\omega_0})\mathbf{E}\end{aligned} \tag{B.15}$$

We can interpret this quantity in parentheses as the relative complex permittivity, i.e.,

$$\tilde{\varepsilon}_r = \epsilon_r^{'} - i\epsilon_r^{''} = \epsilon_r - i\frac{\sigma}{\epsilon_0\omega_0} \quad . \tag{B.16}$$

Based on this, we can redefine the conductivity in terms of the imaginary part of the relative permittivity:

$$\sigma = \omega_0\varepsilon_0\varepsilon_r^{''} \tag{B.17}$$

Substituting this result along with our normalization factors into Equations (B.8) and (B.9), we can write down the nondimensionalized Maxwell's equations:

$$\nabla^{'} \times \mathbf{E}^{'} + \mu_r\frac{\partial \mathbf{H}^{'}}{\partial t^{'}} = \mathbf{M}^{'} \tag{B.18}$$

$$\nabla^{'} \times \mathbf{H}^{'} - \epsilon_r^{''}\mathbf{E} - \varepsilon_r^{'}\frac{\partial \mathbf{E}^{'}}{\partial t^{'}} = \mathbf{J}^{'} \tag{B.19}$$

Given solution to the dimensionless $\mathbf{E}^{'}$ and $\mathbf{H}^{'}$, we can define a mapping between these dimensionless quantities and the "real world" dimensional quantities:

$$\mathbf{E} = E_0\mathbf{E}^{'} \tag{B.20}$$

$$\mathbf{H} = \sqrt{\frac{\varepsilon_0}{\mu_0}}E_0\mathbf{H}^{'} \tag{B.21}$$

$$\mathbf{J} = \varepsilon_0\omega E_0\mathbf{J}^{'} \tag{B.22}$$

$$\mathbf{M} = \sqrt{\varepsilon_0\mu_0}\omega E_0\mathbf{M}^{'} \tag{B.23}$$

It is worth noting that in many situations we only care about computing normalized powers. In such cases, the dimensionless quantities are sufficient on their own.

In some cases (for example when numerically solving the equations in two dimensions) it is convenient to work with the time-harmonic form of Equations (B.18) and (B.19). These equations are easily acquired by assuming the electric and magnetic fields have a time dependence given by $\exp(-i\omega_0 t) = \exp(-i\omega_0 t^{'}/\omega_0) = \exp(-it^{'})$. In this case, we find that the non-dimensional time-harmonic form of Maxwell's equations are simply:

$$\nabla^{'} \times \mathbf{E}^{'} - i\mu_r\mathbf{H}^{'} = \mathbf{M}^{'} \tag{B.24}$$

$$\nabla^{'} \times \mathbf{H}^{'} + i\epsilon_r\mathbf{E}^{'} = \mathbf{J}^{'} \tag{B.25}$$

This resembles the normal form of the time-harmonic Maxwell's equations with the exception that there is no $\omega_0\varepsilon_0$ multiplying $\mathbf{E}$ and $\mathbf{H}$.

Appendix C

Extending CW-FDTD Method to Multiple Wavelengths

In our formulation of the finite difference time domain method, we excite the simulation using a ramped continuous wave source (i.e., single wavelength) and then extract the amplitude and phase of the oscillating fields once they have stabalized by fitting the field at each point in the domain with a sinusoid. This process can easily be extended to multiple simultaneous wavelengths.

Our starting point is to rewrite Equations (3.13)-(3.15) to include multiple sinusoids corresponding to a sum of multiple ramped CW sources:

$$E(t_1) = A_1 \sin(\omega_1 t_1 + \varphi_1) + A_2 \sin(\omega_2 t_1 + \varphi_2) + \cdots + c \tag{C.1}$$
$$E(t_2) = A_1 \sin(\omega_1 t_2 + \varphi_1) + A_2 \sin(\omega_2 t_2 + \varphi_2) + \cdots + c \tag{C.2}$$
$$E(t_3) = A_1 \sin(\omega_1 t_3 + \varphi_1) + A_3 \sin(\omega_2 t_3 + \varphi_2) + \cdots + c \tag{C.3}$$
$$\cdots \tag{C.4}$$
$$E(t_M) = A_1 \sin(\omega_M t_M + \varphi_1) + A_3 \sin(\omega_M t_M + \varphi_2) + \cdots + c \tag{C.5}$$

As it is written, we have a nonlinear equation of $M = 2N_\lambda + 1$ variables (where N_λ is the number of sources driving the fields). In order to make this problem easier to solve, it is convenient to rewrite the sinusoidal terms as a sum of $\sin$ and $\cos$, i.e., $A_j \sin(\omega_j t_i + \varphi_j) = A_j \cos(\omega_j t_i) + B_j \sin(\omega_j t_i)$. With this modification, we can rewrite our previous field expressions as

$$\vec{E} = A_1\vec{I_1} + A_2\vec{I_2} + \cdots + B_1\vec{Q}_1 + B_2\vec{Q}_2 + \cdots + \vec{c} \tag{C.6}$$

where $\vec{I_j} = [\cos\omega_j t_1, \cos\omega_j t_2, \cdots]^T$ is the “in phase" vector and $\vec{Q} = [\sin\omega_j t_1, \sin\omega_j t_2, \cdots]^T$ is the “quadrature” vector. We can rewrite this expression in terms of matrices as,

$$\vec{E} = I\vec{A} + Q\vec{B} + \vec{c} \tag{C.7}$$

where $\vec{A} = [A_1, A_2, \cdots]^T$, $\vec{B} = [B_1, B_2, \cdots]^T$, and I and Q are rectangular matrices with size $M \times N_\lambda$ whose columns are the $\vec{I}_j$'s and $\vec{Q}_j$'s, respectively. To further simplify this problem, let us combine the matrices I and Q into a single matrix $D = [IQ\vec{1}]$ and also combine the vectors $\vec{A}$, $\vec{B}$, $\vec{c}$ into a single vector $F = [\vec{A}^T, \vec{B}^T, \vec{c}^T]^T$. Doing so yields the simple matrix equation:

$$\vec{E} = D\vec{F} \tag{C.8}$$

In order to solve for the amplitude and phase of the field, all we need to do is solve for the vector $\vec{F}$,

$$\vec{F} = D^{-1}\vec{E} \tag{C.9}$$

For even large numbers of wavelengths ($\sim$10), this system of equations is small and quick to solve, making it an effective way to perform multi-wavelength simulations. Furthermore, notice that D is identical for every single point in the simulation domain, while $\vec{E}$ changes for each point. If we factorize D using LU decomposition, then the process of solving for F for each field value requires that we only solve a simple forward and back substitution, which is computationally inexpensive.

Overall, the computational complexity of this method is low. Given N_λ wavelengths, the simulation time will grow as $O(2N_\lambda + C)$ where C will typically be at least an order of magnitude larger than most N_λ of interest. This means that the simulation time typically will not increase significantly as more wavelengths are added to the simulation. Conversely, the storage requirements will grow as $O(2N_\lambda + 1)$ and the time it takes to compute the final frequency-domain fields will grow as $O(N_\lambda^2)$. For large numbers of wavelengths, the memory requirements may become a bottleneck, however for $N_\lambda \leq 10$, this method should be quite effective.

Appendix D

Grating Coupler Derivations

D.1 The Grating Equation

Perhaps the most important tool in designing a basic grating coupler is the grating equation. In its most common 1D form,

$$k_{\text{in}}(\Lambda) - \frac{2\pi m}{\Lambda} = k_c \sin\theta \tag{D.1}$$

the grating equation gives the relationship between the periodicity Λ of the grating coupler and its scattering angle θ. With this equation, we are able to design simpler uniform grating couplers which serve as an excellent starting points for optimizations.

The origin of this relationship is easily explained using the simple geometric arguments presented in Figure D.1. As a wave guided by the grating encounters an etched groove, light is scattered, giving rise to a plane wave which propagates at an angle away from the grating coupler. The phase accumulated by a guided wave over one period of the grating is equal to

$$\varphi_g = k_{\text{in}}\Lambda \tag{D.2}$$

where k_{in} is the effective wavenumber of the guided wave and Λ is the period of the grating. In the time it takes for this guided wave to propagate from one groove to the next, the wave scattered from the previous groove picks up a phase

$$\varphi_c = k_c\Lambda \sin\theta \tag{D.3}$$

where θ is the angle of the scatter waved with respect to the surface normal of the grating coupler. Notice that all of these scattered waves will constructively interfere if they are in phase. This is equivalent to saying the phase accumulated by a scattered wave is within a factor of 2π of the phase accumulated by the guided wave in one period of the grating, i.e.,

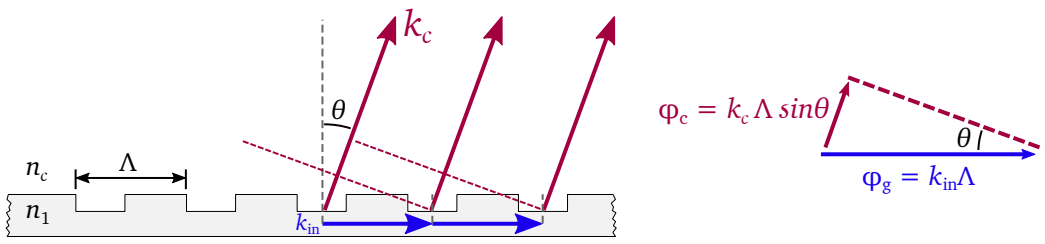

Figure D.1: Diagram of phase matching of scattered waves in 1D grating couplers. Each groove of the grating gives rise to a scattered plane wave. If the periodicity of the grooves is just right, the phase of each scattered wave along the grating will be in phase with the previous scattered wave.

$$\varphi_c + 2\pi m = \varphi_g \tag{D.4}$$

where m is an integer. Geometrically, we see that $\varphi_c = k_c \Lambda \sin\theta$. It follows that

$$k_c \Lambda \sin\theta + 2\pi m = k_{\text{in}} \Lambda \tag{D.5}$$

Dividing by Λ and rearranging terms yields

$$k_{\text{in}} - \frac{2\pi m}{\Lambda} = k_c \sin\theta \tag{D.6}$$

which is the familiar 1D grating equation (or phase matching condition). Notice that the integer m in this equation gives the diffraction order. If the grating period is sufficiently long, then multiple diffraction angles can satisfy the phase matching condition (corresponding to different values for m). In general this is not desirable.

D.2 Focusing Grating Couplers

In reality, grating couplers are never truly one dimensional. Instead, a linear grating coupler can be defined using a wide ($> 10\,\mu\text{m}$) slab which behaves like a 1D grating. In order to couple light into this wide slab, very long (greater than $100\,\mu\text{m}$) tapers are typically required to ensure that light is only coupled into the fundamental mode of the slab. If we make these tapers too short, the wavefront of the guided field will begin to curve, which will result in a scattered field with a spread of undesired wave vectors (i.e., the generated scattered field will not have the desired flat wavefront).

In order to get around this issue and reduce the overall footprint of the grating coupler, we can use shorter tapers and curve the etched grating lines such that the scattered field has only the single desired wave vector. Choosing the correct shape of the grating coupler slots, however, is no longer as simple as applying the phase matching process we introduced in the previous section.

Instead, we need a more general wave vector matching condition that applies to two dimensional structures, which we will derive now. Our starting point for this derivation is the scalar wave equation,

$$\nabla^2 E(\mathbf{r}) + k_0^2 n(\mathbf{r})^2 E(\mathbf{r}) = 0 \tag{D.7}$$

where $E(\mathbf{r})$ is the scalar electric field, $n(\mathbf{r})$ is the spatially-dependent refractive index, and $k_0 = 2\pi/\lambda$ is the wavenumber in vacuum. Note that this is only technically valid if the refractive index varies only weakly. While the derivation relies on this assumption, the end result works reasonably well for mediums with more significant refractive index variations.

For the sake of simplicity, let us assume that the square of the refractive index consists of a constant value which is modulated by a small amount, i.e.,

$$n^2(\mathbf{r}) = n_g^2 + \Delta n^2(\mathbf{r}) \quad . \tag{D.8}$$

As a result of this small modulation in refractive index, the electric field will also be perturbed with respect to the fields in a homogeneous medium with refractive index n_g. To account for this, let us assume that the electric field takes the form

$$E(\mathbf{r}) = E_g(\mathbf{r}) + \Delta E(\mathbf{r}) \tag{D.9}$$

where $E_g(\mathbf{r})$ is the electric field of a homogeneous medium and $\Delta E(\mathbf{r})$ is the perturbation to that electric field due to the modulation in refractive index. Substituting these expressions for $n(\mathbf{r})$ and $E(\mathbf{r})$ into Equation (D.7), yields

$$\nabla^2 \left[E_g(\mathbf{r}) + \Delta E(\mathbf{r})\right] + k_0^2 \left[n_g^2 + \Delta n^2(\mathbf{r})\right] \left[E_g(\mathbf{r}) + \Delta E(\mathbf{r})\right] = 0 \quad . \tag{D.10}$$

If the modulation to the refractive index is sufficiently small, then any products containing multiple Δ terms can be neglected. In this case, we have

$$\nabla^2 \Delta E(\mathbf{r}) + k_0^2 \Delta n^2(\mathbf{r}) E_g(\mathbf{r}) + k_0^2 n_g^2 \Delta E(\mathbf{r}) + \nabla^2 E_g(\mathbf{r}) + k_0^2 n_g^2 E_g^2(\mathbf{r}) = 0 \quad . \tag{D.11}$$

Notice that the right-most terms are simply the homogeneous wave equation. We know that these two terms add to zero. Canceling these terms and moving things around, we arrive at our perturbed scalar wave equation:

$$\nabla^2 \Delta E(\mathbf{r}) + k_0^2 n_g^2 \Delta E(\mathbf{r}) = -k_0^2 \Delta n^2(\mathbf{r}) E_g(\mathbf{r}) \tag{D.12}$$

This tells us that the perturbation to the electric field is the solution to the wave equation in a homogeneous medium which is excited by a source given by the product of the refractive index modulation and homogeneous electric field!

Our next step is to “solve" this equation. To do so, let us first recall that the solution to the scalar wave equation in a homogeneous medium is a plane wave. In other words, the unperturbed field $E_g(\mathbf{r})$ takes the form

$$E_g(\mathbf{r}) = a_0 e^{i\mathbf{k}_{\text{in}}\cdot\mathbf{r}} \tag{D.13}$$

where a_0 is the amplitude of the wave and $\mathbf{k}_{\text{in}}$ is its wave vector. Because the homogeneous field is a plane wave, it is convenient to also represent the perturbed field and modulation to the refractive index using a sum of plane waves. In particular, let us write the perturbed electric field and modulation to the refractive index as Fourier transforms, i.e.,

$$\Delta E(\mathbf{r}) = \int d\mathbf{k} B(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{D.14}$$

and

$$\Delta n^2(\mathbf{r}) = \int d\mathbf{k} N(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{D.15}$$

where the limits of integration are assumed to be the same in both cases. Although not yet apparent, we stand to learn a lot by solving for the coefficients $B(\mathbf{k})$ and $N(\mathbf{k})$. We accomplish this by substituting these definitions into Equation (D.12), which yields

$$\nabla^2 \int d\mathbf{k} B(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} + k_0^2 n_g^2 \int d\mathbf{k} B(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} = -k_0^2 a_0 e^{i\mathbf{k}_{\text{in}}\cdot\mathbf{r}} \int d\mathbf{k} N(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{D.16}$$

Since the integrals are over $\mathbf{k}$, we are free to move spatially-dependent quantities inside the integrals. Doing so yields a slightly simpler expression:

$$\int d\mathbf{k} \left[\nabla^2 B(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} + k_0^2 n_g^2 B(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}\right] = -k_0^2 a_0 \int d\mathbf{k} N(\mathbf{k} - \mathbf{k}_{\text{in}}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{D.17}$$

where in the integral on the right hand side we have also made the substitution $\mathbf{k} \to \mathbf{k} - \mathbf{k}_{\text{in}}$. Notice that up until this point, we have placed no constraints on the limits of the integrals in Equation (D.14) and Equation (D.15) except that they must be the same. Equation (D.17) must therefore hold regardless of the limits of integration. According to the fundamental theorem of calculus, it follows that the integrands must be equal as well, i.e.,

$$\nabla^2 B(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} + k_0^2 n_g^2 B(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} = -k_0^2 a_0 N(\mathbf{k} - \mathbf{k}_{\text{in}}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{D.18}$$

Notice that $\nabla^2 e^{i\mathbf{k}\cdot\mathbf{r}} = -|\mathbf{k}|^2 e^{i\mathbf{k}\cdot\mathbf{r}}$. Using this observation, we can solve Equation (D.18) for $B(\mathbf{k})$:

$$B(\mathbf{k}) = \frac{-k_0^2 a_0 N(\mathbf{k} - \mathbf{k}_{\text{in}})}{k_0^2 n_g^2 - |\mathbf{k}|^2} \tag{D.19}$$

This expression effectively gives us the amplitude of plane waves which sum to form the perturbed electric field. Notice that if the modulation to the refractive index contains a spatial frequency $\mathbf{k}_{\Delta n}$ (i.e., $N(\mathbf{k}_{\Delta n}) \neq 0$), then the electric field will contain a wave vector $\mathbf{k} = \mathbf{k}_{\Delta n} + \mathbf{k}_{\text{in}}$ with a corresponding non-zero amplitude.

It is important to note that this applies only to fields which exist within the periodically modulated medium. In the case of grating couplers, which we can model as a fixed effective index with a finite-thickness periodic modulation, the field outside of the grating coupler is not constrained to this relationship. Nonetheless, any fields that exist outside of the grating coupler must satisfy boundary conditions at the grating coupler surface. This is equivalent to saying that the *in-plane component of the wave vector of the scattered fields* (defined as the plane containing $\mathbf{k}_{\Delta n}$ and $\mathbf{k}_{\text{in}}$) *must match the wave vector of the perturbed fields* that we just found. In other words, the in-plane component of the scattered wave vector must satisfy the condition

$$\mathbf{k}_{\parallel} = \mathbf{k}_{\text{in}} + \mathbf{k}_{\Delta n} \tag{D.20}$$

If the grating has a periodicity of Λ and the direction of modulation is $\hat{k}_g$, then this expression can be rewritten in a slightly more enlightening way as

$$\mathbf{k}_{\parallel} = \mathbf{k}_{\text{in}} + \frac{2\pi m}{\Lambda}\hat{k}_g \tag{D.21}$$

where m is an integer. Note that the magnitude of the whole scattered wave vector outside of the grating coupler must also have a magnitude according to the medium its in, i.e.

$$|\mathbf{k}| = \sqrt{|\mathbf{k}_{\parallel}|^2 + |\mathbf{k}_{\perp}|^2} = \frac{2\pi n_c}{\lambda} \tag{D.22}$$

where n_c is the refractive index of the cladding outside of the grating coupler and λ is the vacuum wavelength.

Equations (D.20) and (D.21) are the most general forms of the grating coupler equation. When working with either 1D linear grating couplers or 2D grating couplers, these results can be reduced to more useful expressions.

In the case of the non-uniform 2D gratings, our first step in rewriting this equation in a more enlightening way is to note that as long as the grating does not change too rapidly, then we can treat it as being locally periodic. In this case, the wave vectors in Equation (D.21) are spatially dependent, i.e.,

$$\mathbf{k}_{\parallel}(x, y) = \mathbf{k}_{\text{in}}(x, y) + \frac{2\pi m}{\Lambda}\hat{k}_g(x, y) \tag{D.23}$$

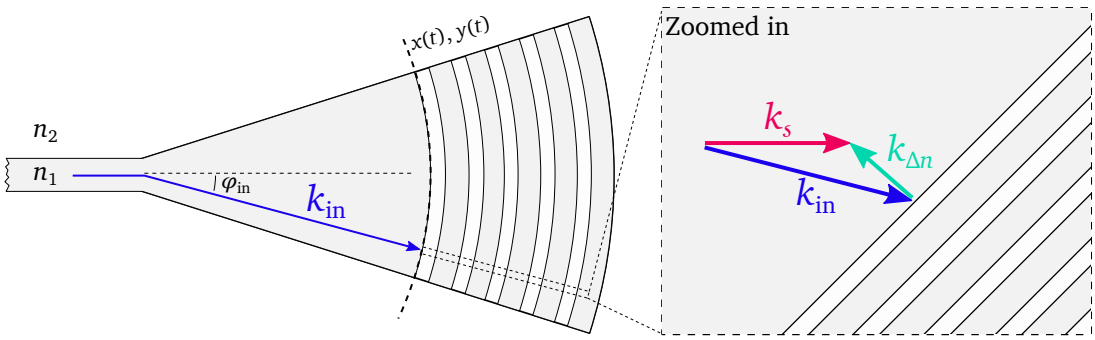

Figure D.2: Diagram of the wave vectors in a focusing grating coupler. Light entering a focusing grating diffracts as it propagates towards the grating line. This leads to different wave vectors $\mathbf{k}_{\text{in}}$ at different points along the grating lines. The grating lines are curved such that the in-plane component of the wave vector of the scattered light always points along the $\hat{x}$ direction.

where we assume the grating is in the x, y plane. If the input wave vectors are known and the desired output wave vectors are specified, then the normal direction of the grating lines can be found by solving for $\hat{k}_g(x, y)$. Doing so will yield a shape for our non-uniform grating coupler.

A special type of non-uniform gratings which is of particular interest in silicon photonics is the *focusing grating*. The focusing grating, depicted in Figure D.2, consists of a wedge with relatively steep angles which has a curved grating etched into it. Light entering this wedge from the input waveguide will diffract, causing the wavefronts to curve. By curving the grating lines, we can compensate for the curved wavefront in order to scatter into a beam with a single in-plane wavevector component (were we to use straight grating lines, the scattered light would have a curved wavefront which is not well matched to an optical fiber).

For the sake of simplicity, let us assume that the wavefront of the light in the wedge slab has a circular wavefront. In this case, the incident wavevector has the simple spatial dependence given by

$$\mathbf{k}_{\text{in}}(\varphi_{\text{in}}) = k_{\text{in}} \cos\varphi_{\text{in}} \hat{x} + k_{\text{in}} \sin\varphi_{\text{in}} \hat{y} \tag{D.24}$$

where φ_{in} is labeled in Figure D.2. Furthermore, the goal of focusing gratings is typically to couple light into an optical fiber which is tiled at an angle with respect to the z axis and lies within the xz plane. In this case, the in-plane part of the scattered wave vector has a single component in the x direction. In particular, the in-plane wave vector has the form $\mathbf{k}_{\parallel} = \hat{x} k_s \sin\theta_{\text{out}}$ where θ_{out} is the angle of the fiber with respect to the z axis. Based on these assumptions, we can rewrite Equation (D.23) in terms of its $\hat{x}$ and $\hat{y}$ components as

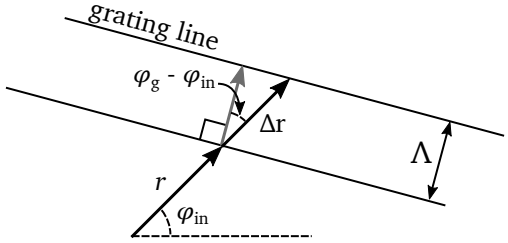

Figure D.3: Simple geometric arguments used for simplifying focusing grating equation.

$$k_s \sin\theta = k_{\text{in}} \cos\varphi_{\text{in}} + \frac{2\pi m}{\Lambda} k_{gx} \tag{D.25}$$

$$0 = k_{\text{in}} \sin\varphi_{\text{in}} + \frac{2\pi m}{\Lambda} k_{gy} \tag{D.26}$$

where k_{gx} and k_{gy} are defined such that $\hat{k}_g = k_{gx}\hat{x} + k_{gy}\hat{y}$. Determining the shape of our focusing grating is equivalent to finding its period and the trajectories of the individual grating lines. In order to find the period in terms of convenient quantities, we can square and add Equations (D.25) and (D.26) which yields

$$\Lambda = \frac{m\lambda}{\sqrt{n_s^2 + n_{\text{in}}^2 \sin^2\varphi_{\text{in}} - 2n_s n_{\text{in}} \cos\varphi_{\text{in}} \sin\theta_{\text{out}}}} \tag{D.27}$$

where λ is the vacuum wavelength, n_s is the refractive index of the cladding, and n_{in} is the effective index of the incident guided wave within the grating. Notice that if we set $\varphi_{\text{in}} = 0$, this expression reduces to the period obtained using the 1D grating equation.

With the period known, we can now solve for the trajectories of the grating lines. To do so, we first solve for the normal direction of the grating lines, $\hat{k}_g$, which is

$$k_{gx} = \frac{n_{\text{in}} \cos\varphi_{\text{in}} - n_s \sin\theta_{\text{out}}}{\sqrt{n_s^2 + n_{\text{in}}^2 \sin^2\varphi_{\text{in}} - 2n_s n_{\text{in}} \cos\varphi_{\text{in}} \sin\theta_{\text{out}}}} \tag{D.28}$$

$$k_{gy} = -\frac{n_{\text{in}} \sin\varphi_{\text{in}}}{\sqrt{n_s^2 + n_{\text{in}}^2 \sin^2\varphi_{\text{in}} - 2n_s n_{\text{in}} \cos\varphi_{\text{in}} \sin\theta_{\text{out}}}} \tag{D.29}$$

These two equations define the *normal* direction of the non-uniform grating lines, which is enough to define the shape of the focusing grating. In particular, we can define the tangent vector $\hat{T}_g = -k_{gy}\hat{x} + k_{gx}\hat{y}$ of the grating lines. In order to construct the grating lines, we can use a simple iterative process based on these spatially-dependent tangent vectors:

$$\begin{aligned} x_g(t+1) &= x_g(t) - k_{gy}\Delta t \\ y_g(t+1) &= y_g(t) + k_{gx}\Delta t \end{aligned} \tag{D.30}$$

where $x_g(t)$ and $y_g(t)$ are parametric function which describe the positions of points which lie along the grating lines. Starting with some initial position (x_0, y_0), we can use this expression to draw grating lines.

A more common expression for the shapes of the grating lines, however, can be derived using a few simple geometric arguments. Consider Figure D.3 which shows an incident “ray" on the grating which has traveled a distance r. Geometrically, we see that the period and the distance this ray travels between two period of the grating are related by the equations

$$\begin{aligned} \Lambda &= \Delta r \cos(\varphi_g - \varphi_{\text{in}}) \\ \Rightarrow \Lambda &= \Delta r \left(\cos\varphi_g \cos\varphi_{\text{in}} + \sin\varphi_g \sin\varphi_{\text{in}}\right) \end{aligned} \tag{D.31}$$

where we have defined a new angle φ_g which describes the normal direction of the grating lines, i.e., $\hat{k}_g = \cos\varphi_g\hat{x} + \sin\varphi_g\hat{y}$. Substituting in Equations (D.27), (D.28), and (D.29), we find that distance Δr is given by

$$\Delta r = \frac{m\lambda}{n_{\text{in}} - n_{\text{s}} \sin\theta_{\text{out}} \cos\varphi_{\text{in}}} \tag{D.32}$$

Notice because Λ and Δr are proportional, this expression holds for any multiple of the grating period. Based on this, we can write an arbitrary radial position of the grating lines $r = q\Delta r$ as

$$r = \frac{qm\lambda}{n_{\text{in}} - n_{\text{s}} \sin\theta_{\text{out}} \cos\varphi_{\text{in}}} \tag{D.33}$$

where q is a real number. This expression defines ellipses for the focusing grating lines and is consistent with other expressions presented in the literature [93]. In general, however, Equation (D.30) may be easier to use.

Bibliography

[1] R. T. Haftka and R. V. Grandhi, "Structural shape optimization–A survey," *Computer Methods in Applied Mechanics and Engineering*, vol. 57, pp. 91–106, Aug. 1986.

[2] M. P. Bendsøe and N. Kikuchi, "Generating optimal topologies in structural design using a homogenization method," *Computer Methods in Applied Mechanics and Engineering*, vol. 71, pp. 197–224, Nov. 1988.

[3] O. Sigmund, "On the Design of Compliant Mechanisms Using Topology Optimization," *Mechanics of Structures and Machines*, vol. 25, pp. 493–524, Jan. 1997.

[4] G. Allaire, F. Jouve, and A.-M. Toader, "Structural optimization using sensitivity analysis and a level-set method," *Journal of Computational Physics*, vol. 194, pp. 363–393, Feb. 2004.

[5] H.-B. Lee and T. Itoh, "A systematic optimum design of waveguide-to-microstrip transition," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, pp. 803–809, May 1997.

[6] Y.-S. Chung, C. Cheon, I.-H. Park, and S.-Y. Hahn, "Optimal shape design of microwave device using FDTD and design sensitivity analysis," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, pp. 2289–2296, Dec. 2000.

[7] P. I. Borel, A. Harpøth, L. H. Frandsen, M. Kristensen, P. Shi, J. S. Jensen, and O. Sigmund, "Topology optimization and fabrication of photonic crystal structures," *Opt. Express*, vol. 12, pp. 1996–2001, May 2004.

[8] J. Jensen and O. Sigmund, "Topology optimization for nano-photonics," *Laser & Photon. Rev.*, vol. 5, pp. 308–321, Mar. 2011.

[9] J. Lu and J. Vučković, "Nanophotonic computational design," *Opt. Express*, vol. 21, pp. 13351–13367, June 2013.

[10] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, "Adjoint shape optimization applied to electromagnetic design," *Optics Express*, vol. 21, pp. 21693–21701, Sept. 2013.

[11] L. H. Frandsen, Y. Elesin, L. F. Frellsen, M. Mitrovic, Y. Ding, O. Sigmund, and K. Yvind, "Topology optimized mode conversion in a photonic crystal waveguide fabricated in silicon-on-insulator material," *Opt. Express*, vol. 22, pp. 8525–8532, Apr. 2014.

[12] A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," *Nature Photonics*, vol. 9, pp. 374–377, June 2015.

[13] B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with 2.4 $\times$ 2.4 $\mu\mathrm{m}^2$ footprint," *Nature Photonics*, vol. 9, p. 378, 2015.

[14] O. Sigmund, J. S. Jensen, and L. H. Frandsen, "On nanostructured silicon success," *Nat. Photon*, vol. 10, pp. 142–143, Mar. 2016.

[15] L. F. Frellsen, Y. Ding, O. Sigmund, and L. H. Frandsen, "Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides," *Optics Express*, vol. 24, pp. 16866–16873, July 2016.

[16] A. Y. Piggott, J. Petykiewicz, L. Su, and J. Vučković, "Fabrication-constrained nanophotonic inverse design," *Scientific Reports*, vol. 7, p. 1786, May 2017.

[17] Z. Lin, M. Lončar, and A. W. Rodriguez, "Topology optimization of multi-track ring resonators and 2d microcavities for nonlinear frequency conversion," *Opt. Lett.*, vol. 42, pp. 2818–2821, Jul 2017.

[18] J. Wang, Y. Shi, T. Hughes, Z. Zhao, and S. Fan, "Adjoint-based optimization of active nanophotonic devices," *Opt. Express*, vol. 26, pp. 3236–3248, Feb 2018.

[19] Y. Elesin, B. S. Lazarov, J. S. Jensen, and O. Sigmund, "Design of robust and efficient photonic switches using topology optimization," *Photonics and Nanostructures - Fundamentals and Applications*, vol. 10, pp. 153–165, Jan. 2012.

[20] M. H. Bakr and N. K. Nikolova, "An adjoint variable method for frequency domain TLM problems with conducting boundaries," *IEEE Microwave and Wireless Components Letters*, vol. 13, pp. 408–410, Sept. 2003.

[21] G. Veronis, R. W. Dutton, and S. Fan, "Method for sensitivity analysis of photonic crystal devices," *Opt. Lett.*, vol. 29, pp. 2288–2290, Oct 2004.

[22] E. A. Soliman, M. H. Bakr, and N. K. Nikolova, "An adjoint variable method for sensitivity calculations of multiport devices," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, pp. 589–599, Feb. 2004.

[23] L. Pang, Y. Liu, and D. Abrams, "Inverse lithography technology (ILT): What is the impact to the photomask industry?," in *International Society for Optics and Photonics*, vol. 6283, may 2006.

[24] T. Nomura, K. Sato, K. Taguchi, T. Kashiwa, and S. Nishiwaki, "Structural topology optimization for the design of broadband dielectric resonator antennas using the finite difference time domain technique," *International Journal for Numerical Methods in Engineering*, vol. 71, no. 11, pp. 1261–1296, 2007.

[25] T. G. Jurgens, A. Taflove, K. Umashankar, and T. G. Moore, "Finite-difference time-domain modeling of curved surfaces [EM scattering]," *IEEE Transactions on Antennas and Propagation*, vol. 40, pp. 357–366, Apr. 1992.

[26] S. Dey and R. Mittra, "A modified locally conformal finite-difference time-domain algorithm for modeling three-dimensional perfectly conducting objects," *Microwave and Optical Technology Letters*, vol. 17, pp. 349–352, Apr. 1998.

[27] S. Dey and R. Mittra, "A conformal finite-difference time-domain technique for modeling cylindrical dielectric resonators," *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, pp. 1737–1739, Sept. 1999.

[28] N. Kaneda, B. Houshmand, and T. Itoh, "FDTD analysis of dielectric resonators with curved surfaces," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, pp. 1645–1649, Sept. 1997.

[29] A. Farjadpour, D. Roundy, A. Rodriguez, M. Ibanescu, P. Bermel, J. D. Joannopoulos, S. G. Johnson, and G. W. Burr, "Improving accuracy by subpixel smoothing in the finite-difference time domain," *Optics Letters*, vol. 31, pp. 2972–2974, Oct. 2006.

[30] A. F. Oskooi, C. Kottke, and S. G. Johnson, "Accurate finite-difference time-domain simulation of anisotropic media by subpixel smoothing," *Optics Letters*, vol. 34, pp. 2778–2780, Sept. 2009.

[31] F. P. Preparata and M. Shamos, *Computational Geometry: An Introduction*. Monographs in Computer Science, Springer-Verlag, 1985.

[32] O. Sigmund and K. Maute, "Topology optimization approaches," *Structural and Multidisciplinary Optimization*, vol. 48, pp. 1031–1055, Dec 2013.

[33] W. Zhang, J. Yuan, J. Zhang, and X. Guo, "A new topology optimization approach based on moving morphable components (mmc) and the ersatz material model," *Structural and Multidisciplinary Optimization*, vol. 53, pp. 1243–1260, Jun 2016.

[34] A. Michaels and E. Yablonovitch, "Leveraging continuous material averaging for inverse electromagnetic design," *Optics Express*, vol. 26, pp. 31717–31737, Nov. 2018.

[35] A. Michaels, "Emopt: Electromagnetic optimization toolbox." https://github.com/anstmichaels/emopt, 2018.

[36] and and C. A. Balanis, "Analysis of singly and doubly periodic absorbers by frequency-domain finite-difference method," *IEEE Transactions on Antennas and Propagation*, vol. 44, pp. 798–805, June 1996.

[37] N. J. Champagne, J. G. Berryman, and H. M. Buettner, "FDFD: A 3d Finite-Difference Frequency-Domain Code for Electromagnetic Induction Tomography," *Journal of Computational Physics*, vol. 170, pp. 830–848, July 2001.

[38] W. Shin and S. Fan, "Choice of the perfectly matched layer boundary condition for frequency-domain Maxwell's equations solvers," *Journal of Computational Physics*, vol. 231, pp. 3406–3431, Apr. 2012.

[39] W. Shin and S. Fan, "Accelerated solution of the frequency-domain Maxwell's equations by engineering the eigenvalue distribution of the operator," *Optics Express*, vol. 21, pp. 22578–22595, Sept. 2013.

[40] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes, R. T. Mills, T. Munson, K. Rupp, P. Sanan, B. F. Smith, S. Zampini, H. Zhang, and H. Zhang, "PETSc Web page." http://www.mcs.anl.gov/petsc, 2018.

[41] W. Briggs, V. Henson, and S. McCormick, *A Multigrid Tutorial, 2nd Edition*. 01 2000.

[42] P. R. Amestoy, I. S. Duff, J. Koster, and J.-Y. L'Excellent, "A fully asynchronous multifrontal solver using distributed dynamic scheduling," *SIAM Journal on Matrix Analysis and Applications*, vol. 23, no. 1, pp. 15–41, 2001.

[43] P. R. Amestoy, A. Guermouche, J.-Y. L'Excellent, and S. Pralet, "Hybrid scheduling for the parallel solution of linear systems," *Parallel Computing*, vol. 32, no. 2, pp. 136–156, 2006.

[44] K. Yee, "Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media," *IEEE Transactions on Antennas and Propagation*, vol. 14, pp. 302–307, May 1966.

[45] A. Taflove and S. C. Hagness, *Computational electrodynamics: the finite-difference time-domain method; 3rd ed.* Artech House antennas and propagation library, Boston, MA: Artech House, 2005.

[46] V. Hernandez, J. E. Roman, and V. Vidal, "SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems," *ACM Trans. Math. Software*, vol. 31, no. 3, pp. 351–362, 2005.

[47] B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with 2.4 $\times$ 2.4 $\mu\mathrm{m}^2$ footprint," *Nature Photonics*, vol. 9, p. 378, May 2015.

[48] Y. Elesin, B. S. Lazarov, J. S. Jensen, and O. Sigmund, "Time domain topology optimization of 3d nanophotonic devices," *Photonics and Nanostructures - Fundamentals and Applications*, vol. 12, pp. 23–33, Feb. 2014.

[49] J. Lu and J. Vučković, "Objective-first design of high-efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes," *Optics Express*, vol. 20, pp. 7221–7236, Mar. 2012.

[50] L. Su, R. Trivedi, N. V. Sapra, A. Y. Piggott, D. Vercruysse, and J. Vučković, "Fully-automated optimization of grating couplers," *Optics Express*, vol. 26, pp. 4023–4034, Feb. 2018.

[51] R. M. Jenkins, R. W. J. Devereux, and J. M. Heaton, "Waveguide beam splitters and recombiners based on multimode propagation phenomena," *Optics Letters*, vol. 17, pp. 991–993, July 1992.

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

[53] S. H. Tao, Q. Fang, J. F. Song, M. B. Yu, G. Q. Lo, and D. L. Kwong, "Cascade wide-angle Y-junction 1 $\times$ 16 optical power splitter based on silicon wire waveguides on silicon-on-insulator," *Optics Express*, vol. 16, pp. 21456–21461, Dec. 2008.

[54] M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons, "Deep Neural Network Inverse Design of Integrated Nanophotonic Devices," *arXiv:1809.03555 [cond-mat, physics:physics]*, Sept. 2018. arXiv: 1809.03555.

[55] Q. Wang, J. Lu, and S. He, "Optimal design method of a low-loss broadband Y branch with a multimode waveguide section," *Applied Optics*, vol. 41, pp. 7644–7649, Dec. 2002.

[56] L. H. Frandsen, P. I. Borel, Y. X. Zhuang, A. Harpøth, M. Thorhauge, M. Kristensen, W. Bogaerts, P. Dumon, R. Baets, V. Wiaux, J. Wouters, and S. Beckx, "Ultralow-loss 3-dB photonic crystal waveguide splitter," *Optics Letters*, vol. 29, pp. 1623–1625, July 2004.

[57] Y. Wang, S. Gao, K. Wang, and E. Skafidas, "Ultra-broadband and low-loss $3\,\mathrm{dB}$ optical power splitter based on adiabatic tapered silicon waveguides," *Optics Letters*, vol. 41, pp. 2053–2056, May 2016.

[58] J. K. Doylend, M. J. R. Heck, J. T. Bovington, J. D. Peters, L. A. Coldren, and J. E. Bowers, "Two-dimensional free-space beam steering with an optical phased array on silicon-on-insulator," *Optics Express*, vol. 19, pp. 21595–21604, Oct. 2011.

[59] J. Sun, E. Timurdogan, A. Yaacobi, E. S. Hosseini, and M. R. Watts, "Large-scale nanophotonic phased array," *Nature*, vol. 493, pp. 195–199, Jan. 2013.

[60] C. V. Poulton, P. Russo, E. Timurdogan, M. Whitson, M. J. Byrd, E. Hosseini, B. Moss, Z. Su, D. Vermeulen, and M. R. Watts, "High-Performance Integrated Optical Phased Arrays for Chip-Scale Beam Steering and LiDAR," in *Conference on Lasers and Electro-Optics (2018), paper ATu3R.2*, p. ATu3R.2, Optical Society of America, May 2018.

[61] S. Han, T. J. Seok, N. Quack, B.-W. Yoo, and M. C. Wu, "Large-scale silicon photonic switches with movable directional couplers," *Optica*, vol. 2, pp. 370–375, Apr. 2015.

[62] T. J. Seok, N. Quack, S. Han, R. S. Muller, and M. C. Wu, "Large-scale broadband digital silicon photonic switches with vertical adiabatic couplers," *Optica*, vol. 3, pp. 64–70, Jan. 2016.

[63] A. Joshi, C. Batten, Y. Kwon, S. Beamer, I. Shamim, K. Asanovic, and V. Stojanovic, "Silicon-photonic clos networks for global on-chip communication," in *2009 3rd ACM/IEEE International Symposium on Networks-on-Chip*, pp. 124–133, May 2009.

[64] D. Nikolova, S. Rumley, D. Calhoun, Q. Li, R. Hendry, P. Samadi, and K. Bergman, "Scaling silicon photonic switch fabrics for data center interconnection networks," *Optics Express*, vol. 23, pp. 1159–1175, Jan. 2015.

[65] J. L. O'Brien, "Optical Quantum Computing," *Science*, vol. 318, pp. 1567–1570, Dec. 2007.

[66] X. Qiang, X. Zhou, J. Wang, C. M. Wilkes, T. Loke, S. O'Gara, L. Kling, G. D. Marshall, R. Santagati, T. C. Ralph, J. B. Wang, J. L. O'Brien, M. G. Thompson, and J. C. F. Matthews, "Large-scale silicon quantum photonics implementing arbitrary two-qubit processing," *Nature Photonics*, vol. 12, p. 534, Sept. 2018.

[67] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, "Deep learning with coherent nanophotonic circuits," *Nature Photonics*, vol. 11, pp. 441–446, July 2017.

[68] H. Wei, J. Yu, X. Zhang, and Z. Liu, "Compact 3-dB tapered multimode interference coupler in silicon-on-insulator," *Optics Letters*, vol. 26, pp. 878–880, June 2001.

[69] K. Solehmainen, M. Kapulainen, M. Harjanne, and T. Aalto, "Adiabatic and Multimode Interference Couplers on Silicon-on-Insulator," *IEEE Photonics Technology Letters*, vol. 18, pp. 2287–2289, Nov. 2006.

[70] F. Wang, Y. Jianyi, C. Limei, J. Xiaoqing, and W. Minghua, "Optical switch based on multimode interference coupler," *IEEE Photonics Technology Letters*, vol. 18, pp. 421–423, Jan. 2006.

[71] M. Yong, P. Seoijin, W. Liwei, and H. Seng Tiong, "Ultracompact multimode interference 3-dB coupler with strong lateral confinement by deep dry etching," *IEEE Photonics Technology Letters*, vol. 12, pp. 492–494, May 2000.

[72] H. Yun, W. Shi, Y. Wang, L. Chrostowski, and N. A. F. Jaeger, "2x2 adiabatic 3-dB coupler on silicon-on-insulator rib waveguides," in *Photonics North 2013*, vol. 8915, p. 89150V, International Society for Optics and Photonics, Oct. 2013.

[73] J. Xing, Z. Li, Y. Yu, and J. Yu, "Design of polarization-independent adiabatic splitters fabricated on silicon-on-insulator substrates," *Optics Express*, vol. 21, pp. 26729–26734, Nov. 2013.

[74] J. Xing, K. Xiong, H. Xu, Z. Li, X. Xiao, J. Yu, and Y. Yu, "Silicon-on-insulator-based adiabatic splitter with simultaneous tapering of velocity and coupling," *Optics Letters*, vol. 38, pp. 2221–2223, July 2013.

[75] K. Hassan, C. Durantin, V. Hugues, B. Szelag, and A. Glière, "Robust silicon-on-insulator adiabatic splitter optimized by metamodeling," *Applied Optics*, vol. 56, pp. 2047–2052, Mar. 2017.

[76] H.-C. Chung and S.-Y. Tseng, "Robust silicon 3-dB coupler using inverse engineering based optimization," *Japanese Journal of Applied Physics*, vol. 57, p. 08PC01, June 2018.

[77] W. Bogaerts, P. Dumon, D. V. Thourhout, and R. Baets, "Low-loss, low-cross-talk crossings for silicon-on-insulator nanophotonic waveguides," *Optics Letters*, vol. 32, pp. 2801–2803, Oct. 2007.

[78] P. J. Bock, P. Cheben, J. H. Schmid, J. Lapointe, A. Delâge, D.-X. Xu, S. Janz, A. Densmore, and T. J. Hall, "Subwavelength grating crossings for silicon wire waveguides," *Optics Express*, vol. 18, pp. 16146–16155, July 2010.

[79] Y. Zhang, A. Hosseini, X. Xu, D. Kwong, and R. T. Chen, "Ultralow-loss silicon waveguide crossing using Bloch modes in index-engineered cascaded multimode-interference couplers," *Optics Letters*, vol. 38, pp. 3608–3611, Sept. 2013.

[80] Y. Ma, Y. Zhang, S. Yang, A. Novack, R. Ding, A. E.-J. Lim, G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, "Ultralow loss single layer submicron silicon waveguide crossing for SOI optical interconnect," *Optics Express*, vol. 21, pp. 29374–29382, Dec. 2013.

[81] Y. Zhang, M. Streshinsky, A. Novack, Y. Ma, S. Yang, A. E.-J. Lim, G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, "A compact and low-loss silicon waveguide crossing for O-band optical interconnect," in *Silicon Photonics IX*, vol. 8990, p. 899002, International Society for Optics and Photonics, Mar. 2014.

[82] Y. Liu, J. M. Shainline, X. Zeng, and M. A. Popović, "Ultra-low-loss CMOS-compatible waveguide crossing arrays based on multimode Bloch waves and imaginary coupling," *Optics Letters*, vol. 39, pp. 335–338, Jan. 2014.

[83] Z. Yu, A. Feng, X. Xi, and X. Sun, "Inverse-designed low-loss and wideband polarization-insensitive silicon waveguide crossing," *Optics Letters*, vol. 44, pp. 77–80, Jan. 2019.

[84] A. Mekis, S. Gloeckner, G. Masini, A. Narasimha, T. Pinguet, S. Sahni, and P. D. Dobbelaere, "A Grating-Coupler-Enabled CMOS Photonics Platform," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, pp. 597–608, May 2011.

[85] T. Tamir and S. T. Peng, "Analysis and design of grating couplers," *Applied physics*, vol. 14, pp. 235–254, Nov. 1977.

[86] D. Vermeulen, S. Selvaraja, P. Verheyen, G. Lepage, W. Bogaerts, P. Absil, D. V. Thourhout, and G. Roelkens, "High-efficiency fiber-to-chip grating couplers realized using an advanced CMOS-compatible Silicon-On-Insulator platform," *Optics Express*, vol. 18, pp. 18278–18283, Aug. 2010.

[87] A. Bozzola, L. Carroll, D. Gerace, I. Cristiani, and L. C. Andreani, "Optimising apodized grating couplers in a pure SOI platform to -0.5 dB coupling efficiency," *Optics Express*, vol. 23, June 2015.

[88] D. Vermeulen, Y. D. Koninck, Y. Li, E. Lambert, W. Bogaerts, R. Baets, and G. Roelkens, "Reflectionless grating couplers for Silicon-on-Insulator photonic integrated circuits," *Optics Express*, vol. 20, pp. 22278–22283, Sept. 2012.

[89] J. Notaros, F. Pavanello, M. T. Wade, C. Gentry, A. Atabaki, L. Alloatti, R. J. Ram, and M. Popović, "Ultra-Efficient CMOS Fiber-to-Chip Grating Couplers," in *Optical Fiber Communication Conference (2016), paper M2I.5*, p. M2I.5, Optical Society of America, Mar. 2016.

[90] M. Dai, L. Ma, Y. Xu, M. Lu, X. Liu, and Y. Chen, "Highly efficient and perfectly vertical chip-to-fiber dual-layer grating coupler," *Optics Express*, vol. 23, pp. 1691–1698, Jan. 2015.

[91] R. Waldhäusl, B. Schnabel, P. Dannberg, E.-B. Kley, A. Bräuer, and W. Karthe, "Efficient Coupling into Polymer Waveguides by Gratings," *Applied Optics*, vol. 36, pp. 9383–9390, Dec. 1997.

[92] C.-L. Chen, *Foundations for Guided-Wave Optics*, pp. 417–420. John Wiley & Sons, Inc., 2006.

[93] R. Waldhäusl, B. Schnabel, P. Dannberg, E.-B. Kley, A. Bräuer, and W. Karthe, "Efficient coupling into polymer waveguides by gratings," *Appl. Opt.*, vol. 36, pp. 9383–9390, Dec 1997.