Simulating Photonic Crystal Structures in Highly Anisotropic Materials and Optimizing with the Adjoint Method

A Thesis in Physics

by

John McLaughlin

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Bachelor in Science
With Specialized Honors in Physics

Contents

Abstract

Nanophotonics is a field focused on the control and manipulation of individual photons. This paper reviews the simulation and optimization of a nanophotonic device known as a line-defect cavity, which is responsible for confining photons for an extended period of time in order to enhance light-matter interactions. The cavity is made from the anisotropic material tungsten disulfide. The simulation and optimization of anisotropic materials are of particular interest because such materials have control over light polarization, stronger electromagnetic field confinement capabilities, and allow for directional flow of energy. The inverse design technique known as the adjoint method is used to design and optimize the cavity, employed through the computational electromagnetic field simulator Tidy3D. The results achieve a 234% increase in the cavity quality factor, and the methodology employed reflects promise in simulating highly anisotropic materials.

Chapter 1

Introduction

For centuries, humans have been pursuing a fundamental understanding of one component of the universe that weaves together the very fabric of our existence - light. Earliest recordings about the nature of light date as far back as the 5-6th century B.C. Ancient India was well known for it's philosophical schools of Samkhya and Vaisheshika, both of which proposed a great many theories explaining the nature of our existence. Samkhya theorized light to be one of five elements (referred to as *tanmatra*), from which gross elements of Earth, Water, Fire, Air, and Ether (or Space) are born. Vaisheshika gave an atomic based theory of light in contrast, suggesting that light "consists of fire-like particles moving at a high speed" [1]. Other societies such as ancient Egypt even performed experiments with light, proving that it reflects off of smooth surfaces and bends when passing through materials of different optical densities.

By the 17th century, the argument of whether light is a particle or a wave began. Isaac Newton argued that light was composed of a stream of particles, coining his idea as corpuscular theory (particles were referred to as corpuscles at the time). Other scientists like Dutch physicist Christian Huygens argued that light was a vibrating wave that traveled in an ether [2]. Newton was so greatly respected at the time, however, that it took decades for

any argument to seriously challenge his own. His theory did, of course, have holes in it such as refraction, where he needed to invoke some unknown force to explain why it is that light changes velocity when traveling through mediums like water.

After identifying several more holes in Newton's theory, Thomas Young developed the famous slit experiment in 1801. Young sought to prove that light behaved like a wave, similar to how ripples propagate on a body of water. The experiment involved shining a beam of sunlight through a small slit, which then shined onto two more slits and finally onto a screen. When he moved the two slits further apart and close to the screen, it formed two overlapping patches of light on the screen. When he reduced the distance between the slits, distinct bands of color formed on the screen, separated by dark regions in a serial order. He coined this phenomena as interference fringes, and deduced they could only be produced if light behaved like a wave [3].

In 1905, Einstein was able to explain the phenomenon known as the photoelectric effect, where electrons are emitted from a metal surface when illuminated by light of sufficient frequencies. He clarified that light is composed of discrete particles known as photons, and each photon carries an energy proportional to its frequency [4]. Einstein was able to theorize one of the central foundations of quantum mechanics with this explanation, concluding that light itself is quantized and that it does indeed behave like a particle.

Four years later, in 1909, G.I. Taylor performed another version of the double-slit experiment and showed that a light source as weak as "a candle burning at a distance exceeding a mile" could lead to interference fringes, showing that no matter the dimness of the light source, the serial pattern would still emerge [5]. This was not explained until after the emergence of quantum mechanics years later, which proved the wave-particle duality of light.

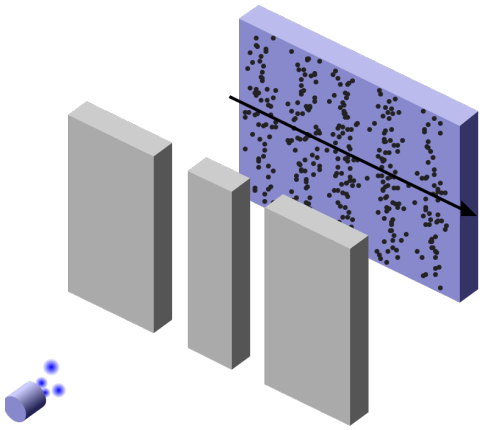

Figure 1.1: Public domain image of the double slit experiment, extracted from Wikimedia Commons. When illuminating the double-slit experiment with low intensity, single particles can be detected on the screen. An interference pattern emerges as these particles continue building up.

Once physicists gained an understanding of photons, they quickly turned to discerning how to control them. Perhaps the earliest example of controlling photons to a precise degree was the introduction of the laser in 1960, which was invented by Theodore Maiman and has the ability to generate an intense, narrow beam of light of a single wavelength [6].

It was clear that scientists were becoming more and more capable of altering physical phenomena at a smaller and smaller scale by this point in time. In 1959, one year prior to the introduction of the laser, the American Physical Society hosted the "Winter Meeting in the West," where Richard Feynman presented several hundred physicists with a question: "What would happen if we could arrange the atoms one by one the way we want them?" Feynman continued by outlining his vision, in which humanity achieves ultra-dense storage, nanobots used for medical applications, imaging at the nanometric scale, and more. Feynman's simple question laid the foundation of nanotechnology, the understanding and control of matter at the nanoscale [7]. This field, coupled with major developments in optical structures such as that of the laser, birthed nanophotonics, where the interaction of individual photons and

atoms are studied and exploited to enable novel devices. The interactions of atoms and photons are now known as light-matter interactions.

Today, we are capable of both designing and manufacturing devices at the nanometer scale that are used to manipulate and control light-matter interactions, a major technological feat in itself. Manufacturing techniques such as electron beam lithography [8] are capable of physically realizing such devices. This method involves utilizing beams of electrons to etch patterns into materials that are small enough to change how photons propagate and interact with the material at the atomic scale. But how are devices meant for controlling individual photons designed at the nanometric scale? How can they be simulated to ensure they achieve their designated purpose? Moreover, how can we ensure these devices function to their fullest potential? In this paper, we explore an efficient approach to designing, simulating, and optimizing the efficiency of nanophotonic devices, focusing largely on a device known as a line-defect cavity.

Figure 1.2: Public domain image of Richard Feynman, the father of nanotechnology. Published in 1986 and extracted from The Big T, the yearbook of the California Institute of Technology.

Chapter 2

Background

Nanophotonic devices are used to manipulate light at sub-wavelength scales. Line cavities are a type of device used for confining light for extended periods of time. This allows for more light-matter interactions between the trapped photons and the atoms of the cavity to occur. These devices are seen in a wide range of applications, ranging from light-based sensors used to characterize health-related biomarkers, to devices that exploit the properties of individual photons to enable quantum computing, and many more. The performance of such cavities are measured by the quality factor of their resonant modes, which is the ratio of stored energy to energy loss over time. High quality factor cavities are generally sought after because they offer higher energy efficiency and narrow resonance linewidths.

Cavities, like all other nanodevices, can be computationally designed, simulated, and optimized for enhanced performance. This chapter reviews the techniques and theory used for this process. Section 2.1 introduces line cavities and discusses the mechanics responsible for confining light within them, as well as the theoretical framework explaining how the quality factor can be improved. Section 2.2 reviews the Finite-Difference Time-Domain method, which is used to simulate nanophotonic devices interacting with light. Finally, Section 2.3 explains the inverse design technique known as the adjoint method, which is

used to optimize the performance of nanophotonic devices.

While line cavities have been studied primarily in isotropic materials, the behavior in anisotropic materials is less understood. Successful simulations of anisotropic materials could lead to the development of more efficient nanophotonic cavities, as these materials may offer stronger and more efficient light-matter interactions.

2.1 Photonic Crystalline Nanodevices

Nanophotonics as it stands today can generally be composed from four material classes; plasmonic, dielectric, two-dimensional, and hybrid systems [8]. The methodology implemented in this paper is developed specifically for devices that employ two-dimensional photonic crystal architectures. Photonic crystals (PhCs) have spatially periodic refractive indices. The periodicity of a two-dimensional PhC is confined to the x/y plane of these structures, which allows for five types of lattices, including oblique, rectangular, tetragonal, trigonal, and hexagonal.

Nanophotonic devices based on PhCs are developed by creating patterned features, which can be either air holes or columns with a different refractive index than that of the host material, into a host material to form a periodic lattice. The appeal of utilizing such lattices is that these structures enable the formation of photonic band gaps. Similar to how crystals in semiconductors control electron flow, photonic band gaps prevent certain wavelengths of light from passing through, allowing for the control of light within PhCs. Defects, or regions that disrupt the patterned structuring of these columns, are then introduced to allow for optical states to exist in regions where band gap frequencies would otherwise be forbidden [9]. By tuning the geometry and positioning of these defects, light can be confined and controlled in desired directions or regions.

Because of their ability to confine light, PhCs are invaluable for creating nanophotonic

devices, primarily waveguides and cavities. Waveguides, created by introducing a line defect through a PhC slab, are used for confining and directing light in certain directions. By altering the positions and radii of the surrounding holes, the amount of energy loss is minimized and the transmitted optical flux maximized. Cavities are used for keeping light confined within the PhC slab. Because these devices are responsible for trapping light for extended periods of time, photons become more likely to interact with the atoms that make up the device, resulting in a higher quantity of light-matter interactions. Adjusting the holes around a cavity allows for less energy loss and longer energy storage.

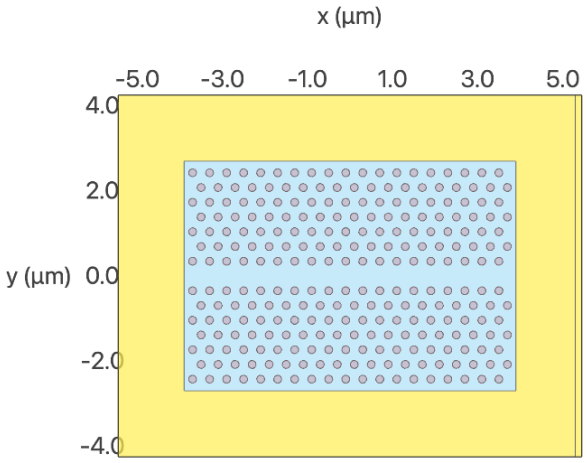

Figure 2.1: Lattice patterning for a simple, unoptimized waveguide, made from silicon. The blue region represents the silicon slab, the gray circles represent air columns at each lattice point, and the yellow represents background material where no explicit structure is defined. Each color represents a different medium, but the undefined yellow region has no medium that would effect the propagation of light. The slab is ∼780 nm long and ∼539 nm wide, with each defect having a radius of ∼20 nm. In this design, light can propagate efficiently in the x direction through the defect region, which is centered between the rows of air columns.

2.1.1 Line Cavities and Quality Factor

As mentioned, photonic band gaps are frequency ranges of light that cannot propagate through a PhC, or regions where no modes can exist [10]. Line-defect cavities (or L-cavities) are made by removing a distinct number of air holes (typically one to three) from a given row within an otherwise perfect PhC structure. When a line defect is introduced to an otherwise perfect lattice, the periodicity of the PhC in one direction is disrupted, resulting in new, allowed bands to exist within the otherwise forbidden band gap. These new bands are referred to as guided modes, or electromagnetic waves whose frequency lies in the band gap but are confined to and propagate along the defect line. The surrounding PhC forbids guided modes from being able to spread anywhere else outside of the defect region.

L-cavities have a distinct quality factor, which describes the temporal confinement of the guided mode within the cavity. A higher Q factor correlates to greater strength of the light-matter interaction within the cavity. L-cavities also have distinct resonant frequencies, or frequencies that lay within the photonic band gap of the PhC and correspond to electromagnetic fields that satisfy the boundary conditions of the PhC and Maxwell's equations (discussed in further detail in Chapter 3), resulting in maximal mode confinement.

The confinement of a guided mode in an L-cavity occurs due to two different optical phenomena. Energy is constricted in the longitudinal direction due to distributed Bragg reflection, and total internal reflection keeps the mode constricted vertically [11]. Bragg reflection in a PhC occurs due to the periodic modulation of refractive index that is enabled by the arrays of air holes employed throughout the host material [12]. When the guided mode confined to the missing defect region (also referred to as the cavity mode) extends into the surrounding PhC, it interacts with the periodic structure and excites scattering from each column site. For the wavelengths that lay in the photonic band gap of the crystal, the scattered contributions interfere constructively. The result is that any propagation in

the lattice region is suppressed, and the PhC itself functions as a Bragg reflector [11], a high-reflectivity mirror that is theoretically lossless and responsible for confining the mode to the defect region.

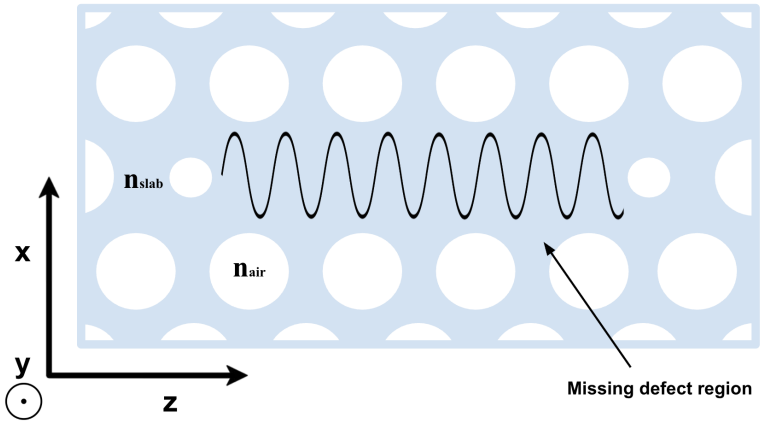

Figure 2.2: The missing defect region of an L-cavity and a guided mode confined within it, depicting how the device acts as a Bragg reflector in the longitudinal direction of the cavity. n_{slab} is the refractive index of the slab, which is higher than n_{air}, the refractive index of the air columns at each lattice point in the PhC. Both mediums meet the conditions for Bragg reflection.

In the vertical direction of an L-cavity, energy is stored due to total internal reflection. When light reflects off a surface that separates two mediums n_1 and n_2, as shown in Figure 2.3, the angle at which the ray strikes the surface is called the incident angle, θ_1, and the angle at which the ray passes through the medium to another is called the angle of refraction, θ_2 [13]. The incident angle becomes a critical angle if it causes the angle of refraction to equal 90°. If the incident angle is greater than the critical angle, all of the light is reflected back into medium n_1, rather than being partially refracted from medium n_1 to n_2. This condition is known as total internal reflection. In an L-cavity, the guided mode is confined vertically due to this condition. n_1 is the refractive index of the PhC, and n_2 is the refractive index of the medium surrounding the slab.

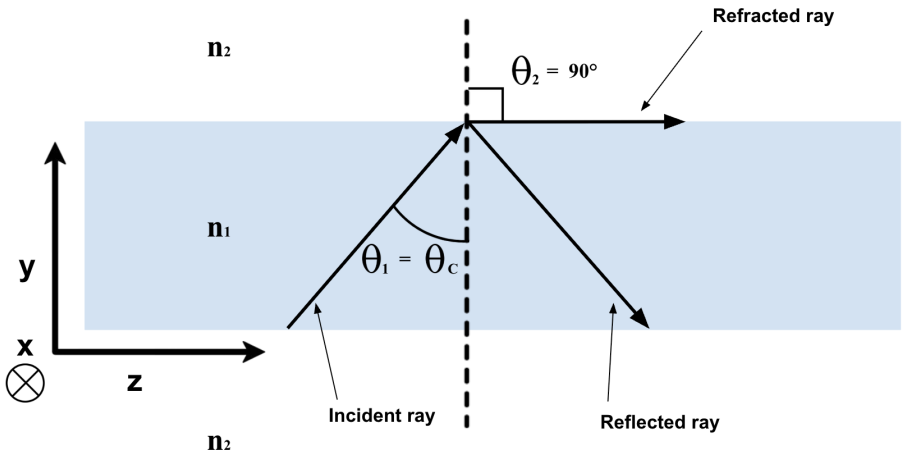

Figure 2.3: Side view of an L-cavity with medium with refractive index n_1, and a surrounding medium with refractive index n_2. The guided mode in the vertical direction has a critical angle such that the angle of refraction is 90°. If the incident ray angle increases past this point, total internal reflection is achieved. Therefore, all the light is reflected back into the slab medium n_1, and stays confined within the cavity region.

However, physically realized cavities do suffer from out-of-plane radiation loss largely due to manufacturing imperfections, among other reasons. Following the analysis in [11], the cavity mode can be decomposed into spatial frequency components described by the wave. vector, $\mathbf{k} = (k_x, k_y, k_z)$. $\mathbf{k}$ comes from the forward Fourier spatial frequency transform for an electric field distribution, which decomposes a spatial electric field function into components of spatial frequencies. The general formula for a forward Fourier transform is:

$$f(k_x) = \int f(x) e^{ik_x x} dx \tag{2.1}$$

Note here that the k can depend on any axis. As depicted in Figure 2.4, the guided mode is confined to the finite region of the cavity and has in-plane wavevector components of magnitude $k_{\|} = \sqrt{k_x^2 + k_z^2}$. The spatial confinement of the mode implies a spectrum of these components. Components with $k_{\|} < \frac{\omega_0}{c} n_2$, where ω_0 is the resonant frequency, c is the

speed of light, and n_2 is the refractive index of the surrounding medium, fall into what is known as the light cone. This is a region that does not fulfill requirements for total internal reflection, and therefore these components can couple to modes outside of the slab. This type of energy dissipation is referred to as out-of-plane radiation loss.

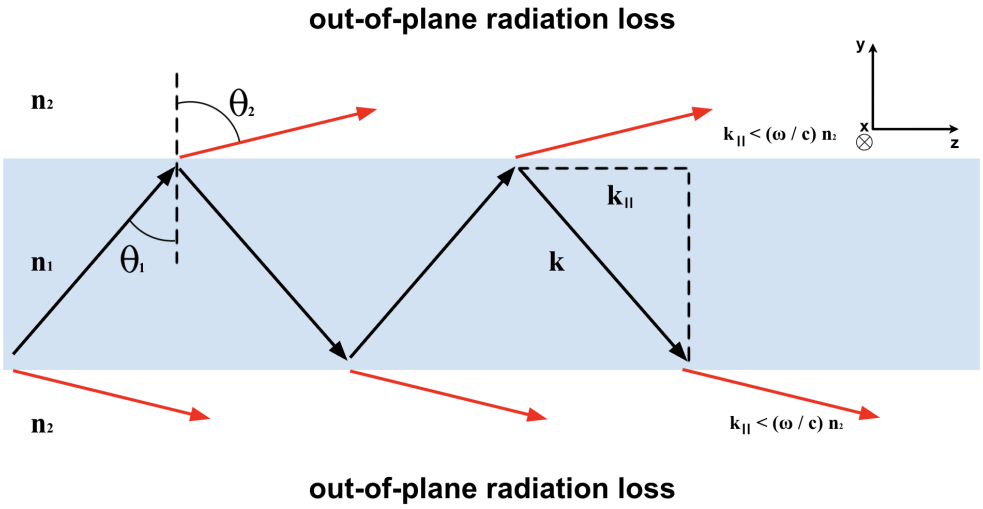

Figure 2.4: When a ray strikes the surface between n_1 and n_2 of a cavity, total internal reflection occurs if the incident angle, θ_1, is greater than critical angle. The waves with $k_{\|}$ components that fall into the light cone region do not fulfill the conditions for total internal reflection, and leak out of the slab in the vertical direction as a result, resulting in out-of-plane radiation loss.

The quality factor of a line cavity can be described by the formula

$$Q = \omega_0 \frac{U}{P} \tag{2.2}$$

where U is the electromagnetic energy stored within the cavity and P is the energy lost per unit time. The goal of designing any cavity is to increase the Q factor as much as possible, which means both increasing total energy stored and decreasing the radiation loss. It is worth noting the Q factor is limited by factors such as the initial cavity design and refractive

index of the PhC, among other parameters.

2.1.2 Improving the Quality Factor of Line Cavities

Maximizing the quality factor of a cavity involves smoothing what is referred to as the spatial envelope, which describes how the amplitude of the mode varies across the cavity, so that the envelope varies smoothly at the edges of the cavity. Following the line of reason from [11], we can consider a one-dimensional cavity mode,

$$\mathbf{E}(x) = \mathbf{A}(x)sin(k_x x) \tag{2.3}$$

where $A(x)$ describes how the amplitude of the cavity mode varies across the missing defect region, referred to as the cavity field envelope. Rapid spatial variations in $A(x)$ would enable a wider range of in-plane wave vector components, via the Fourier transform. If this distribution is wider, such components will fall into the light cone region. As a result, abrupt spatial features correspond to higher out-of-plane radiation loss and a smaller cavity Q factor.

Optimizing the cavity Q factor therefore requires avoiding abrupt spatial variations in the field. This can be achieved by adjusting the holes around the cavity, allowing the mode to extend further into the surrounding lattice and decay more gradually near the edges. The spatial profile of the mode becomes wider and smoother as a result, which narrows the distribution of in-plane wave vector components. Fewer components end up falling into the light cone region, reducing out-of-plane radiation loss and increasing Q.

2.1.3 Anisotropic Materials

Anisotropic materials are birefringent, meaning they have directional dependence of physical properties [14]. In this case, we will be focused on materials with refractive indices that change in the x, y, and z directions. Additionally, materials that have a refractive index dependent on the wavelength of light traveling through it are referred to as dispersive. Materials can be anisotropic and dispersive simultaneously. Little is known about the behavior of anisotropic materials when used for developing nanophotonic devices. Physicists have historically used isotropic materials for the design of nanodevices, or materials that have directional independence of physical properties.

One reason such anisotropic materials are currently of interest in the nanophotonics community is that they offer greater control over light polarization [15]. This is useful for mechanics like optical switching [16], in which the flow of light is controlled like an on and off switch, and is often implemented in photonic integrated circuits. Another useful mechanism offered by realizing anisotropic devices is multiplexing [17], where different spatial modes are used to carry distinct information streams, often used in few-mode fibers. A major appeal of realizing anisotropic line-defect cavities specifically is due to their enhanced energy confinement [18]. These devices can be fine tuned to confine electromagnetic waves in specific directions, allowing for more enhanced light-matter interactions. Similarly, such cavities also allow for directional flow of energy, which can be useful for developing waveguides.

2.2 The FDTD Method

The Finite-Difference Time-Domain method has proven to be an extremely capable computational method for designing and simulating how electromagnetic fields propagate and interact with different structures and mediums at the nanometer scale. Introduced in 1966 by electrical engineer and mathematician Kane Yee, the FDTD method is capable of simu-

lating the interaction of light with materials and optical devices in three-dimensional space. This method has become an industry standard due to its sheer versatility. It can calculate optical properties such as transmission, interference, absorption, and more, of basic materials or complex devices. It is also capable of analyzing the propagation and wavelengths of light, and can even provide moving visualizations of how light travels through these materials and devices in real time [19].

2.2.1 Maxwell's Equations in the Time Domain

To develop a means of computationally simulating how electromagnetic waves interact with nanodevices, the first step Kane Yee took in developing this scheme was selecting mathematical formulas that could be used to calculate electric and magnetic field values as they traverse through different mediums. Fortunately, about a century prior to Yee's scheming, James Clerk Maxwell published "A Dynamical Theory of the Electromagnetic Field," which outlines one of the most influential sets of equations in all of physics. Known as Maxwell's equations, the four equations describe how electric and magnetic fields propagate and interact, as well as how they are influenced by matter. Moreover, they even showed that the speed of light happens to be the same speed of the propagation of electromagnetic waves, meaning that light and electromagnetic waves are one in the same.

Of the four equations, Yee selected the Maxwell-Faraday and Maxwell-Ampère equations, which are shown respectively in differential form in Equation 2.4 and Equation 2.5,

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \tag{2.4}$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \mathbf{J}(\mathbf{r}, t) + \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} \tag{2.5}$$

where $\mathbf{E}$ is the electric field vector (V/m), $\mathbf{B}$ is the total magnetic field vector (T), $\mathbf{H}$ is the free magnetic field vector, or the field resulting from a current source (A/m), $\mathbf{J}$ is the free current density vector (A/m^2), and $\mathbf{D}$ is the electric flux density vector (C/m^2). Each vector quantity here is dependent on position, $\mathbf{r}$, and time, t.

In free space, $\mathbf{B}(\mathbf{r}, t)$ can be defined as

$$\mathbf{B}(\mathbf{r}, t) = \mu_0 \mathbf{H}(\mathbf{r}, t) \tag{2.6}$$

where μ_0 is the permeability of free space. The electric flux density, $\mathbf{D}(\mathbf{r}, t)$, can also be expressed as

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon_r \varepsilon_0 \mathbf{E}(\mathbf{r}, t) \tag{2.7}$$

where ε_r is the relative permittivity of a material, and ε_0 is the permittivity of free space. Substituting Equation 2.6 and Equation 2.7 into Equation 2.4 and Equation 2.5, respectively, yields:

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\mu_0 \frac{\partial \mathbf{H}(\mathbf{r}, t)}{\partial t} \tag{2.8}$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \mathbf{J}(\mathbf{r}, t) + \varepsilon_r \varepsilon_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \tag{2.9}$$

If we know what $\mathbf{E}(\mathbf{r}, t)$ or $\mathbf{H}(\mathbf{r}, t)$ is for a given simulation, information about how the other missing component propagates can be found using the combination of Equation 2.8 and Equation 2.9. Equation 2.9 also contains two values of importance. ε_r is the relative permittivity of a material and can be expressed as $\varepsilon_r = n^2$, where n is the refractive index of a material. The refractive index describes how light propagates through the medium, and is expressed as the ratio of the speed of light in a vacuum to the speed in a medium. In the

context of nanophotonics, ε_r is used to describe the structure of a device. For example, if we are interested in simulating a waveguide mounted on a substrate of a different medium, the simulation differentiates which structure is which based on the relative permittivity of each. The second variable is $\mathbf{J}(\mathbf{r}, t)$, which is zero in a vacuum or free space, and can be used to describe a source current that injects light into a system. An electromagnetic source point can be defined at any region in a simulation, for which $\mathbf{J}(\mathbf{r}, t)$ would become a nonzero vector value.

In the FDTD method, the goal is to solve both Equation 2.8 and Equation 2.9 in order to yield distributions for the electric field, $\mathbf{E}(\mathbf{r}, t)$, and magnetic field, $\mathbf{H}(\mathbf{r}, t)$, throughout the entire simulation, given ε_r and $\mathbf{J}(\mathbf{r}, t)$. This idea is also where the latter half of the name Finite-Difference Time-Domain comes from, as the variables of interest are all time-dependent, rather than frequency-dependent [20].

2.2.2 Discretizing Maxwell's Equations

Maxwell's equations provide the opportunity to calculate information about how light propagates across a space and interacts with different mediums. But how is it possible for field values to be calculated computationally? If field values change and propagate differently at each point in a space, how is this accounted for and recorded as data?

Moreover, to observe the effects of light-matter interactions properly, a simulation that is responsible for calculating $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{H}(\mathbf{r}, t)$ must have a space that is large enough to observe far-field electromagnetic wave effects. The far-field region occurs past one wavelength in distance from an electromagnetic field source point, and is where the electric and magnetic waves are coupled, orthogonal, and propagate together. Within the near-field region, or the distance spanning one wavelength from a source point, electromagnetic waves do not behave in the same way [21]. Here, electric and magnetic waves are not necessarily in phase with

each other and field patterns are difficult to predict as a result. In order to observe far-field effects and avoid any sources of error spanning from the near-field region, a simulation must be larger than one wavelength, and is typically multiple wavelengths large in practice. To this end, how can we ensure that the field information throughout a large simulation size is precise and reflects physically tangible values at each point in space and time?

Fortunately for us, the Finite-Difference half of Yee's Finite-Difference Time-Domain method provides an answer to this question by allowing for Maxwell's equations in the time domain to be discretized and calculated. We will examine the case of computing the electric and magnetic field values within a one-dimensional simulation to explain this method, meaning that all field values depend only on a single axis. Beginning with Maxwell's equations, the cross products of Equations 2.8 and 2.9 are found first in order to describe the force components in the x, y, and z directions:

$$\nabla \times \mathbf{E} = \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right)\hat{x} - \left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z}\right)\hat{y} + \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right)\hat{z} \tag{2.10}$$

$$\nabla \times \mathbf{H} = \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right)\hat{x} - \left(\frac{\partial H_z}{\partial x} - \frac{\partial H_x}{\partial z}\right)\hat{y} + \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\hat{z} \tag{2.11}$$

Because we are dealing with a one-dimensional simulation, we will focus only on the electric and magnetic waves varying along the z axis (meaning that $\frac{\partial}{\partial x} = \frac{\partial}{\partial y} = 0$ here, and field components $E_x(z,t)$, $E_y(z,t)$, $E_z(z,t)$ may still exist). The cross products therefore reduce to

$$\nabla \times \mathbf{E} = -\frac{\partial E_y}{\partial z}\hat{x} + \frac{\partial E_x}{\partial z}\hat{y} \tag{2.12}$$

$$\nabla \times \mathbf{H} = -\frac{\partial H_y}{\partial z}\hat{x} + \frac{\partial H_x}{\partial z}\hat{y} \tag{2.13}$$

Substituting Equations 2.12 and 2.13 into Equations 2.8 and 2.9 respectively yields

$$\begin{aligned} \frac{\partial E_x}{\partial t} &= -\frac{1}{\epsilon_0}\frac{\partial H_y}{\partial z} \qquad \frac{\partial H_y}{\partial t} = -\frac{1}{\mu_0}\frac{\partial E_x}{\partial z} \\ \frac{\partial E_y}{\partial t} &= \frac{1}{\epsilon_0}\frac{\partial H_x}{\partial z} \qquad \frac{\partial H_x}{\partial t} = \frac{1}{\mu_0}\frac{\partial E_y}{\partial z} \end{aligned} \tag{2.14}$$

These are Maxwell's equations in Cartesian form, describing how each field component propagates in a one dimensional space. The expressions are two pairs of first order, coupled partial differential equations, meaning E or H could be calculated with one expression and substituted into the other to find the corresponding unknown value. It is also worth noting that $\mathbf{J}(\mathbf{r}, t) = 0$ and $\varepsilon = 1$ for this problem. In practice, ε would have some value that does not equal one and the vector $\mathbf{J}(\mathbf{r}, t)$ would have some magnitude and direction. Both values can easily be integrated into Equation 2.14.

Because these expressions are continuous, however, it would be impossible for a simulation to derive the values for E and H as the fields propagate and interact with materials in a large dimensional space. Yee was able to address this by discretizing the entire simulation into a grid made of "Yee Cells." Each cell contains spatial dimensions and a time dimension. In our one-dimensional case, the spatial dimension is $z = k\Delta z$, with k being an integer spatial point, and the time dimension is $t = n\Delta t$, with n being an integer time step. Note that $f(k\Delta z, n\Delta t) = f_k^n$. The size of a Yee Cell in one dimension is defined by Δz, with each integer spatial point k denoting the grid cell in which a given field component is evaluated. If the simulation were three-dimensional, each cell size would be defined by Δx, Δy, and Δz, and the same would be true of its spatial integer points in all three directions. Typical dimensions of these cells range from one twentieth to one fiftieth of a wavelength, which allows for very precise measurements of the fields across the entire grid.

In order to discretize the simulation into Yee Cells, we first take the derivatives of Equation 2.14 and replace them with derivative approximations to yield new expressions. For the

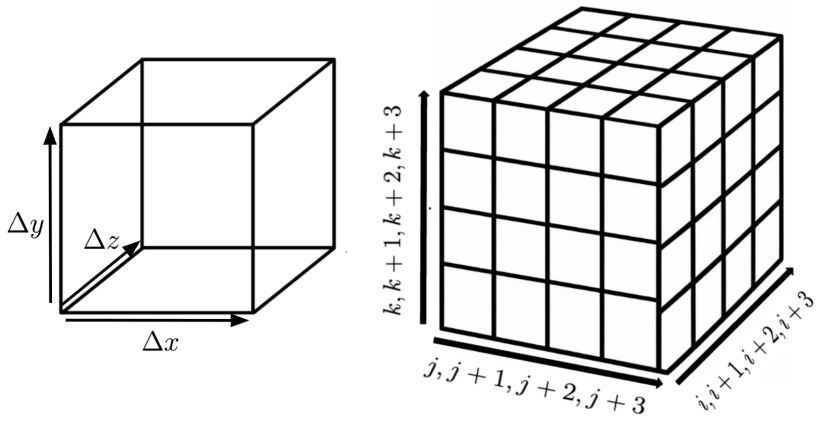

Figure 2.5: Left: Yee Cell with dimensions Δx, Δy, and Δz. Each cell in a simulation also has a time dimension. For each time step, a cell calculates E and H. Right: A three-dimensional simulation discretized into Yee Cells, with each cell denoted by spatial coordinates i, j, and k. Because each cell deciphers E and H for each time step, information about how electromagnetic fields propagate through an entire simulation space over time can be calculated computationally.

sake of explanation, we will only do so for the top two expressions in Equation 2.14:

$$\frac{E_{x,k}^{n+1} - E_{x,k}^{n}}{\Delta t} = \frac{1}{\epsilon_0} \frac{H_{y,k+1}^{n+1/2} - H_{y,k}^{n+1/2}}{\Delta z} \tag{2.15}$$

$$\frac{H_{y,k}^{n+1/2} - H_{y,k}^{n-1/2}}{\Delta t} = -\frac{1}{\mu_0} \frac{E_{x,k+1/2}^{n} - E_{x,k-1/2}^{n}}{\Delta z} \tag{2.16}$$

Examining Equation 2.15, the time derivative of E_x is centered at $n + 1/2$. This means that if H_y is known at time $n+1/2$, then E_x can be updated from time n to $n+1$ correspondingly. Knowing that we now have found E_x for time steps n and $n+1$, we can then evaluate the time derivative of Equation 2.16, H_y, which is centered at n. Using our known values of E_x, H_y can be updated from time $n - 1/2$ to $n + 1/2$. As a result, we can circle back to Equation 2.15 using our now known values of H_y for the following time coordinates, $n+1/2$ and $n - 1/2$, to find E_x at the next time coordinates of $n + 1$ and n. The process continues

over and over again until all values of E_x and H_y are computed for every time point. This method is known as a leapfrog algorithm; starting with E_x^n, $H_y^{n+1/2}$ can be computed, which is then used to compute E_y^{n+1}, and the cycle continues throughout the entire time and space domain. Each field is updated using the other field that was evaluated half a step earlier in time.

Each variable in these expressions can now be calculated computationally for every cell and time step throughout the entire simulation. The same substitution made here can, of course, be made for the second pair of partial differential equations in Expression 2.14. Even more impressively, this scheme does not require storing values in a matrix, which means it is quite computationally efficient in comparison to other simulation methods. Instead, the field values are stored in arrays, allowing for rather simplistic data access for plotting energy spatial distributions, energy decay plots, and much more.

The final form of Equations 2.15 and 2.16 used within the FDTD method, as described by [22], are

$$E_{x,k}^{n+1} = \frac{\Delta t}{\epsilon_0 \Delta z}\left(H_{y,k+1}^{n+1/2} - H_{y,k}^{n+1/2}\right) + E_{x,k}^{n} \tag{2.17}$$

$$H_{y,k}^{n+1/2} = \frac{\Delta t}{\epsilon_0 \Delta z}\left(E_{x,k-1/2}^{n} - E_{x,i+1/2}^{n}\right) + H_{y,k}^{n-1/2} \tag{2.18}$$

In real simulations, ϵ is typically not equal to one and $\mathbf{J}(\mathbf{r}, t)$ will have some value. Incorporating both values into Equations 2.17 and 2.18 yields:

$$E_{x,k}^{n+1} = \frac{\Delta t}{\epsilon\epsilon_0 \Delta z}\left(H_{y,k+1}^{n+1/2} - H_{y,k}^{n+1/2}\right) + E_{x,k}^{n} - \frac{\Delta t}{\epsilon\epsilon_0 \Delta z} J_{x,k}^{n} \tag{2.19}$$

$$H_{y,k}^{n+1/2} = \frac{\Delta t}{\epsilon\epsilon_0 \Delta z}\left(E_{x,k-1/2}^{n} - E_{x,k+1/2}^{n}\right) + H_{y,k}^{n-1/2} \tag{2.20}$$

To show that this process still applies to three-dimensional simulation, the continuous form of Maxwell's equations in three-dimensional space shown to be

$$\begin{aligned}
\frac{\partial E_x}{\partial t} &= \frac{1}{\epsilon\varepsilon_0}\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right) - J_x, \\
\frac{\partial E_y}{\partial t} &= \frac{1}{\epsilon\varepsilon_0}\left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right) - J_y, \\
\frac{\partial E_z}{\partial t} &= \frac{1}{\epsilon\varepsilon_0}\left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right) - J_z, \\
\frac{\partial H_x}{\partial t} &= \frac{1}{\epsilon\varepsilon_0}\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right), \\
\frac{\partial H_y}{\partial t} &= \frac{1}{\epsilon\varepsilon_0}\left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right), \\
\frac{\partial H_z}{\partial t} &= \frac{1}{\epsilon\varepsilon_0}\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right)
\end{aligned} \tag{2.21}$$

and the process by which their derivative approximation forms are found is the same. The spatial coordinates for such a system would be $f(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = f^n_{i,j,k}$, meaning each cell contains three spatial dimensions and one time dimension.

2.3 The Adjoint Method for Nanophotonic Optimization

We have outlined the framework for how nanophotonic devices are designed in theory, specifically line-defect cavities, and how they can be fine-tuned to enhance the quality factor, Q. We have also determined how such devices are simulated using the FDTD method. This method is responsible for breaking up a simulation into Yee Cells and calculating the E and H field components for each cell and time step throughout the entire simulation, which allows us to see how electromagnetic fields interact with cavities. Using these two methodological frameworks, we can now examine how a computational optimization process is performed to

maximize the Q of a given cavity. This section describes how the parameters of any nanophotonic device can be adjusted in order to optimize the device, with a machine learning and inverse design technique known as the adjoint method.

2.3.1 Forward and Adjoint Simulations

In order to understand the adjoint method, we first must understand the processes behind what are referred to as forward simulations and adjoint simulations. Let us examine the case of optimizing a line-defect cavity, where the goal is to maximize the quality factor Q. As we know, this is achieved by adjusting the positioning of the missing-defect region to smooth the spatial envelope and reduce out-of-plane radiation loss. This can be turned into a mathematical optimization problem, where Q is the objective function, and the hole positions are the parameters that are adjusted to maximize the objective function. Because adjusting the parameters changes the sensitivity of the objective function, this sensitivity can be quantized as a gradient, and adjusting the parameters by some amount affects the objective function in turn. This means that the optimization problem corresponds to calculating the gradient and moving along the objective function curve in the direction determined by the gradient to reach the maxima.

As noted by Equation 2.2, our objective function is determined via the ratio of stored energy within the cavity to the energy leakage per cycle. A forward simulation, which is just a normal simulation carried out with the FDTD method, is responsible for measuring both quantities. In an forward simulation, each can be measured with field monitors, or spatial regions that record the field components at a specified time or frequency [23]. The stored energy monitor encompasses the missing defect region of the cavity, and the energy leakage monitor is placed directly above the cavity along the axis where the field components travel out of the device. Finally, the field source point is placed in the center of the cavity.

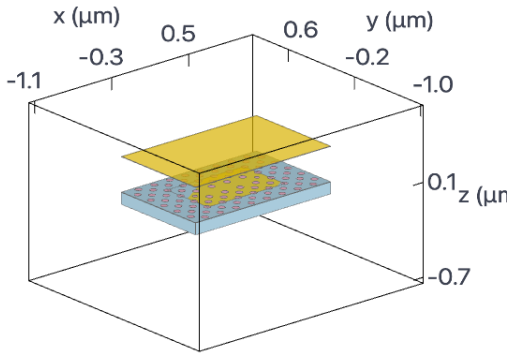

Figure 2.6: Three-dimensional simulation of a line defect cavity. Above the cavity is an energy leakage monitor along the x and y axes, which is responsible for measuring the out-of-plane radiation loss. Centered around the missing defect region is a three-dimensional stored energy monitor, responsible for measuring the electromagnetic energy stored in the region over time. These measurements are recorded during a forward FDTD simulation.

The objective function can be called $\mathbf{F}(\mathbf{e}(p))$, as $\mathbf{F}$ is dependent on the electromagnetic field distribution $\mathbf{e}$, which is dependent on parameter p. To compute how $\mathbf{F}$ changes with respect to p, we first take the derivative, accounting for both the real and imaginary parts of the field, resulting in

$$\frac{d\mathbf{F}}{dp} = \frac{\partial \mathbf{F}}{\partial \mathbf{e}}\frac{d\mathbf{e}}{dp} + \frac{\partial \mathbf{F}}{\partial \mathbf{e}^*}\frac{d\mathbf{e}^*}{dp} + \frac{\partial \mathbf{F}}{\partial p} \tag{2.22}$$

which simplifies to

$$\frac{d\mathbf{F}}{dp} = 2\Re\left(\frac{\partial \mathbf{F}}{\partial \mathbf{e}}\frac{d\mathbf{e}}{dp}\right) + \frac{\partial \mathbf{F}}{\partial p} \tag{2.23}$$

The $\frac{\partial \mathbf{F}}{\partial \mathbf{e}}$ term is easy to find. This value is simply a normal FDTD simulation that computes the electromagnetic field distribution from a physical light source. The challenging part is calculating $\frac{d\mathbf{e}}{dp}$, as there are n amounts of parameters that could be positioned in n

amounts of directions. To computationally calculate this value in an efficient manner, we first must derive a single wave equation that maps the propagation of an electromagnetic field. First, we substitute Equations 2.6 and 2.7 into Equations 2.4 and 2.5, respectively, yielding:

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\mu_0 \frac{\partial \mathbf{H}(\mathbf{r}, t)}{\partial t} \tag{2.24}$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \mathbf{J}(\mathbf{r}, t) + \epsilon_0 \epsilon(\mathbf{r}) \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \tag{2.25}$$

Note here that ϵ is dependent on parameter r here because the permittivity is dependent on the spatial point of a simulation. We next calculate the curl of Equation 2.24, which results in

$$\nabla \times (\nabla \times \mathbf{E}(\mathbf{r}, t)) = -\mu_0 \frac{\partial}{\partial t} (\nabla \times \mathbf{H}(\mathbf{r}, t)) \tag{2.26}$$

Next, substituting Equation 2.25 into Equation 2.26:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}, t) = -\mu_0 \frac{\partial}{\partial t} \left(\mathbf{J}(\mathbf{r}, t) + \epsilon_0 \epsilon(\mathbf{r}) \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \right) \tag{2.27}$$

Distributing and rearranging the derivative produces the expression:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}, t) + \mu_0 \epsilon_0 \epsilon(\mathbf{r}) \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} = -\mu_0 \frac{\partial \mathbf{J}(\mathbf{r}, t)}{\partial t} \tag{2.28}$$

This expression is referred to as a time-domain inhomogeneous equation of an electromagnetic field [24]. It describes electromagnetic wave propagation caused by a source, denoted by the term on the right side of the equation. However, the expression needs to be rewritten in terms of frequency rather than time. To achieve this, a Fourier transformation is used:

$$\mathbf{E}(\mathbf{r},\omega) = \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r},t)e^{i\omega t}dt \tag{2.29}$$

where ω is frequency (Hz). Note that we are performing the Fourier transform on $\mathbf{E}(\mathbf{r},t)$ but it may be used on any time-dependent variable. The Fourier transform of a time derivative, as expressed in Equation 2.28, is shown to be

$$\mathcal{F}\left(\frac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t}\right) = \int_{-\infty}^{\infty} \frac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t} e^{i\omega t}dt = -i\omega \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r},t)e^{i\omega t}dt \tag{2.30}$$

In simplistic terms, a time derivative when converted to frequency [25] via the Fourier transform can be denoted as

$$\mathcal{F}\left(\frac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t}\right) = i\omega \mathbf{E}(\omega) \tag{2.31}$$

Using this substitution, Equation 2.28 can be rewritten as:

$$(\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - k_0^2 \epsilon(r))\mathbf{E}(\mathbf{r}) = i\omega\mu_0 \mathbf{J}(\mathbf{r}) \tag{2.32}$$

where $k_0^2 = \omega^2 \mu_0 \epsilon_0$. Finally, we perform one last substitution using the relation

$$\mathbf{J}(\mathbf{r}) = \frac{\partial \mathbf{P}(\mathbf{r},t)}{\partial t} = -i\omega \mathbf{P}(\mathbf{r}) \tag{2.33}$$

where $\mathbf{P}(\mathbf{r})$ is a positional dipole moment (C·m). Our final frequency-domain inhomogeneous equation of an electromagnetic field [26] is

$$(\nabla \times \nabla \times -k_0^2 \epsilon_r(\mathbf{r}))\mathbf{E}(\mathbf{r}) = \omega^2 \mu_0 \mathbf{P}(\mathbf{r}) \tag{2.34}$$

which describes the propagation of an electromagnetic wave due due to a dipole moment source. This is the expression that is used to calculate $\frac{de}{dp}$. In order to implement this wave

equation into the adjoint method, we next linearize Equation 2.34 to

$$A\mathbf{e} = \mathbf{b} \tag{2.35}$$

where $A = (\nabla \times \nabla \times -k_0^2\epsilon_r(\mathbf{r}))$, $\mathbf{e} = \mathbf{E}(\mathbf{r})$, and $\mathbf{b} = \omega^2\mu_0\mathbf{P}(\mathbf{r})$. The A component is in discretized operator form, and encodes the how the fields propagate and couple to points in the simulation space. e is a vector holding all the field components at all spatial points, and $\mathbf{b}$ is the field source point. Rewriting 2.35 yields

$$\mathbf{e} = A^{-1}\mathbf{b} \tag{2.36}$$

and tells us that $\mathbf{E}(\mathbf{r})$ can be calculated given A^{-1} and $\mathbf{b}$. Differentiating and rearranging Equation 2.35,

$$\frac{d\mathbf{e}}{dp} = -A^{-1}\frac{dA}{dp}\mathbf{e} \tag{2.37}$$

which can then be substituted into 2.23 to produce

$$\frac{d\mathbf{F}}{dp} = 2\Re\left(-\frac{\partial\mathbf{F}}{\partial\mathbf{e}}A^{-1}\frac{dA}{dp}\mathbf{e}\right) + \frac{\partial\mathbf{F}}{\partial p} \tag{2.38}$$

In order to make this expression more computationally efficient, a new source known as the adjoint source is defined, where

$$\mathbf{e}_{adj}^T = -\left(\frac{\partial\mathbf{F}}{\partial\mathbf{e}}\right)A^{-1} \tag{2.39}$$

$$A\mathbf{e}_{adj} = -\left(\frac{\partial\mathbf{F}}{\partial\mathbf{e}}\right)^T \tag{2.40}$$

which is in the same form as the linear system defined in Equation 2.35. The adjoint source

here is defined as the negative transpose of the objective gradient with respect to the field. It acts as a virtual source that is placed at the output of the objective function, or the monitors of a given simulation. An adjoint simulation then solves Maxwell's equations from this new virtual source point to compute how sensitive the objective function is to changes in the parameters of the device. Substituting this expression into Equation 2.38,

$$\frac{d\mathbf{F}}{dp} = -2\Re\left(\mathbf{e}_{adj}^{T}\frac{dA}{dp}\mathbf{e}\right) + \frac{\partial\mathbf{F}}{\partial p} \tag{2.41}$$

Fortunately, the remaining $\frac{dA}{dp}$ and $\mathbf{e}$ term can be resolved rather simplistically. The A component from the linearized form of Maxwell's equation 2.35 notes that

$$A \equiv \nabla \times \nabla \times -k_0^2\epsilon_r(\mathbf{r}) \tag{2.42}$$

and its derivative with respect to p would be

$$A \equiv -k_0^2\frac{d\epsilon_r(\mathbf{r})}{d\mathbf{r}} \tag{2.43}$$

The permittivity $\epsilon_r(\mathbf{r})$ is defined at each point in the simulation as well. The structure being simulated will have some permittivity value, and outside of the structure it will become zero. Computationally, this is treated as a diagonal matrix that stores the value 1 for points inside the structure and 0 for points outside of it. Therefore, it can be written as a Kronicker delta function:

$$\frac{d\epsilon_r(\mathbf{r})}{d\mathbf{r}} = \delta_{r\in p} \tag{2.44}$$

and substituted back into Formula 2.41 to produce

$$\frac{d\mathbf{F}}{dp} = 2\Re\left(\mathbf{e}_{adj}^{T}\delta_{r\in p}\mathbf{e}\right) + \frac{\partial\mathbf{F}}{\partial p} \tag{2.45}$$

which simplifies to the final expression for the adjoint method [27],

$$\frac{d\mathbf{F}}{dp} = -2k_0^2 \sum_{i \in p} \Re \left(\mathbf{e}_{adj}^{(i)} \mathbf{e}^{(i)} \right) + \frac{\partial \mathbf{F}}{\partial p} \tag{2.46}$$

The beauty behind the adjoint method is how computationally efficient it truly is. As expressed in Equation 2.46, the full sensitivity of the objective function with respect to every single parameter can be obtained using only one forward simulation and one adjoint simulation. This means that instead of running thousands of simulations in order to determine how each parameter adjustment affects the overall objective function, one forward simulation and one adjoint simulation can be run instead to yield a sensitivity map that encodes how shifts in every parameter would effect the value returned by the objective function. The gradient produced is simply the product of the two values, and sums only over the regions where the parameters have influence on the function.

Chapter 3

Methodology

This chapter describes the computational methodology utilized to design, simulate, and optimize an anisotropic line-defect cavity, based on the theory described in Chapter 2. Section 3.1 reviews the parameters that were implemented to design the initial cavity, how the anisotropic nature of the PhC material is accounted for, and the source point and the monitors parameters used to enable the simulations. Section 3.2 discusses how the objective function used for the adjoint method optimization of the cavity was designed. Finally, Section 3.3 describes both the optimizer that was implemented to enable the adjoint method of optimization, as well as the parameters that were utilized for the optimization iteration loop. This procedure was developed through the usage of the FDTD simulation platform known as Tidy3D [23].

3.1 Device Parameters

The cavity is missing three holes from it's center (referred to as an L3 cavity) and is made from tungsten disulfide. The initial design of the L3 cavity implemented a triangular PhC lattice. The spatial constant between lattice points, d, was approximated to be equal to the

lattice constant, a, for the purpose of simplifying the parameters of the cavity structure. d was determined with a simple approximation,

$$d = \frac{\lambda}{2n_d} \tag{3.1}$$

where n_d is the medium refractive index at the central wavelength, m is the diffraction order ($m = 1, 2, 3...$), and λ is the wavelength passing through the PhC medium. The resulting spacing periodicity used was 80.295 nm. This formula is a simplified version of the Bragg condition [28], where

$$2d \sin\theta = m\lambda \tag{3.2}$$

with θ being the diffraction angle. In Formula 3.2, we assume that $\sin\theta = 1$ and that $m = 1$. Another reason this expression is an oversimplification is because it does not take into account the interplanar spacing of the crystal structure. This is a relation used in crystallography to describe the orientation of crystal planes and directions in a lattice [29]. This framework also establishes a relationship between the spatial periodicity d and the lattice constant a; as mentioned, here we are assuming the two to be equal, which is not traditionally done in practice. Details of the relation between d and a for a triangular lattice in a physically tangible crystal structure can be found in [30].

The radius of holes is typically between $0.25a$ and $0.35a$ for most cavities and waveguides [31]. Initial simulations yielded strong energy confinement utilizing a radius of $0.35a$, which corresponds to 28.103 nm. Nine rows and eleven columns of holes were chosen for this device. Initial simulations indicated that a higher number of rows and columns corresponded to a higher quality factor, but Q began dipping after surpassing these values. Adding spacing to the sides of these rows and columns, in order to guarantee confinement of the holes along the sides of the devices, resulted in a slab size of 1400 $\times$ 896 $\times$ 100 nm. It is worth

noting here that the slab height of 100 nm is relatively arbitrary; initial simulations yielded a stronger, isolated resonant frequency peak with this thickness. It is also worth noticing that the air columns to the left and right of the missing defect region are smaller than the other columns. This is a design tactic meant to minimize leakage at the cavity boundaries. As mentioned in Section 2.1, tuning these two columns in particular smooths the spatial envelope and reduces out-of-plane radiation loss. Initial simulations yielded strong Q values when these columns were given radii of 20 nm.

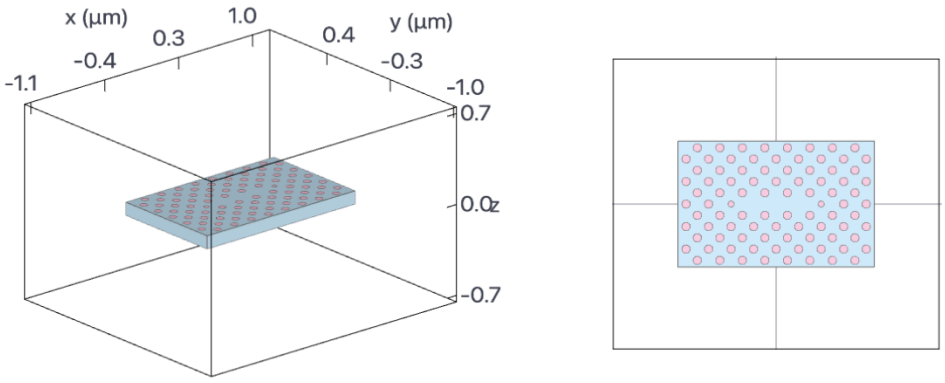

Figure 3.1: Left: Three dimensional view of the initial WS_2 L3 cavity design, of dimensions $1400 \times 896 \times 100$ nm. The lattice is triangular and the center has three missing holes. The spatial period $a = 80.295$ nm, and the radius of each column is $0.35a$. The surrounding medium is a vacuum, defined as a non-dispersive, isotropic material with relative permittivity of 1.0, allowing for light to travel freely in this region. Right: Top view of the cavity structure. The columns to the left and right of the missing defect region each have a radius of 20 nm.

The cavity is made from an anisotropic material called tungsten disulfide, WS_2. The refractive indices of WS_2 are the same in the x and y directions of the material, but differ in the z direction. The refractive index of WS_2 in the z direction is shown in Figure 3.2. To account for the anisotropic nature of WS_2, the AnisotropicMedium class offered by the Tidy3D API [23] was utilized:

```
ws2_medium = td.AnisotropicMedium(
    xx=medium_xy,
    yy=medium_xy,
    zz=medium_z
)
```

The medium in the x and y directions are denoted as `medium_xy` here, and the medium in the z direction is denoted as `medium_z`. It is worth noting that AnisotropicMedium only allows for diagonal anisotropy; the refractive index values are stored in a diagonal matrix reflecting different material properties in the x, y, and z directions without any coupling between directions. Note that medium for WS_2 is fixed in the x and y directions for a wavelength of 700 nm, which correlates to a refractive index value of $n \approx 4.357$. The z direction corresponded to a refractive index value of $n \approx 2.85$ [32].

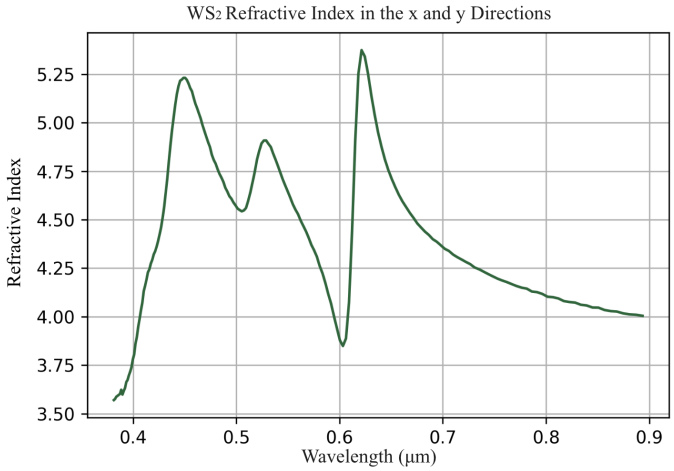

Figure 3.2: Refractive index in the x and y direction of WS_2.

A point dipole source was utilized with the time dependence described by a Gaussian pulse of frequency width of 3×10^{13} Hz and central frequency of 4.833×10^{14} Hz, encapsulating a resonant frequency located at 4.838205×10^{14} Hz [23]. This point dipole is polarized in the y-direction (E_y), as it matches the dominant field component of the cavity mode. A field monitor is centered around the region surrounding the three missing holes, of size 800 $\times$ 500 $\times$ 100 nm, to measure the energy storage over time. A secondary field monitor is implemented to measure the out-of-plane radiation loss. This monitor is two-dimensional, with dimensions 1400 $\times$ 896 nm, and is located 350 nm above the cavity along the axis. Finally, a dimensionless field time monitor was implemented to measure the field decay over time.

Running an initial forward simulation revealed that a sharp resonance peak occurred at 4.838×10^{14} Hz, as shown in Figure 3.3, corresponding to a quality factor of approximately 5058. Quality factor is calculated here using another Tidy3D class known as ResonanceFinder [33], which uses a time-series function to derive the frequency, decay rate, Q-factor, amplitude, and phase of a frequency plot.

3.2 Objective Function

For this study, initial simulations indicated the highest Q factor results when selecting the closest 58 holes to the cavity region as optimizable, meaning these cavities could be moved in the x and y directions while the other columns remained fixed during optimization. In a perfect cavity, typically only the two holes to the left and right of the missing defect region are adjusted. In real world simulations, however, cavities are imposed by physical limits such as the degree of precision obtainable during the manufacturing of nanophotonic devices.

When defining an objective function (or the function that returns a value we want to optimize based on our parameters, which are the positions of the 58 columns surrounding

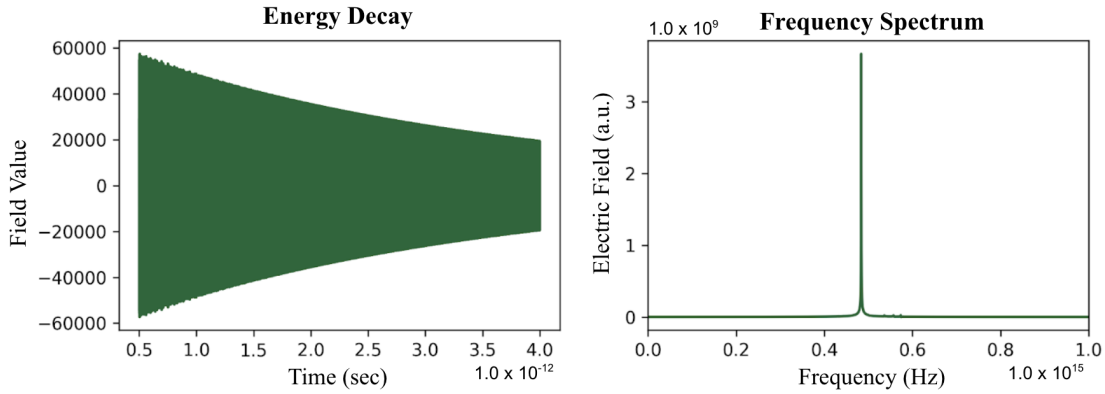

Figure 3.3: Left: Electromagnetic field decay of the unoptimized WS_2 L3 cavity, with respect to time. Right: Graph depicting where the resonant frequency is located, which has an electric field amplitude of approximately 1.4×10^{10} V/m. The graph plots the electric field magnitude in the y direction as it is the dominant field component.

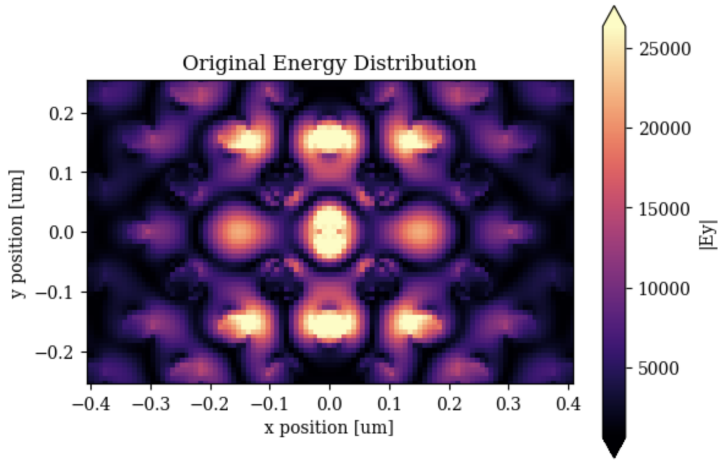

Figure 3.4: The initial distribution of the electric field around the missing defect region of the WS_2 L3 cavity. $|E_y|$ is The spatial profile of the cavity mode can be seen "spilling" into the surrounding PhC lattice region as it drifts further from the missing defect region.

the missing defect region), the value returned by the function must be both scalar and differentiable so that the gradient can be calculated. For this reason, implementing the Tidy3D class ResonanceFinder proved to be unusable because it produces an output that is not differentiable with respect to our parameters. Instead, by using the monitors responsible for measuring both the energy confinement in the missing defect region and the out-of-plane radiation loss, the objective function was defined as $Q = U/P$, as expressed in Equation 2.2. Using this function allowed for the stored energy, U, to be maximized, and the radiation loss, P, to be minimized, during the optimization process. The objective function is defined below as:

```python
#Objective function
def objective(params: anp.ndarray):
    sim = make_sim(params)
    sim_data = web.run(sim, task_name="l_cavity_adjoint", verbose=False)

    field_data = sim_data["field_monitor"]

    #Calculate U, stored electromagnetic field within cavity region
    Ex = anp.array(field_data.Ex.values)
    Ey = anp.array(field_data.Ey.values)
    Ez = anp.array(field_data.Ez.values)
    Hx = anp.array(field_data.Hx.values)
    Hy = anp.array(field_data.Hy.values)
    Hz = anp.array(field_data.Hz.values)
    Utot = anp.sum(
        anp.abs(Ex)**2 +
        anp.abs(Ey)**2 +
        anp.abs(Ez)**2 +
        anp.abs(Hx)**2 +
        anp.abs(Hy)**2 +
        anp.abs(Hz)**2
    )

    #Calculate P, power dissipation out of the cavity
    flux_field = sim_data["flux_monitor1"]
    Ptot = anp.sum(anp.abs(flux_field.flux.values))

    #Calculate Q
    print("U =", Utot)
    print("P =", Ptot)
    print("U/P =", Utot/Ptot)

    return -(np.log(Utot) - np.log(Ptot))
```

Figure 3.5: Function implemented for the objective function.

The variable sim_data contains all the data obtained by the simulation run, from which all E and H field components are collected. The total stored energy U_{tot} corresponds to the formula for the total electromagnetic energy stored within a volume,

$$U = \frac{1}{2}\int(\epsilon_0 E^2 + \mu_0 H^2) \tag{3.3}$$

The constant values here are disregarded. Even though the dipole is polarized in the E_y direction, E and H components in all directions may still be induced by the source point, which is why they are taken into account here. The power dissipation P_{tot} is simply collected

by summing all the energy flux values through the field monitor located above the cavity region. Finally, instead of calculating Q as U/P directly, we instead take the logarithm of this equation, resulting in

$$Q = -(\log(U_{tot}) - \log(P_{tot})) \tag{3.4}$$

This is done in order to avoid numerical instability; U_{tot} can sometimes be a very large number whereas P_{tot} can sometimes be very small. Taking the logarithm turns the ratio into a difference, preventing extreme values while yielding a smoother optimization curve. Note that the value returned is also negative; this is done because the optimizer used in this methodology naturally tries to minimize a function rather than maximize. Returning the negative of a value ensures the optimizer maximizes instead.

3.3 Optimizers

Parameter updates, or the displacement of air column positions surrounding the missing defect region, were performed using the Adam optimizer implemented in Optax [34]. Adam is a popular and efficient gradient based optimizer used for machine learning algorithms. The optimization ran for 10 iterations. For each iteration, a forward and adjoint simulation was run. The resulting gradient and objective function return value were calculated and the parameters were updated for each step. The Autograd value_and_grad function was utilized to extract the function return value and gradient.

To realize physically enforceable geometries, hole displacements were constrained to a maximum magnitude of one half of the lattice spacing, corresponding to 40.1475 nm. The learning rate, which determines the magnitude of each parameter update, was set to one tenth of the lattice spacing, which equals 8.0295. This parameter was selected empirically as it allowed for the maxima to be obtained without overshooting and risking lower values of Q.

Chapter 4

Results

As mentioned, a loop consisting of ten iterations of adjoint simulations was run in order to adjust the positioning of the air columns surrounding the missing defect region and maximize quality factor. Results indicated an increase in Q from approximately 5058 to 16906, as calculated by ResonanceFinder, corresponding to a 234% increase. The objective function convergence plots are shown below, as well as the field distributions and field decay plots for cavities before and after the optimization process:

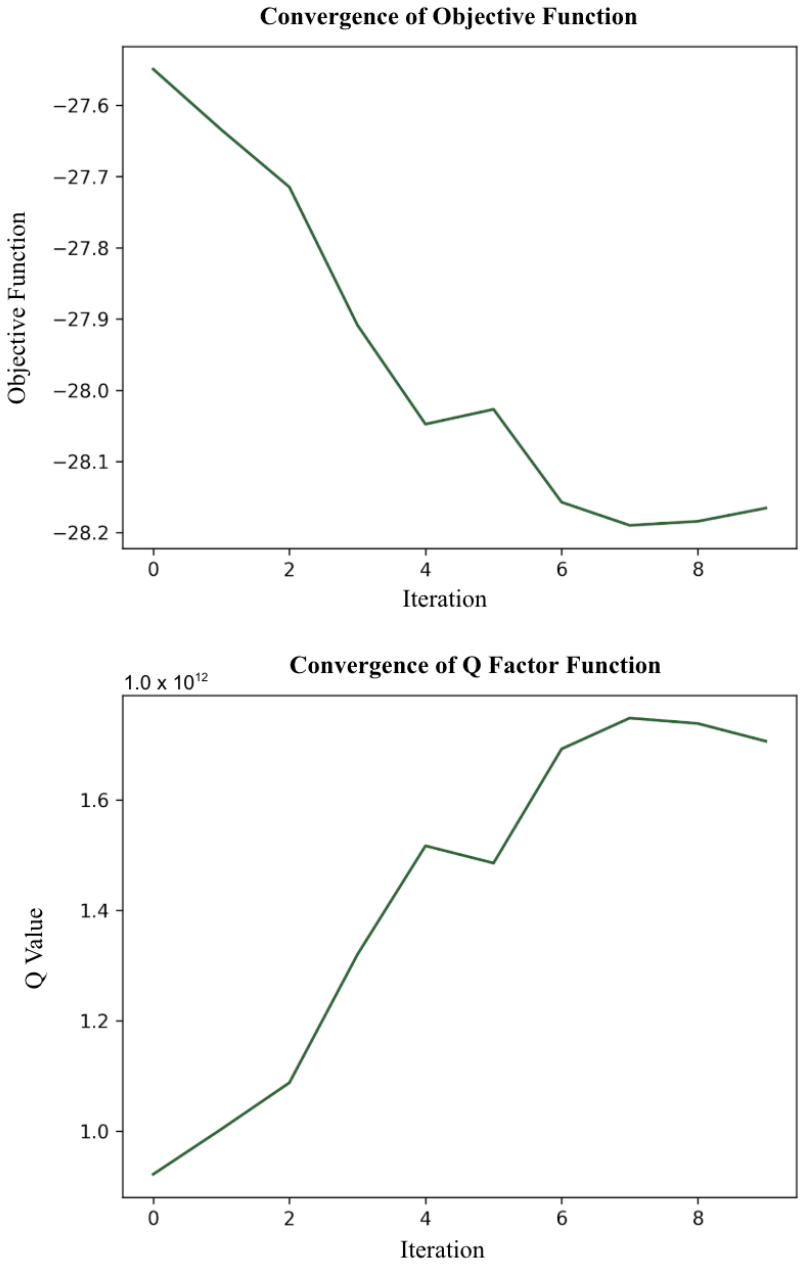

Figure 4.1: Top: Plot displaying the objective function return value versus optimization sequence iteration number. The Adam optimizer naturally performs minimization, so the objective function was defined as the negative of the physical performance metric. Minimizing the function therefore corresponds to maximizing the desired quantity, Q. Note here that the objective corresponds to a log function, so the values along the y axis only correspond to an order of magnitude of 10. Bottom: Plot displaying the Q factor value versus optimization sequence iteration number. This reflects the true value of Q that was maximized during the optimization process.

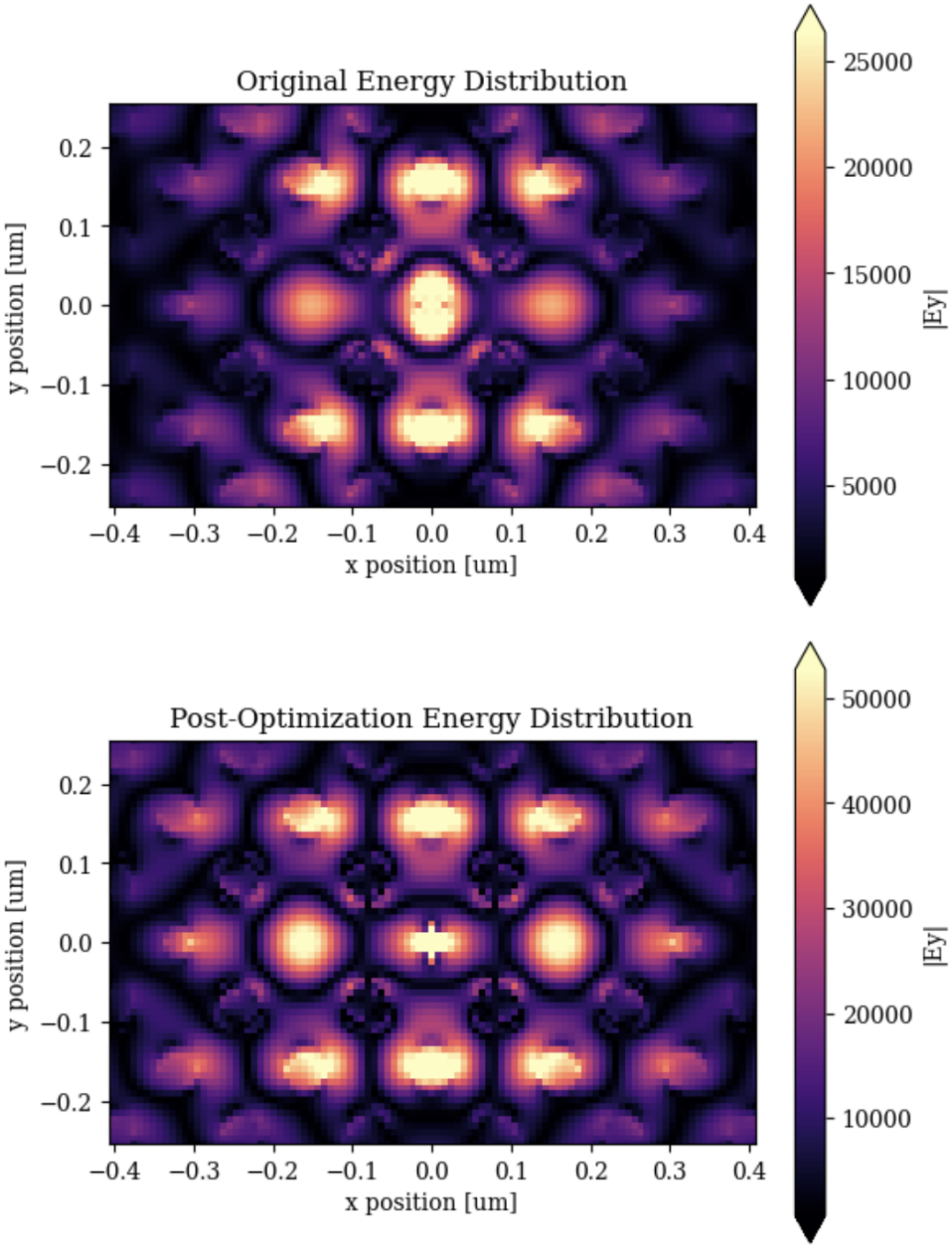

Figure 4.2: Top: Distribution of the magnitude of the electric field in the y direction across the cavity region, prior to optimization. Bottom: Field distribution after optimization. The post-optimization mode profile shows that the energy confinement increased overall, and also reflects a shift in the spatial pattern itself due to the adjustments in the column positions.

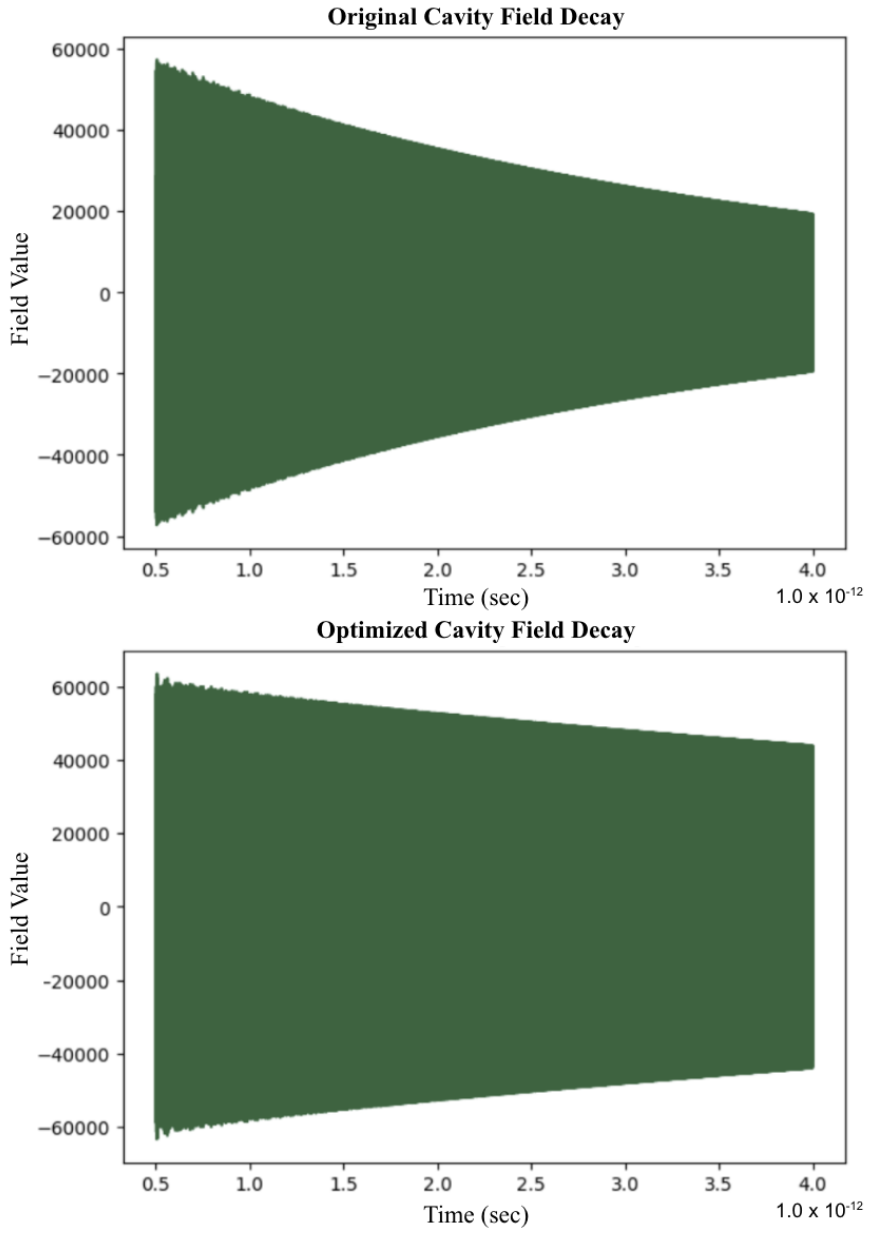

Figure 4.3: Top: Field decay of the unoptimized cavity over time. Bottom: Field decay of the optimized cavity over time. The decay is more gradual for the optimized plot than it is for the unoptimized, indicating stronger field confinement and less leakage over time.

Chapter 5

Conclusion

The 234% increase in quality factor indicates the significant difference anisotropic material response can have on line-defect cavity performance. In contrast, original simulations that did not account for the anisotropic medium of WS_2 (i.e. the refractive index being the same in all three directions) resulted in quality factors that would average near 5500 post-optimization, highlighting the importance of accounting for the permittivity tensor material properties in inverse design. These results suggest that anisotropic materials successfully provide additional degrees of freedom to tailor electromagnetic field confinement in line-defect cavities. Utilizing highly anisotropic materials in the design of cavities and waveguides can therefore prove to be invaluable due to their ability to confine electromagnetic waves in specific directions.

Ongoing work involves simulating and optimizing other highly anisotropic materials, primarily CrSBr. Moreover, the methodology discussed here indicates the promise of using the adjoint method to simulate and optimize anisotropic materials, but these results do not take into account the dispersive medium of materials such as WS_2. In this study, the medium was a fixed constant corresponding to the refractive index for a wavelength of 700 nm. Because resonance was located near approximately 620 nm, the refractive index at

this wavelength is not accurately reflected in the simulations, and the results here are not physically tangible. Incorporating the dispersive nature of WS_2 would mean that as the wavelength changes, the values of n change as well. Tidy3D's AnisotropicMedium class cannot hold non-constant values in its permittivity tensor, and ensuring such a simulation is computationally efficient is challenging in itself.

Bibliography

[1] Richard Sottie. "The History of Light and the Birth of Quantum Physics". In: *UNESCO* (2025).

[2] American Physical Society. *May 1801: Thomas Young and the Nature of Light*. 2008. URL: https://www.aps.org/archives/publications/apsnews/200805/physicshistory.cfm.

[3] Matthew Parry-Hill and Michael W. Davidson. *Thomas Young's Double Slit Experiment*. https://evidentscientific.com/en/microscoperesource/tutorials/doubleslitwavefronts.

[4] Karl Giberson. "Einstein Describes the Photoelectric Effect". In: *EBSCO* (2023).

[5] Akira Tonomura. "The double-slit experiment". In: *Hiden Analytical* (2002).

[6] Charles H. Townes. "The first laser". In: *The University of Chicago Press* (2003).

[7] National Nanotechnology Coordination Office. *About Nanotechnology*. https://www.nano.gov/about-nanotechnology/. 2026.

[8] Yuan-Fong Chou Chau. "Nanophotonic Materials and Devices: Recent Advances and Emerging Applications". In: *MDPI* (2025).

[9] Paul V. Braun, Stephanie A. Rinne, and Florencio García-Santamaría. "Introducing Defects in 3D Photonic Crystals: State of the Art". In: *Advanced Materials* (2006).

[10] Rajesh V. Nair and R. Vijaya. "Photonic crystal sensors: An overview". In: *ScienceDirect* (2010).

[11] Marco Saldutti et al. "Modal Properties of Photonic Crystal Cavities and Applications to Lasers". In: *MDPI* (2021).

[12] W Heiss et al. "Epitaxial Bragg mirrors for the mid-infrared and their applications". In: *ScienceDirect* (2002).

[13] OpenStax. "Total Internal Reflection". In: *LibreTexts* (2022).

[14] Anqi Ma et al. "Anisotropic quantum transport in a programmable photonic topological insulator". In: *Optica* (2024).

[15] Aniket Patra et al. "Design and Polarization Control of the Modal Splitting in Hybrid Anisotropic Nanocavities". In: *Wiley Advanced* (2023).

[16] R. Paschotta. *Optical Switches*. 2025. URL: https://www.rp-photonics.com/optical_switches.html.

[17] R. Paschotta. *Mode Division Multiplexing*. 2019. URL: https://www.rp-photonics.com/mode_division_multiplexing.html.

[18] Yosep Shin et al. "Anisotropic metamaterials for scalable photonic integrated circuits: a review on subwavelength gratings for high-density integration". In: *National Library of Medicine* (2025).

[19] Dakota E. McCoy et al. "Finite-difference Time-domain (FDTD) Optical Simulations: A Primer for the Life Sciences and Bio-Inspired Engineering". In: *ScienceDirect* (2021).

[20] Tyler Hughes, Zongfu Yu, and Shanhui Fan. *Introduction to FDTD Simulation*. 2022. URL: https://www.flexcompute.com/fdtd101/Lecture-1-Introduction-to-FDTD-Simulation/.

[21] Dr. Rüdiger Paschotta. *Near Field and Far Field*. URL: https://www.rp-photonics.com/near_field_and_far_field.html.

[22] William H. Weedon and Carey M. Rappaport. "A General Method for FDTD Modeling of Wave Propagation in Arbitrary Frequency-Dispersive Media". In: *IEEE* (1997).

[23] Flexcompute. *Tidy3D API*. https://docs.flexcompute.com/projects/tidy3d/en/latest/api/index.html.

[24] ScienceDirect Topics. *Inhomogeneous Equation*. URL: https://www.sciencedirect.com/topics/mathematics/inhomogeneous-equation.

[25] George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists*. Sixth. Elsevier / Academic Press, 2005. URL: https://msashigri.wordpress.com/wp-content/uploads/2016/11/methods-of-mathemacial-for-physicists.pdf.

[26] Lukas Novotny and Bert Hecht. *Principles of Nano-Optics*. 2nd. Cambridge University Press, 2012. URL: https://einstein.nju.edu.cn/upload/kindeditor/file/20210703/20210703140803_3676.pdf.

[27] Tyler Hughes, Zongfu Yu, and Shanhui Fan. "Inverse Design in Photonics Lecture 2: Adjoint Method". In: *FlexCompute* (2024).

[28] Chemistry LibreTexts. *Bragg's Law*. URL: %5Curl%7Bhttps://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_(Analytical_Chemistry)/Instrumentation_and_Analysis/Diffraction_Scattering_Techniques/Bragg's_Law%7D.

[29] Roger Nix. *Miller Indices (hkl)*. https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Surface_Science_(Nix)/01\%3A_Structure_of_Solid_Surfaces/1.02\%3A_Miller_Indices_(hkl).

[30] Anthony Kelly and Kevin M. Knowles. *Appendix 3*. 2012.

[31] Hyun-Joo Chang et al. "A photonic-crystal optical antenna for extremely large local-field enhancement". In: *Optics Express* (2010).

[32] Panaiot G. Zotev et al. “Nanophotonics with multilayer van der Waals materials”. In: *Nature Photonics* (2025).

[33] *Extracting resonance information using Resonance Finder*. `https://www.flexcompute.com/tidy3d/examples/notebooks/ResonanceFinder/`.

[34] *Optax Documentation*. `https://optax.readthedocs.io/en/latest/`.