Computational and Graphical Methods in Nano-Photonic Device Design

By

Raymond Anthony Wambold

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Electrical and Computer Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2021

Date of final oral examination: 5/7/2021

The dissertation is approved by the following members of the Final Oral Committee:

Mikhail A. Kats, Associate Professor, Electrical and Computer Engineering
Zongfu Yu, Associate Professor, Electrical and Computer Engineering
Mark Eriksson, Professor, Physics
Jennifer Choy, Assistant Professor, Engineering Physics

To my family

Carolyn, your loving and unwavering support has sustained me through innumerable challenges

Keira Grace, your smile, laughter, and love of life brings me immense joy

Rowan Anthony, your perseverance is an inspiration (and sometimes a frustration) to me

You are all my favorites

Abstract

Computational and Graphical Methods in Nano-Photonic Device Design

This thesis is primarily concerned with using computational methods to design novel photonic components and use these results to improve our intuition of device design. In this work, we seek to leverage optimization techniques, optical simulations, and an understanding of electromagnetic theory to demonstrate the broad range of capabilities we now have in designing novel photonic devices as well as developing new optical platforms.

Chapter 2 is focused on using doped semiconductors as a novel platform for optical components in the mid-IR. We design and demonstrate Fresnel zone plates and frequency selective surfaces embedded into silicon by selective doping. We explore using diffusion-doped indium arsenide to create gradient-index profiles within the semiconductors to redirect the flow of light. Due to the complexity of design, we implemented a multi-physics optimization routine to design a lens and beam deflector.

In Chapter 3 we describe the use of a powerful design technique known as adjoint optimization to design structures that enhance the emission from nitrogen vacancy (NV) color centers in diamond. Here, we show how we used adjoint-optimization to create a nanoscale light extractor that can extract up to 35× the amount of light from a single NV compared to NVs below un-patterned diamond. Then we present an array of devices designed for NVs in (111) crystalline diamond that showcase a wider range of capabilities enabled by adjoint optimization. We present

an extractor that selectively deflects emission based on NV orientation and an NLE that minimizes the intensity of the excitation laser by up to 70% to protect photosensitive samples.

Finally, in Chapter 4, we focus on using graphical techniques alongside full-wave simulations to understand the behavior of reflect-array metasurfaces. We show how these metasurfaces can achieve a full 2π scattered field phase coverage with the addition of the reflective back plane. A characteristic feature of these metasurfaces is that they can become perfect absorbers and, around this point, the scattered field phase coverage is greatly reduced. We show how this phenomenon occurs and what can be done to engineer this specific response.

Contents

List of Figures

Acknowledgements

One of my favorite quotes is by Isaac Newton in a letter to Robert Hooke “If I have seen further, it is by standing on the shoulders of Giants”. The saying references a metaphor of dwarves standing on the shoulders of giants which implies our current knowledge and discoveries are only made possible due to the great work of those who have come before. This saying is true for me in a broad sense, in that this thesis builds upon the foundation of knowledge put forward by many great scientists of both large and small renown, but I also hold this to be true on a personal level: I am where I am today due to the generosity, sacrifice, and support of countless others in my life.

I would first like to thank my advisor Mikhail. Firstly, I am grateful that he took me on as a student when I was looking to finish completing my PhD after Penn State and gave me the opportunity to be a part of a “real” grad program. Throughout the course of my tenure at UW, I learned so much from Mikhail about how to think about physics, the scientific process, and general insights on life. I am also extremely grateful for the freedom that he gave me to pursue the science I was interested in while at the same time guiding me and offering valuable insights (such as ctrl+f for “device”). Also, I’m sorry for maybe breaking your rib during basketball.

I would like to secondly thank Dr. Z (Dr. Darin Zimmerman), my previous advisor at Penn State. Without his support early in my academic journey, I would not be where I am today. My time in Dr. Z’s lab was truly one of the best seasons in my life thus far. Along with Darin, I need to thank Dr. Gary Weisel for his influence and mentorship during my time at Penn State. I always enjoyed our philosophical conversations about the world over some Great China (or cigars).

This work would also not be where it is without the collaboration we had with the RAISE-TAQS grant with Bob Hamers, Mark Eriksson, Jennifer Choy, and Shimon Kolkowitz.

Specifically, I am very grateful for conversations with Jen and Shimon who taught me about NV centers and told me what devices would be useful for experiments. I also want to acknowledge the students who were a part of this project, primarily Ben Bachman and Gabriel Jaffe for their insights and contributions. This collaboration and working with these talented individuals were truly a highlight of my time at UW. I also thank Zongfu Yu for being a part of my committee along with Mikhail, Mark and Jen and forcing me to really understand adjoint optimization.

And many thanks go out to my colleagues and friends in the Kats Group at UW. Our cadre of students gave me the support and community that I never had before and was something truly special. I would like to specifically thank Jad Salman for his friendship and mentorship when I first started. Jad is one of the most selfless people I know and is always willing to teach or talk through any issue or just shoot the breeze. I also want to thank the "first-wave" of students in Mikhail's group: Chenghao, April, Patrick, Alireza, and Brad. It was a great experience growing into productive scientists with you all. Yuzhe is also a part of that group but as a mentor as he came in as a post-doc. My appreciation is also extended to Brian, May, Demeng, Purva and Jon; the "second-wave" who are now becoming senior members in the group (time-flies!). Acknowledgement must also be made to the shorter-term members of our group with whom I worked: Michel, Anthony, Raphael, Graham and Manuel.

My final thanks go out to those in my personal life who supported me through this journey. I thank my parents who didn't freak out too much when we said we were moving to Wisconsin. Their selfless love and support through my whole life is the foundation for where I am today. I also want to thank my in-laws for their encouragement and love during this season of our lives. I also would not have made it without the support of my friends back home: Will, Luke, and Eric,

and my new friends here in Madison: Michael, Ben, Karl, Jake, and many others. Sometimes you just need to talk about something other than science.

Lastly, but most significantly, I want to thank my wife, Carolyn. She was willing to move 750 miles from home so that I could pursue this opportunity of my getting my Ph.D. Her faithfulness and steadfastness got me through many hard seasons here and her consistent friendship gave me the strength to push through each day. Words cannot do justice to her support; she is the real hero of this story.

1 Introduction

Light is a fundamental part of how we, as humans, interact with and understand the world. For most people, we carry two remarkable optical detectors (admittedly with varying efficiencies) with us wherever we go that allow us to experience the beauty of the world around us. Beyond our biological capabilities, harnessing electromagnetic waves and engineering devices that generate, manipulate, and/or detect them have led to great technological progress, particularly in the past 100 years. Common household microwaves are the result of progress in the development of RADAR technology during World War II, television screens have evolved significantly from (now ancient) cathode ray tubes to liquid crystal displays enabled by light emitting diodes, and, of significant mention, the great strides made in medical diagnosis and procedures due to the development of such electromagnetic-based technologies as magnetic resonance imaging (MRI), computed tomography (CT) scanning, or the precision removal of destructive tissue through laser therapies.

Our capacity to exploit light for effective engineering applications is not particularly limited by our understanding of how light behaves, but by the materials and technologies we can employ to control, detect, and direct it. In a classical and quantum sense, our understanding of light is quite mature. However, our design capabilities have largely relied on intuition born from that knowledge. This had led to great progress in optical engineering yet our design space is limited to what our minds can contrive, which, while generally impressive, does not always reach the global optimum. To move photonic engineering and design to the next level, we must marry our intuition based on electromagnetic theory with the advantages of computational optimization and the

freedom from constraints that this brings. This thesis is primarily concerned with using said computational methods to design novel photonic devices thus improve our intuition of device design. Our improved intuition can then become a feedback loop to improve our computational design methods. We seek to leverage optimization techniques, optical simulations, and an understanding of electromagnetic theory to demonstrate the broad range of capabilities we now have in designing novel photonic devices as well as developing new optical platforms.

Chapter 2 of this thesis is focused on using doped semiconductors as a novel platform for optical components in the mid-IR. This section starts by presenting some of the work led by Jad Salman in Mikhail Kats's group in creating plasmonic and diffractive optics using selectively doped silicon. Using selectively doped semiconductors, we design and implement a Fresnel zone plate as an example of a diffractive device and a frequency selective surface as an example of the plasmonic capabilities of this material platform. We show the fabrication, implantation and measurement methods used to create and characterize these devices.

With the plasmonic and diffractive capabilities established, we explore using diffusion-doped indium arsenide to create gradient-index profiles within the semiconductors to redirect the flow of light. Here, we moderately dope the semiconductor such that the plasma wavelength is longer than our desired operating wavelength. This gives us control over the index of refraction without the high losses that are present when operating beyond the plasma wavelength, in the plasmonic regime. These devices are designed to be fabricated using localized spin-on glass dopant sources which are thermally diffused into the semiconductor substrate.

Due to the complexity of design parameters, we implemented a multi-physics optimization routine comprised of a diffusion simulation paired with a full-wave optical solver embedded within a global optimization algorithm to design a lens and beam deflector using this doped-

semiconductor platform. We attempted fabrication using tin-doped indium-arsenide, but, unfortunately were not able to overcome a few challenges specific to our material system (namely outgassing of arsenic from the substrate) with the fabrication tools at our disposal. Yet this platform shows promise, and we leave room for investigating alternate material systems (such as using alternate semiconductor that does not outgas as easily) using the computation design methods we present.

In Chapter 3 I describe the use of a new and powerful design technique for photonics known as adjoint optimization to design structures that enhance the light output from nitrogen vacancy (NV) color centers in diamond. NV centers in diamond form stable quantum systems up to room temperature and are thus useful for quantum computing applications as well as sensing local electric and magnetic fields, and even local temperature and strain. A key limitation of sensing based on NV centers is the high refractive index of diamond, which makes it difficult to extract a significant fraction of light emitted by the NV out of the diamond substrate. In this section we show how we used adjoint-optimization to design a nanoscale light extractor that can extract up to 35× the amount of light from a single NV compared NVs below an un-patterned diamond surface. We go in depth about the details of the adjoint method and explain some of the parameters and techniques we used to ensure a fabricable devices and designs robust to imperfections.

Then we present an array of devices designed for NVs in (111) crystalline diamond that showcase a wider range of capabilities enabled by the adjoint optimization method. We present an extractor that selectively deflects the emission based on NV orientation as well as an NLE that minimizes the intensity of the excitation laser by up to 70% to protect photosensitive samples from the pump laser. The plethora of adjoint figures-of-merit used here are very useful for designing a

variety of photonic structures and can be easily extended to other photonic applications *(e.g.* waveguides, metasurfaces, beam deflectors, etc..).

Finally, in Chapter 4, we focus on using graphical techniques alongside full-wave simulations to understand the behavior of a specific type of metasurface called a reflect-array metasurface. This metasurface is comprised of metallic nanorods (or nanostrips in 2D) sitting on a dielectric spacer above a metallic ground plane. We show and explain how these metasurfaces can achieve a full 2π scattered field phase coverage with the addition of the reflective back plane. One of the characteristic features of these metasurfaces is that, with specific geometric dimensions, they become perfect absorbers, where they reflect almost no light. Around this point, the scattered field phase coverage is greatly reduced. We show how this phenomenon occurs and what can be done to engineer this specific response.

2 Monolithic doped-semiconductor platform IR photonic devices

2.1 Introduction

Manipulating optical wavefronts has been accomplished using a variety of approaches throughout the development of the field of optics. Optical technology has evolved and progressed significantly from early lenses and magnifiers made from rock crystals and glass[1], [2] to modern optical metasurfaces, optical fibers and gratings. Here we present a novel technology that uses selectively doped semiconductors to engineering their optical properties to create photonic devices. We primarily focus on engineering optical devices in the mid-infrared regime of the electromagnetic spectrum due to the position of the plasma frequency of the doped semiconductors that we are using. To overcome challenging design constraints, we implemented a multi-physics optimization algorithm that calculates the optical properties of a semiconductor slab from the diffusion profile of a given dopant source pattern. The optimization algorithm is responsible for finding the optimum dopant pattern for a desired output field (e.g., creating a lens).

2.2 Theory of semiconductor refractive-index engineering

How a given material interacts with an electromagnetic wave is derived from the atomic composition and structure of the material. In different frequency ranges, different aspects of the material composition wax and wane in their importance in the interaction mechanisms with

light[3]. For metals, the permittivity can be described using a frequency-dependent model based on how free electrons interact with light[4].

Semiconductors are an interesting in-between case, as their name implies. Intrinsically these materials have a band of forbidden electron states between the electronic valence band and the conduction band[5], [6]. That gives these materials unique electrical and optical properties. Primarily, the properties of interest arise when these materials are doped with small concentrations of electronic donors or acceptors which fills the electronic states in the semiconductor which in turn modifies its optical properties.

The Drude model gives a good approximation for the permittivity of metals and semiconductors in the infrared[7]. The frequency-dependent permittivity is primarily governed by the concentration of free electrons and interband transitions that are available for certain energy thresholds.

If we assume we are operative at a frequency regime away from interband transitions, the complex permittivity given by the Drude model is given as follows[7]:

$$\epsilon_{Drude}(\omega) = \epsilon_{\infty} - \frac{\omega_p^2}{\omega^2 + i\Gamma\omega} \tag{2.1}$$

where ϵ_{drude} is the Drude permittivity for a given frequency, ω. Here, ϵ_{∞} is the high-frequency permittivity, Γ is a damping term proportional to the Fermi velocity and inversely proportional to the electron mean free path. The plasma frequency, ω_p is the natural resonance frequency of the free electrons in the material and is given by $\omega_p = \sqrt{ne^2/m_e\epsilon_0}$. In this equation n, represents the density of electrons in the plasma, while e is the elementary charge, m_e is the mass an electron and ϵ_0 is the free-space permittivity.

In rough terms, the plasma frequency is the frequency at which the material transitions between interacting with light as a dielectric to interacting with light more like a metal[7]. Namely, it is the resonance frequency of the free-electron sea in the material. At lower frequencies, the electrons cancel out the incident electromagnetic wave within a very small distance whereas at higher frequencies the wave is able to propagate through the plasma. In more specific terms, this frequency is the frequency at which the real part of permittivity approaches 0 and therefore, at higher frequencies the refractive index becomes imaginary, and the wave equation solution becomes an evanescent wave.

From this equation, it can be seen that the frequency-dependent permittivity can be modified through the altering of the plasma frequency. For example, an increase in the concentration of electrons, n, gives rise to a higher plasma frequency which in turn reduces the permittivity. For most metals (*e.g.*, gold, silver, aluminum), the plasma frequency resides around the UV to visible region of the electromagnetic spectrum. For intrinsic silicon, the plasma frequency is well into the terahertz region of the spectrum with $f_p = 1.6\,THz$ corresponding to a plasma wavelength of $\lambda_p = 11.8\,mm$. However, with sufficient dopant concentrations, the plasma wavelength can be shifted well into the mid-IR spectrum[8]. This control over the plasma frequency via free-carrier concentration opens the door to using semiconductors as plasmonic device platform in the mid to far IR.

To calculate the resulting index of refraction (through the permittivity) using equation 2.1 for semiconductors, we must calculate the change in scattering time with increasing free-carrier concentration. The scattering time is inversely proportional to the mobility while the mobility is inversely proportional to the carrier concentration. This is given by the following equations[6], [9]:

$$\Gamma = \frac{\omega}{\tau} \tag{2.2}$$

$$\tau = \mu \cdot \frac{m_{eff}}{q} \tag{2.3}$$

$$\mu = \frac{1}{q \cdot N_d \cdot \rho} \tag{2.4}$$

where τ is the electron scattering time, μ is the electron mobility and ρ is the empirically measured resistivity which is also a function of carrier concentration.

Materials such as conducting nitrides and oxides[10]–[13] and traditional group IV[8], [14]–[23] and III–V [24]–[27] semiconductors can have doping-tunable plasma wavelengths spanning the infrared spectrum and have become promising for plasmonics applications. Some of these materials, e.g., titanium nitride, have the added benefit of high-temperature resistance and CMOS compatibility[10]. Prior work using these materials mainly involves uniformly doped substrates [19], films [11], [12], [27], [28], or nanocrystals [29], [30], and structuring is usually achieved through selective etching [15], [27].

Figure 2.1 shows the calculated index of refraction of silicon as a function of wavelength in the mid IR for varying dopant concentrations using equations 2.1-2.4. As the dopant concentration is increased, the plasma wavelength blue-shifts through the spectrum (here the plasma wavelength can be visualized roughly as the point where $n = \kappa$) . With extremely high dopant concentrations, the plasma wavelength is pushed further toward the near-IR thus showing the versatility of this approach to create metal-like doped semiconductors over a large swath of the mid-far IR.

One thing to note is how much the refractive index shifts as a function of concentration at a given wavelength. For example, if a slice at $\lambda = 10.6\ \mu m$ is taken, one can see the broad sweep

in index values as the concentration is changed. Thus, not only is there control over where the plasma frequency resides, but also one can have control of the value of the refractive index at a particular wavelength with one knob, the free-electron concentration.

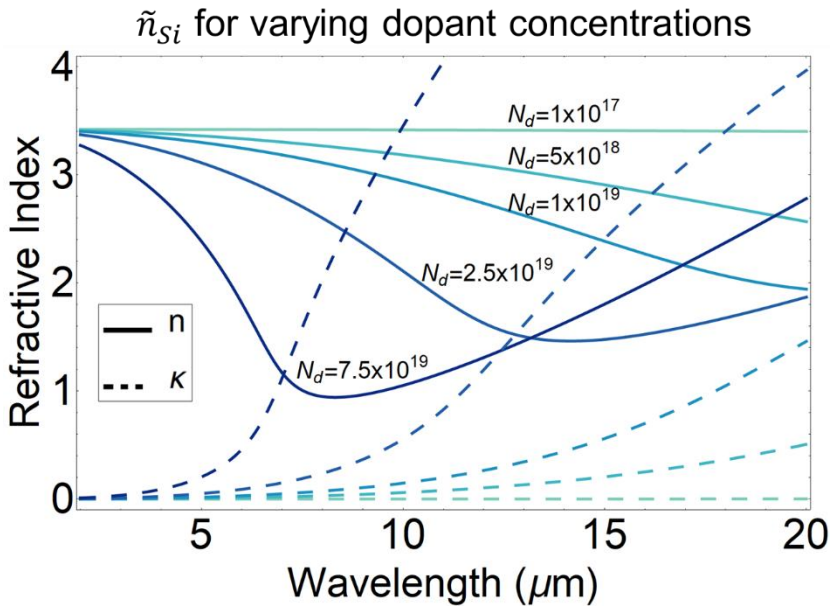

Figure 2.1 - Calculated refractive index of Si for varying free-carrier concentrations through the mid IR spectrum. The real part of the index is denoted by the solid curves while the imaginary part is denoted by the dashed curves.

If we take the plasma frequency as a type of dividing line, we can split the spectrum into the plasmonic regime and the gradient-index regime. We can leverage the same platform to create diverse devices operating in each regime. In the plasmonic regime, we can target wavelengths longer than the plasma frequency where we can have metal-like behavior of the doped elements and thus can create plasmonic-resonance-based devices, while in the gradient-index regime there is much flexibility for leveraging dopant concentration to engineer the index of refraction.

The potential advantages of optics based on doped-semiconductors are: 1) the ability to withstand harsh physical, chemical and thermal environments and 2) the ability to create flat devices on a chip. These advantages are made possible by the fact that the devices can be fully embedded inside a semiconductor wafer, making them resilient and enabling efficient heat

dissipation [31], [32]. GRIN devices enabled by this technology are potentially more versatile compared to conventional refractive optics (e.g., they could combine the functionality of several existing refractive components such as lenses and prisms in one element). Both GRIN and plasmonic devices based on this approach share the advantage that they are flat, and thus can be made to fit in compact systems. Our aim is to introduce this versatile doped-semiconductor platform for use in creating infrared optical devices.

The next section, Chapter 2.2, will touch on the work that we have done to create plasmonic devices from doped Si while Chapter 2.3 will touch on our theoretical work of engineering and optimizing dielectric devices through controlling the refractive index through doping.

2.3 Diffractive and plasmonic doped semiconductor devices

The work presented in this section was performed alongside our collaborators in Prof. Carsten Roning's Group at Friedrich Schiller University Jena in Germany. Also, the device design, characterization and experimental results were led by Jad Salman from Mikhail Kats's research group. Most figures and much of the text in this section is adapted from our publication: J. Salman et al., "Flat Optical and Plasmonic Devices Using Area-Selective Ion-Beam Doping of Silicon," Advanced Optical Materials, vol. 6, no. 5, p. 1701027, Mar. 2018, doi: 10.1002/adom.201701027.

2.3.1 Introduction

Commonly, noble metals (e.g., gold or silver) are used for plasmonic applications but they come with a few drawbacks, particularly in the mid to far IR[7], [33]–[35]. First, the plasma

frequency of noble metals lies in the visible spectrum thus causing them to behave more closely like a perfect conductor in the IR spectrum. This means they cannot sustain the necessary volumetric charge-density oscillations for plasmonic resonances. Further, metal-based devices are inherently not monolithic[32], [36]. They are typically layered structures where the metal resonators are placed on a dielectric substrate. This can leave them susceptible to physical and thermal damage while operating in non-laboratory environments. Finally, the plasma frequency and index of refraction is not tunable in the IR for noble metals and therefore, a degree of design freedom is lost.

In this study, we demonstrated the potential of doping-patterned silicon as a CMOS-compatible platform for mid- and far-infrared optics and plasmonics. We utilized ion-beam doping in conjunction with lithography to locally tune the optical properties of silicon to create embedded diffractive optical elements (e.g., Fresnel zone plates) and plasmonic components (e.g., frequency selective surfaces) that operate in the mid to far-IR regime. The resulting devices are monolithic, flat, resilient to thermal and physical damage, and can be easily integrated into other silicon-based platforms. A schematic showing a representative implantation process for the FSSs can be seen in Figure 2.2.

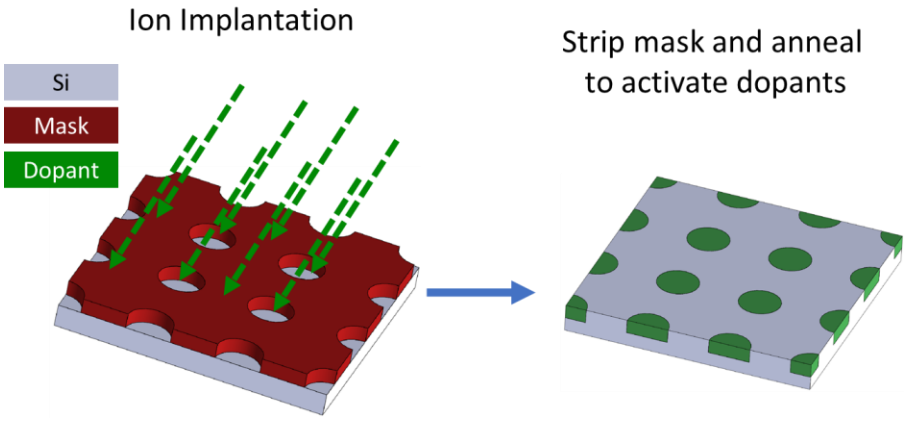

Figure 2.2 – Schematic representing the implantation process for creating doped-Si based FSSs. The photolithographically designed mask defines the implantation regions such that after stripping and annealing, only the unmasked areas are doped (represented by the green areas).

Previous studies that characterized the optical properties of ion-beam-doped silicon have used single, low-concentration or few, low-energy implants [8], [18]. Here we utilized multiple ion implant energies in the hundreds of keV range to create uniformly doped layers, improving the reliability of optical characterization.

High resistivity, single-side-polished silicon wafers (100 - 150 Ω-cm, N = 3×10^{13} cm^{-3}) were implanted with phosphorus (P) at a 7° tilt to create n^+-doped layers. The tilt angle was chosen to prevent channeling [27]. Based on calculations, a roughly uniform doping profile extending ~550 nm into the substrate was achieved by sequentially implanting each sample with five different ion energies (30, 70, 120, 220, and 380 keV), as shown in Figure 2.3b.

After ~550 nm, the doping concentration quickly decreases. The required ion fluences were calculated using Monte Carlo simulations (Stopping and Range of Ions in Matter, SRIM[39]). Four samples with peak doping concentrations ranging from 10^{18} to 10^{21} cm^{-3} were fabricated. After

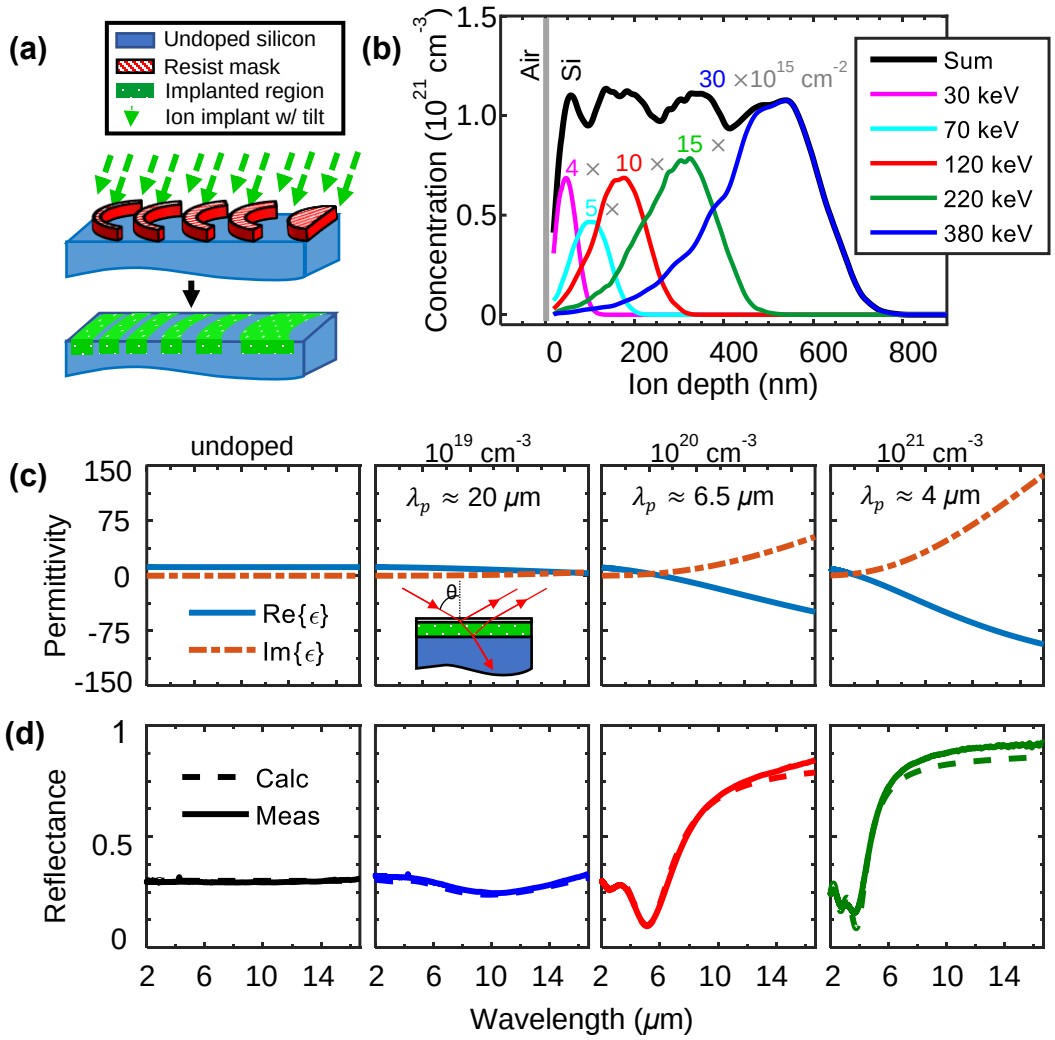

Figure 2.3 - a) Area-selective ion implantation using lithographically defined implant masks. b) Simulated as-implanted doping profile using five irradiation energies for the highest-doped sample in this work, with corresponding ion fluence labeled in the plot. c) The permittivities of undoped and highly doped silicon samples measured using spectroscopic ellipsometry. d) The measured (solid) and calculated (dashed) normal-incidence reflectance of three doped samples compared to undoped silicon.

implantation, the samples were annealed at 900 °C for 15 minutes under high vacuum (< 4×10^{-4}

mbar). Glow-discharge optical emission spectroscopy (GDOES) showed the formation of a uniformly doped layer that extends slightly deeper than the calculated thickness due to diffusion.

With increasing doping concentration, the crossover point ($\mathrm{Re}(\epsilon) = 0$) shifts towards shorter wavelengths, and reaches ~4.0 μm for the highest-doped sample, slightly shorter than previously reported values [18], [35]. From Drude-model fitting, we obtained a maximum activated doping concentration of over 2.1×10^{20} cm^{-3}, with an electron mobility greater than 50 $cm^2\ V^{-1}\ s^{-1}$, which is reasonable for such large concentrations[37]. We note the fraction of activated dopants decreased significantly at the highest doping concentration as the solid solubility limit was approached (~1×10^{21} cm^{-3} phosphorus in silicon[9]). To validate the extracted optical properties, reflectance measurements (Figure 2.3d) were performed at near-normal incidence using a Bruker V70 Fourier transform infrared (FTIR) spectrometer coupled to a Hyperion 2000 IR microscope, with an objective with numerical aperture (NA) = 0.4. The long-wave IR reflectance of the highest-doped sample is greater than 80% and is in good agreement with calculations based on ellipsometry data.

We used area-selective doping via photolithography and ion implantation to realize single-material, flat, embedded Fresnel zone plates (FZPs) to focus mid-IR light. In general, FZPs comprise alternating rings of transparent and opaque zones, and are often made using patterned metals[38]. At high doping densities, silicon is nearly opaque to mid-IR wavelengths, in contrast with intrinsic silicon which is broadly transparent; therefore, our selective doping process can be used to monolithically embed FZPs into a silicon slab.

2.3.2 Embedded mid-infrared Fresnel zone plates

We designed and fabricated three embedded 9-mm-diameter FZPs to focus 10.6-μm light at focal lengths of 10 cm, 6 cm, and 2 cm, with respective numerical apertures (NA) of 0.045, 0.075, and 0.23 (Figure 2.4b-i). A two-dimensional finite-difference time-domain (FDTD, using Lumerical FDTD Solutions[39]) simulation of the 2-cm-focal length FZP using the permittivity values extracted previously (Figure 2.4c) confirms the focusing properties of our design (Figure 2.4b). The implanted zones are nearly invisible to the naked eye and become barely visible under an optical microscope (Figure 2.4c). This demonstrates that the surface remains mostly flat after implantation and the optical properties of the doped regions in the visible do not differ substantially from those of undoped silicon [40].

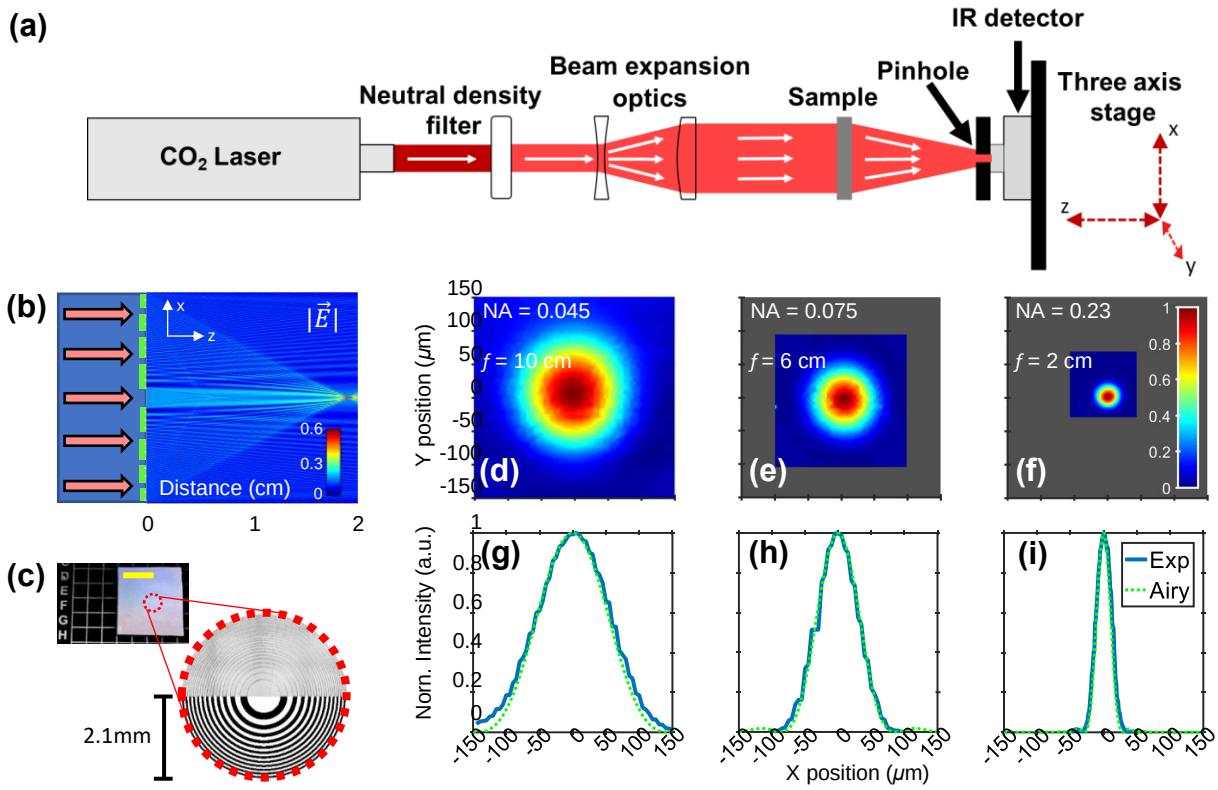

Figure 2.4 - a) Beam-profiling setup. From left to right: CO_2 laser source, neutral-density filter, beam expander, the FZP with the implanted side facing to the right, and a 30 μm pinhole in front of an MCT detector on a computer-controlled three-axis stage. b) Schematic of a Fresnel zone plate (FZP) made of ion-doped regions and a 2D FDTD simulation of the normalized field intensity in air at a wavelength of 10.6 μm, given normal-incidence light from the substrate. c) Photograph of a fabricated FZP (yellow scale bar = 6 mm). The magnified portion is a composite of the zone-plate design (bottom) and an optical image of the zone-plate surface (top). The dark rings are the doped regions. d-f) 2D plots of the measured intensity with respect to the lateral stage position at the focal plane for three FZPs, each plot normalized to 1. The numerical apertures (NA) and focal lengths (f) are noted at the top of each plot. g-i) Cross section of the intensity profile (solid blue) and diffraction-limited Airy disk calculation (dashed green) for each FZP.

To measure the focusing ability of the FZPs, we built a beam profiler consisting of a carbon dioxide (CO_2) laser (1 W continuous-wave, Access Laser, λ = 10.6 μm, Δf = 100 kHz), a beam expander, a sample holder for the FZPs, and a mercury-cadmium-telluride (MCT) detector (Figure 2.4). The detector was mounted on an automated three-axis stage with <50 nm positioning resolution, and the detection area was restricted using a 30-μm pinhole in front of the detector. The non-zero diameter of the pinhole has little impact on the fidelity of the reconstructed intensity profiles. The expanded beam had a diameter of 14 mm to completely encompass the 9-mm zone

plates, thus approximating an incident plane wave. Three-dimensional intensity profiles acquired by raster scanning were used to find the focal planes of the FZPs, which were in excellent agreement with the designs. The measured intensity plots, normalized to the peak value for each FZP at the focal plane, are shown in Figure 2.4d-f. Near-diffraction-limited performance is observed by comparing the cross-sectional intensity plots with diffraction-limited Airy functions (Figure 2.4g-i). The simulated focusing efficiencies ranged from 10% to 13%. Further improvements in efficiency can be realized by implementing phase shifting within zones [41], [42], potentially using the doping dependence of the refractive index (Figure 2.1).

2.3.3 Fresnel zone plate experimental and simulated results

Binary amplitude transmission zone plates have a theoretical focusing efficiency of $1/\pi$ ($\sim$10%) assuming the zones are completely opaque, very thin (much thinner than the wavelength of light), and infinitely large in diameter [43]. In practice, this is defined as the integrated intensity of the focal spot in the focal plane, normalized to the integrated intensity from the source through an open aperture of equal area to the zone plate [44], [45]. Experimentally this is done using a photodetector to record intensity and small aperture to eliminate any signal from other diffracted orders that are considered background (known as an order sorting aperture) [45].

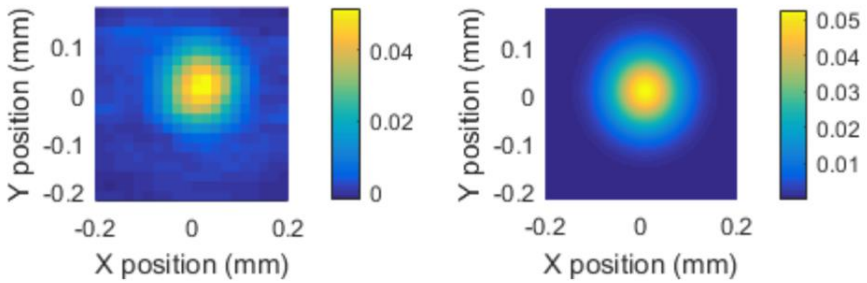

Figure 2.5 - a) The focused intensity profile of the laser source through a 9 mm open aperture with a corresponding 2-D Gaussian fit shown in b).

We experimentally calculated the focusing efficiencies of our zone plates as follows. First, an open aperture equivalent to the diameter of our zone plates (9 mm) was placed in front of the expanded source beam and an undoped silicon wafer. A plano-convex zinc selenide lens (anti-reflection coated, >98% transmission at 10.6 μm) was used to focus the light towards the 30 μm pinhole mounted in front of our detector. Due to thermal fluctuations in the laser, the output power and emitted wavelength could vary on the order a few percent. Therefore, we varied the cooling temperature of the laser between 27.2 °C and 30 °C in order to stabilize the output. The intensity profile of the focused light (Figure 2.5) was numerically integrated to get an effective power. Then the open aperture was replaced with each of our zone plates with the setup shown in Figure 2.4a. The focal spot intensity profiles are taken to be the first-order diffracted light from the zone plates. By integrating each of these intensity profiles and normalizing to the integrated intensity profile through the aperture, we calculated the focusing efficiencies. The measurements were repeated several times with Table 1 summarizing the results.

The range of measured efficiencies in Table 1 is quite wide. We attribute this significant variability to serval factors. The stability of the laser output power was difficult to maintain. As mentioned earlier, the output power and wavelength of the laser is controlled by setting the laser cavity temperature. Drifts or jumps in output power due to temperature drifts were detected with

our infrared detector and were difficult to account for. Due to the finite thickness of the semiconductor substrate (~50 wavelengths thick) and the narrow spectral bandwidth of the source, significant Fabry-Perot interference effects dominated transmission through the silicon substrates. Slight changes in positioning of the experimental samples or undoped reference slab created large variations in detected transmission.

Zone plate design	Experimental efficiency (%)	Gaussian fit efficiency (%)
2 cm focal length	8.8 – 14.9	9.7 – 15.7
6 cm focal length	5.7 -31.9	6.4 -30.3
10 cm focal length	7.4 [a)]	8.2*

Table 1. Range of measured efficiencies. Experimental efficiency is through numeric integration of the raw profile data, while Gaussian fit efficiency is the calculated efficiency after fitting a 2D Gaussian to both the zone plate and source data from Figure S8.

[a)] Only a single measurement was performed. Sample was damaged afterwards.

Reducing this experimental variability for the efficiency measurements proved difficult, so instead we performed full-wave simulations (Lumerical FDTD Solutions [39]) of each zone plate to determine our device performance. Using the designed zone pattern in combination with our ellipsometrically extracted optical properties for the doped silicon, two-dimensional simulations were performed to determine the focal spot intensity profile. The intensity profile of a plane wave source through a 9 mm open aperture was also calculated. The power within the region from the peak to the first minimum of the focus spot was integrated and normalized to the integrated intensity of the plane-wave source through the open aperture to calculate the efficiencies.

Our embedded doped material performed at >97% of an ideal perfect conductor. The discrepancy between the theoretical efficiency and our simulated efficiencies has been attributed to both the finite nature of our devices (not infinite in size) and the 2D simulation approximation of a three-dimension symmetric device.

2.3.4 Mid- and far-infrared frequency selective surfaces

In section 2.3.3, the highly doped silicon was used as an opaque material for diffractive optics. Now we examine the use of area-selective doping for plasmonic frequency selective surfaces (FSSs) that are monolithically embedded within the substrate. FSSs are thin, periodic structures that are designed such that their electromagnetic resonances either absorb, transmit or reflect electromagnetic fields in a given EM spectrum based on the frequency of the incident field [46]. Here we look at a specific case of FSSs designed to be bandstop filters. FSSs designed as a bandstop filter are engineered such that at their resonant frequency they either absorb or reflect the incident while allow frequencies off-resonance to pass through. We designed and fabricated band-stop FSSs (similar to those often implemented using metals [46]) for two different wavelengths (Figure 2.6), each comprising a periodic array of conductive patch antennas, defined by photolithography and ion implantation. At mid- to far-IR wavelengths much longer than the plasma wavelength (in our case, N ~10^{21} cm^{-3}, λ_p ~ 4.0 μm, see Figure 2.3d), the doped patches act as metallic antennas, for which the resonant wavelength is expected to be approximately twice the antenna length within the dielectric medium [47]. The FSS acts as a band-stop filter around this resonance [48]. The two FSS designs presented here consist of 3-μm-diameter antennas with 3-μm gaps, and 4-μm-diameter antennas with 4-μm gaps. The thickness of the antennas was assumed to be ~750 nm (same thickness as in section 2.3.3). At these dimensions, the resonances are expected to be at ~6 μm and 8 μm within the surrounding dielectric (Re[ϵ] ~ 11.7), which corresponds to ~21 μm and 27 μm in free space.

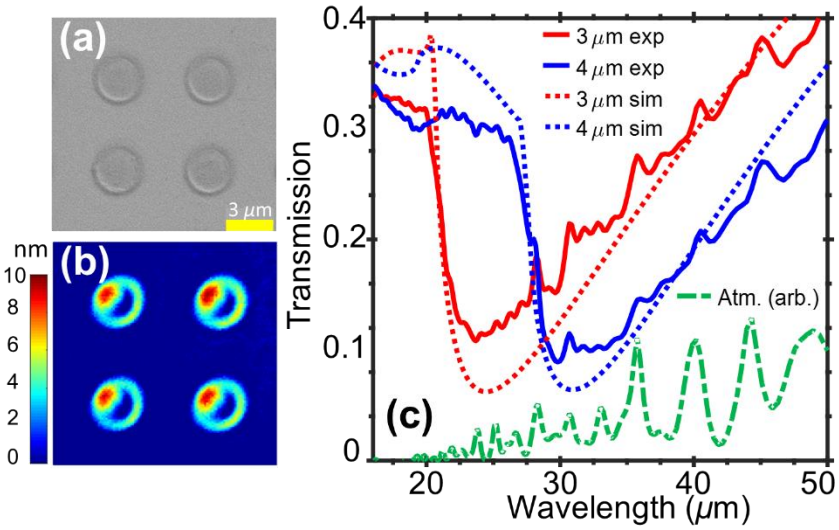

Figure 2.6 - a, b) Scanning electron microscope (SEM) image (top) and topographical map acquired using an interferometric optical profiler (bottom) of the surface of the 3-μm FSS. The surface is mainly smooth, but some height variation is visible in the highly doped region due to swelling and surface damage. The asymmetry in the height is attributed to the off-axis nature of the implant. c) FSS transmission measurements (solid) and corresponding FDTD simulations (dotted lines). The dashed green line shows the atmospheric absorption peaks that correlated well to the distinct features seen in the transmission.

We performed transmission measurements of the FSSs using an FTIR with a far-infrared detector and a Globar source. The experimental results, shown in Figure 2.6c, are normalized to transmission through undoped silicon and show the anticipated band-stop regions. The small features seen in the measurements were correlated with peaks in the atmospheric absorption. The transmission minima for the "3 μm" and "4 μm" FSSs are observed at 24 μm and 30 μm – red-shifted from our original prediction due to the increased size of the patches.

An FDTD model was made to approximate the post-implant devices by including an outer half-doped ring with a maximum diameter of 140% of the design diameter. These simulation results match well with the experimental measurements (Figure 2.6c). By tightening fabrication tolerances and minimizing dopant diffusion using laser [49] or flash-lamp [16] annealing, the shift of the FSS resonances from the original design can be minimized.

2.3.5 Conclusion

In summary, we established area-selective ion-implanted silicon as a platform for flat, monolithic mid-infrared optical components that can be easily integrated into other silicon-based technologies. We designed, fabricated, and tested Fresnel zone plates (FZPs) and frequency selective surfaces (FSSs), demonstrating our ability to locally tune the optical properties of silicon via area-selective ion-beam doping.

Even higher carrier densities (beyond the solubility limit) can be achieved using laser [49] or flash-lamp [16] annealing, potentially yielding plasma wavelengths in the near-infrared range and mitigating unwanted dopant diffusion. Semiconductors with higher carrier mobilities and lower effective masses, such as indium arsenide, are amenable to ion implantation [50]–[52], and can be substituted for silicon to reduce losses and improve plasmonic device performance. Finally, this technique to create monolithic plasmonic devices can be extended to the spatial control of dopants in all three dimensions by combining lithographically defined masks with control of the energy of implanted ions.

2.4 Gradient-index-doped semiconductor devices

In addition to operating in the plasmonic regime as previously described, we also sought to investigate what could be accomplished by operating at wavelengths higher than the plasma frequency. As stated previously, the refractive index in this regime can be modified based on the dopant concentration [24]. Referring back to Figure 2.1, we can see the how refractive index of silicon changes as a function of concentration. One thing to note is that the imaginary part of the

refractive index, which governs the amount of loss, is relatively high for silicon. Therefore, we decided to investigate other materials in order to minimize the loss.

Silicon is generally lossy in this regime due to the low electron mobility [9]; other semiconductors such as germanium and some III-V compounds have much higher mobilities. For our demonstration of this technology, we choose to try indium arsenide (InAs), a III-V semiconductor [53]. The advantage of this material is an electron mobility of $\mu_{InAs} = 33000\ cm^2V^{-1}s^{-1}$ compared to $\mu_{Si} = 1900\ cm^2V^{-1}s^{-1}$ of silicon (both values at taken at 300

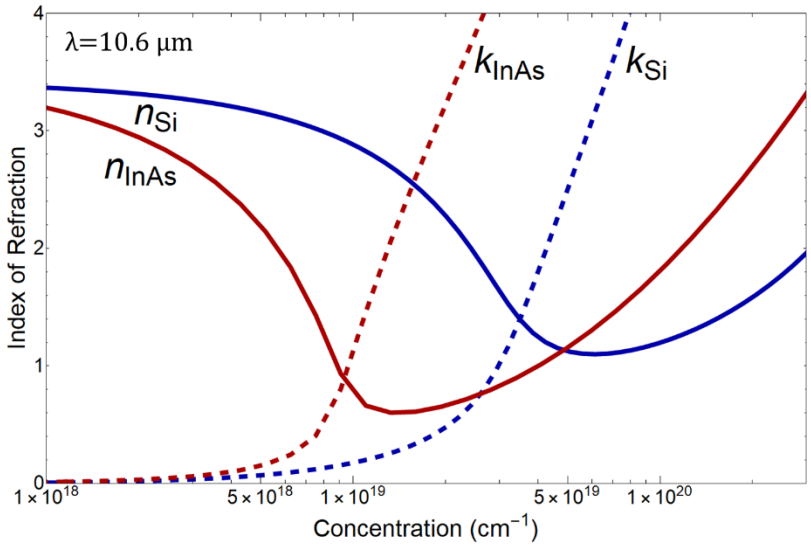

Figure 2.7 – Calculated complex refractive index as a function of free carrier concentration for Si and InAs. InAs has a larger real index change for a given value of the imaginary due to the higher electron mobilities over Si.

°K) leading to significantly less loss [54]. This can be seen in Figure 2.7 where the refractive index of both materials are calculated using equations 2.1-2.4 with empirical values for the material constants coming from [9], [53], [55]. The reduced losses can be seen in that the crossover point, where $n = \kappa$, corresponds to lower value of κ for InAs over that of Si. Further, the amount of real refractive index change is much larger for InAs for a given value of κ meaning that we can have a higher index gradient with less loss if we use InAs.

For typical doping processes in semiconductors (e.g., ion-implantation, diffusion, or epitaxy), there is natural diffusion of dopants during the annealing process [40], [56], [57]. This diffusion produces a gradient in dopant concentration (which may be larger or smaller depending on anneal process, length, or temperature). Our aim was to leverage this gradient in dopant concentration in order to produce a gradient in refractive index. The gradient in refractive index would allow us bend and manipulate the optical wavefront similar to how more, well-established, gradient index (GRIN) optical components work.

GRIN devices work because a change in the refractive index leads to a change in propagation velocity of the electromagnetic wavefront. Light propagates slower in a higher index medium versus a lower index medium[58]. Therefore, a constant phase front of the wave propagating in a GRIN medium would naturally bend towards a positive index gradient (bending from low index to high). This effect has been utilized to create common optical components such as waveguides and lenses.

To maximize the dopant concentration gradient, we decided to investigate and model surface diffusion to introduce dopants into the semiconductor[56], [59]. For InAs a good dopant source is tin (Tn). There are commonly available Tn-doped spin-on-glass (SoG) commercial products that can be used for this purpose.

The proposed design process can be visualized in Figure 2.8. First, a SiN_x (unknown stoichiometry) masking layer is placed down on the surface. This masking layer is to prevent outgassing of As during the anneal process, which is a known issue for III-V semiconductors[56], [59]–[62]. A layer of photoresist will then be deposited on top of the SiN_x layer to enable photolithography-based patterning of the dopant source locations. After the lithography step, the locations for the dopant sources will be etched via reactive-ion etching. The resist will then be

removed and the SoG will be deposited and baked at 200 °C to set. Then a high-temperature thermal anneal can be performed to drive in and activate the dopants. For InAs, the proposed anneal temperature is 650 °C for two hours. The SoG can then be stripped with a hydrofluoric acid etch and the SiN_x can be removed with another wet etch. This proposed process will then result in a final device that is monolithic with select areas doped and annealed to locally modify the index of refraction.

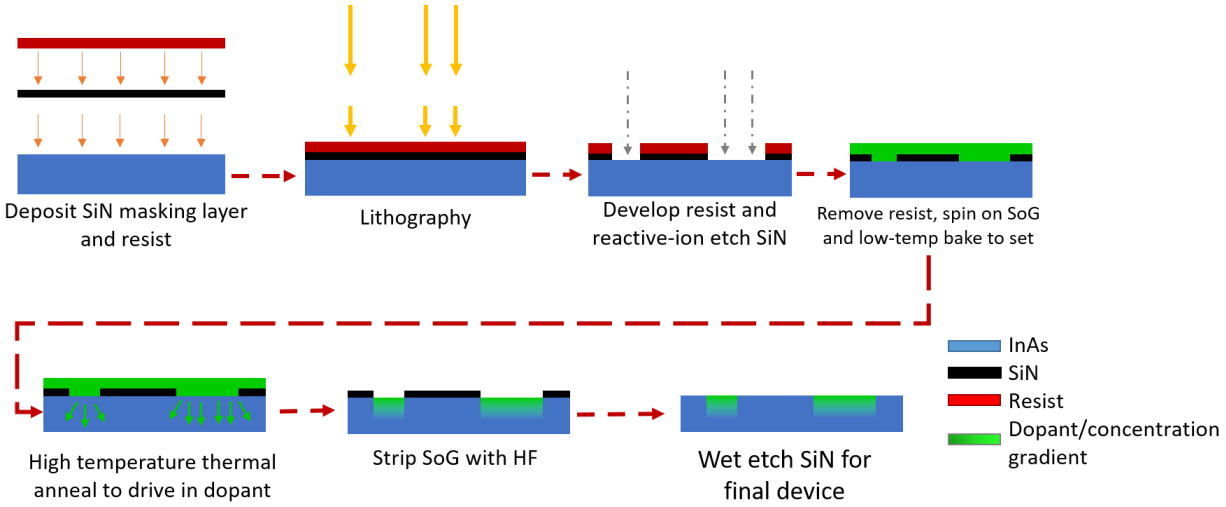

Figure 2.8 – Proposed fabrication process of GRIN-based doped semiconductor optical devices. A SiN_x masking layer is placed down under a photoresist layer. The pattern is then lithographically exposed, developed, and etched, leaving holes where the dopant sources can penetrate into the semiconductor. The spin-on-glass is applied and the device is annealed to drive in and activate the dopants. Finally, the SoG and SiN_x layer is stripped, leaving a monolithic doped device.

However, there still remains a significant question to answer: where and how large does one place the dopant sources in order to achieve a desired wavefront? This question is not straightforward to answer analytically and there are many parts that contribute to the overall performance especially when dealing with dopant sources that are on the order of the wavelength of interest. For a single source, it is hard to analytically calculate how the light will interact with an index gradient on that length scale. Granted, one can calculate the diffusion profile, the

refractive index profile, and then simulate the resulting wavefront. However, multiple diffusion sources are needed, which interact with one another and spread out over a large, multi-wavelength area of the surface. Because of this, building intuition on how place multiple diffusion to achieve a desired result is difficult.

Furthermore, there are many tradeoffs when designing such a device. To have more control over the wavefront, one can either diffuse the dopants deeper into the substrate or aim for a higher concentration, which would give a larger index contrast. If the dopants are diffused deeper, they will also diffuse wider which would lead to a loss in lateral resolution and less fine control over the wavefront phase. If a higher concentration is used, the imaginary part of the refractive index becomes larger, and more loss is introduced leading to a less-efficient device.

To overcome these tradeoffs, we implemented a global optimization algorithm to solve for the location of our dopant sources to create a desired wavefront. The entire algorithm can be visualized in Figure 2.9. The optimization loop begins with randomly placed and randomly sized dopant sources placed on the surface. The algorithm attempts to solve for the ideal position and width of each dopant source. First, a finite-element simulation is run to calculate the diffusion profile of the dopant sources (details below). Once the dopant concentration is obtained, we then calculate the index of refraction throughout the structure according to Eqs. 2.1-2.4. The calculated index is then inserted into a full-wave FDTD simulation (Lumerical FDTD [39]) which simulates the resulting wavefront after it passes through the device. From that simulation, we can obtain a figure-of-merit (FoM) which is a user-defined quantity that defines the performance of the device. This FoM could be percent power into a diffracted order or power directed to a focal spot. Based on the feedback from the FoM, we can then modify the positions of the dopant sources and run another iteration.

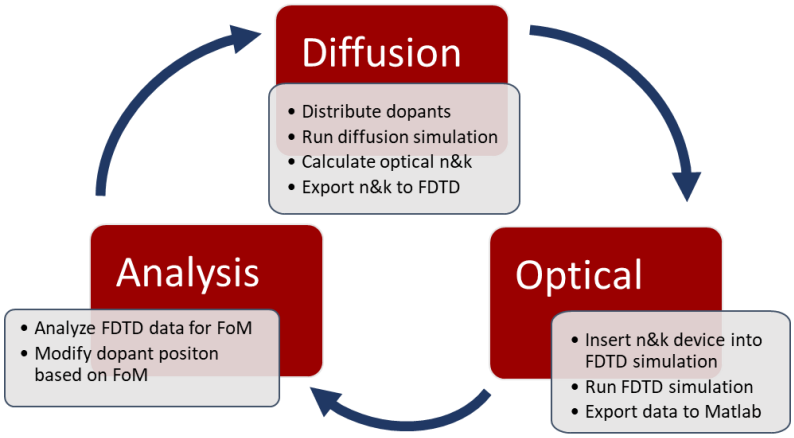

Figure 2.9 – Visualization of the optimization algorithm to design doped-semiconductor GRIN devices. The optimization is seeded with random distribution of dopants. The diffusion simulation is run followed by a refractive index calculation. That result is passed to an FDTD simulation where the output E-fields are analyzed and used to modify the position of the dopants. This algorithm is compatible with multiple global optimization methods such as Swarms, Evolutionary algorithms or Simulated Annealing.

We used MATLAB's built-in finite-element solver to simulate the diffusion profile for dopants placed on the surface. The diffusion of particles into solids is given by Fick's second law. The concentration dependent diffusion equation is then given by the partial differential equation (PDE)[56], [63]:

$$\frac{\partial C}{\partial t} = \nabla \cdot \left(D_{eff} \cdot \nabla C\right) \tag{2.5}$$

where

$$D_{eff} = D_0 \cdot \left[\frac{C(x,y)}{C_0}\right]^{\Delta q} \tag{2.6}$$

Here, C is the dopant concentration, C_0 is the concentration at the surface, D_{eff} is the concentration-dependent diffusion coefficient, D_0 is the diffusion coefficient for the concentration at the surface, and Δq is the difference in charge states of the impurity in a substitutional vs

interstitial location, which is found empirically for Sn in III-V compounds to be 2 [56]. Equation 2.6 givens the value of the diffusion coefficient which is inserted into Eq. 2.5.

This nonlinear PDE can be solved numerically, but care is needed when defining the simulation boundaries. At the top surface (*i.e.*, the surface at which the dopants are introduced) we use a combination of Dirichlet and Neumann boundaries. For PDE's, Dirichlet boundaries hold a constant value while Neumann boundaries define the flux through the boundary[64]. Therefore, we set the boundary areas with dopant sources as Dirichlet boundaries with the max solubility value of the semiconductor substrate (for InAs this is around 10^{19}). The surface areas without a dopant source were defined as Neumann boundaries. We the set the flux through the Neumann boundaries equal to 0 under the assumption that no impurities flow out of the substrate through those boundaries. This is a reasonable assumption given a capping layer, although this model could be modified to allow some diffusion through these boundaries if that occurs.

The full optimization loop is compatible with any global optimization algorithm and the appropriate algorithm can be determined from the design parameters. For instance, if the device simulations are quick to solve, one can choose a more globally robust algorithm such as swarm-based or evolutionary algorithms [65]–[68]. These algorithms utilize many particles in each generation and therefore can be run effectively in parallel. However, if the simulations are larger, it may be advantageous to use an algorithm such as simulated annealing, as that method uses a single simulation per generation [69].

As a proof of concept, we aimed to develop a focusing lens from gradient-doped InAs using this optimization algorithm. To start, we aimed for a surface doping concentration of 10^{19} as this is the upper solubility of InAs using spin-on doping techniques [25], [56]. We targeted a wavelength of 10.6 μm because it is a common wavelength for mid-IR applications and InAs has

relatively low loss at that wavelength at our chosen max doping concentration. The target focal length 300 μm with a lens diameter of 250 μm giving a numerical aperture of 0.38. These parameters were chosen arbitrarily as a proof of concept and to gauge the performance possibilities.

To optimize this particular device, we choose the particle-swarm optimization method. In addition to the dopant source locations being free parameters, the anneal time was also allowed to vary between 30 mins and 10 hours. The swarm optimization consisted of 270 particles and was allowed to run for a maximum of 1300 generations (although FoM convergence happened around generation 300) [70].

At the outset of the algorithm, the dopant sources are randomly placed on the surface. Figure 2.10a shows the simulated electric field intensity distribution of the devices at the beginning

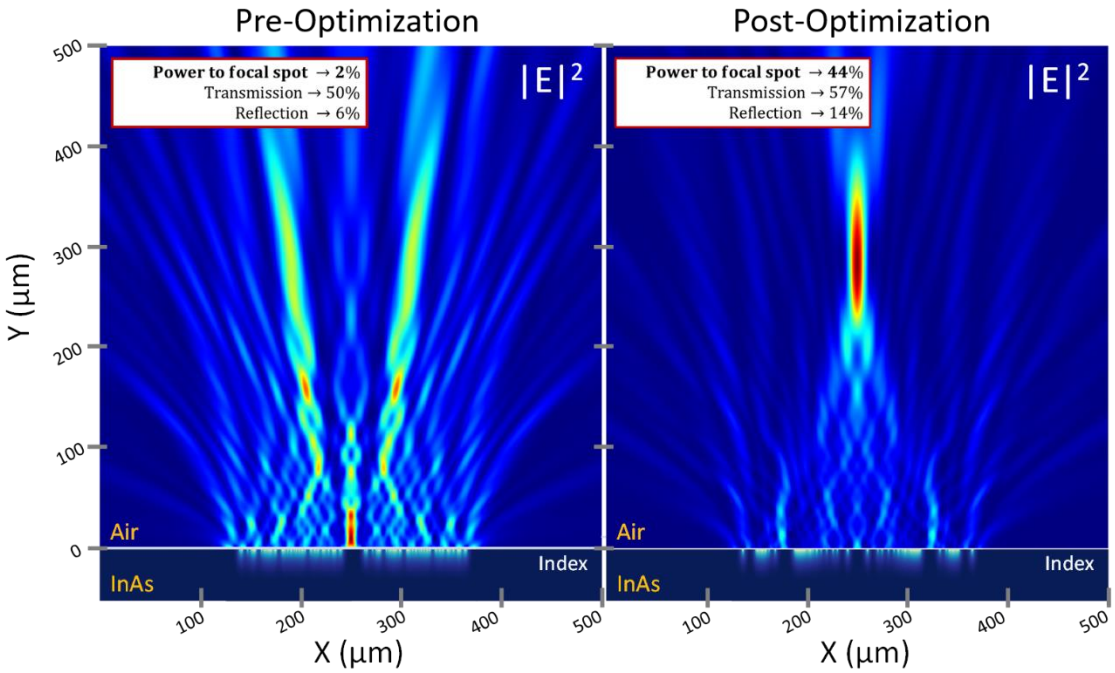

Figure 2.10 – a) Concentration-based refractive index profile along with electric field intensity of the GRIN-based lens prior to optimization. The field is randomly diffracted with minimal power going to the focal region. b) The field profile after undergoing the optimization routine which closely resembles that of a lens. The index profile also seems to mimic that of a Fresnel lens. 77% of the transmitted power is directed into the focal spot region.

of the optimization run. It can be seen here that the field is generally random with light being randomly spread into the far field. The focal spot region (defined as the region $\pm 2\lambda$ in the x and y direction centered at the target focal spot) contained only 2% of the injected power. Figure 2.10b shows the resulting $|E|^2$ after the optimization has converged. It can be seen that the fields strongly converge to the focal point and 77% of the transmitted light reaches the focal spot. Of the total injected power, 44% reaches the focal point region with 14% reflected and 29% absorbed. The remaining power (13%) is transmitted but does not reach the focal point.

It is interesting to note the final dopant concentration in the InAs as it gives some intuition as to the design of the structure. It does appear that the central region remained relatively free of any dopants while the dopants on the side create almost Fresnel-lens like fringes along the side. This makes intuitive sense as the change in refractive index would create a diffraction pattern similar to a Fresnel lens.

From a proof-of-concept standpoint, this device works well. We also tested how the optimization algorithm could perform for a beam deflector. Beam deflectors control the propagation direction of light passing through them and are a useful benchmark of metasurfaces and other flat optical technologies as they require full phase control of the incident beam and the phase control needed is proportional to deflection angle (with higher angles requiring more precise control over the phase)[71]–[74]. For our device, the resulting far-field pattern can be seen in Figure 2.11. Here 96% of the transmitted power is diffracted into the target direction. The inset shows the resulting index distribution for this device.

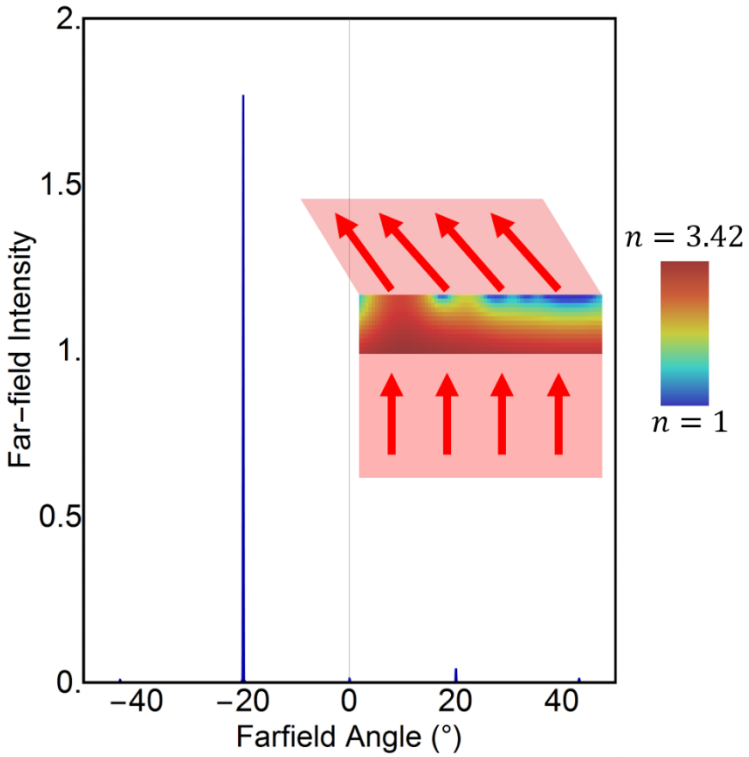

Figure 2.11 – Far-field plot of a 20° beam deflector simulated using the GRIN optimization algorithm. The inset shows the refractive index distribution with blue being lower index and red being higher. 96% of the transmitted light was diffracted into the target order.

Since these proofs of concepts were promising, we attempted to move forward with fabricating a real device, ideally a lens based on the optimization algorithm. However, we ran into many challenges during the fabrication process. The main issue that we ran into was the outgassing of As during the anneal process. We spun on Sn-doped SoG from Desert Silicon (Sn-0365HP) prior to attempting a 2 hour anneal at 650 °C. After the anneal process, it was found that most of the As in the top of our sample evaporated. We preformed energy dispersive X-ray (EDX) analysis on the InAs sample pre- and post-anneal. Figure 2.12 shows the resulting spectra and SEM images. It can clearly be seen that the characteristic As line has disappeared post anneal and the surface quality of the sample has been significantly degraded. It did appear, however, that some Sn made its way into the sample, evidenced by the appearance of the characteristic Sn line.

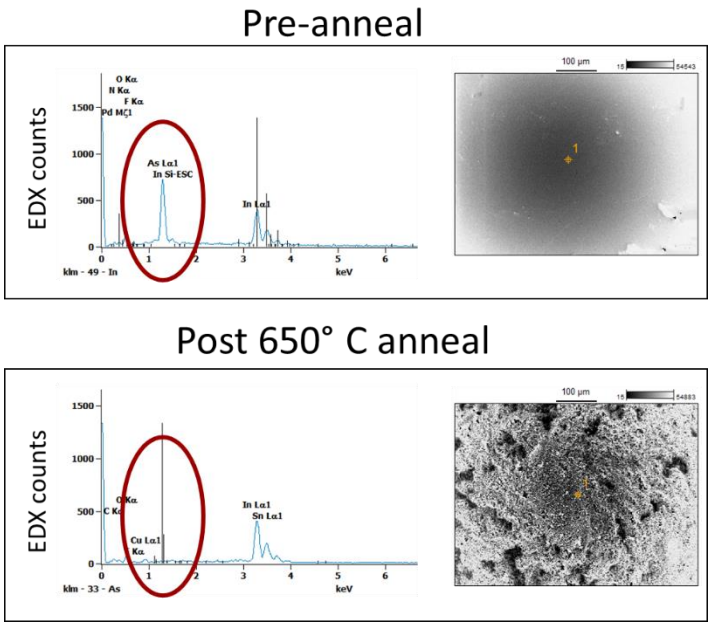

Figure 2.12 – Pre and post anneal EDX and SEM images of our InAs sample coated with Tn-doped SoG. Extreme outgassing of As is present along with significant degradation of the surface due to the outgassing.

To prevent the As outgassing, we attempted to utilize a SiN capping layer[56], [59]–[61]. We deposited 120 nm film of SiN on top of the InAs wafer via physical vapor deposition. We then attempted the anneal process again (650 °C for 2 hours). Figure 2.13 shows the results of this anneal. It can be seen that while some areas remained protected and pristine, there were significant pockets of outgassing that seemed to "explode" through the capping layer. Figure 2.13A shows an optical microscope image of a portion of the sample. A zoomed in SEM image of a single defect can be seen in Figure 2.13B. The surface pattern seen here is very similar to the surface pattern seen in Figure 2.12. Looking at the EDX images in Figure 2.13C and D it can be seen that the area around the defect still maintains significant As concentrations. However, the defect area is again devoid of As due to outgassing.

Due to shifting projects and limited time investment available, we were not able to ultimately fix this issue. However, we do believe that there are a few steps that can be taking to continue this work towards a working prototype. First, we believe the InAs annealing could

probably work well if the anneal was done in a furnace or capsule that had significant As overpressure[60]. This would prevent the As from being able to outgas and hopefully would maintain the surface quality. Since that is a more complicated fab process, the alternative would be to try another semiconductor/dopant material system. It could be beneficial to focus on an elemental semiconductor, such as Ge, as there is no risk of outgassing during the anneal process. Ge also has a higher electron mobility than Si ($\mu_{Ge} = 8800\ cm^2/Vs$ at 300 °K) [54] and therefore could also have lower loss than Si.

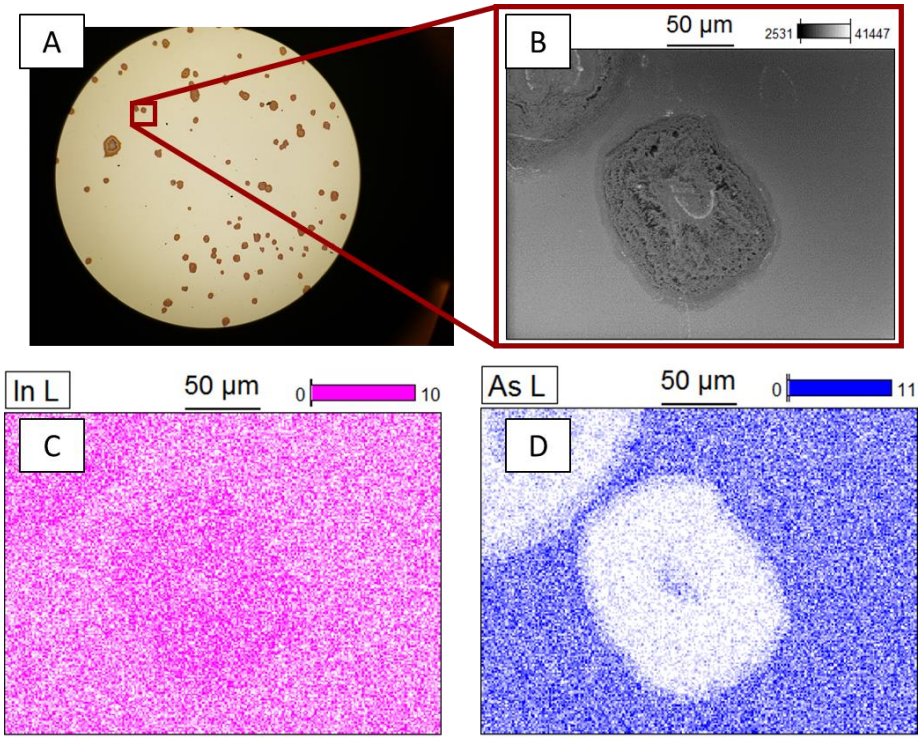

Figure 2.13 – A) Optical microscope image of the InAs surface post a 2 hour, 650 °C anneal. Many pockets of outgassing are visible. B) SEM image of one pocket of As outgassing showing surface degradation. C) and D) EDX counts for In and As in the area of B). It can be seen that while there is As in the pristine sample areas, the outgassed area has little to no As concentration.

In conclusion, we presented a novel optical materials platform based on gradient-doped semiconductors that can bend and shape light in a monolithic, stress-resistant package. We developed a multi-physics optimization algorithm which can solve the dopant location problem

and create relatively high-performing devices. We believe further work in this could lead to the creation of extremely robust, thin, and bespoke optical elements for compact applications.

3 Adjoint optimization of nanoscale light extractors for NV centers in diamond

3.1 Introduction to adjoint optimization

Maxwell's equations encode our understanding of how light interacts with matter. They tell us, for example, how a propagating wave will interact with a specific material arranged in a given shape or pattern. The solution to such a problem is relatively straightforward and can be computationally solved through a variety of methods such as Finite Difference Time Domain (FDTD) simulations and the Finite Element Method (FEM). However, in scientific or engineering applications, the solution to this problem has limited utility. This arises because we are concerned with designing the best structure to conform to a particular need. The issue is that, while we need to know how a given structure interacts with light, we primarily want to know what structure will give us the most *optimal* interaction, given a certain figure of merit or target goal. This difference is understood as the difference between the *forward* problem (what is the field produced by a particular device) and the *inverse* problem (what is the device that best produces a particular field).

The challenge here is that the forward problem has a singular solution. Due to the nature of Maxwell's equations, there can be only one resulting field given a specific structure and input field. The inverse problem, however, is not guaranteed to have a physically realizable solution and, furthermore, if the solution exists, it cannot be guaranteed to be unique; there may be other

geometries that satisfy the design constraints equally well. And so, left with this conundrum, as scientists and engineers we seek the best physically realizable solution to our problem.

In the literature, the design of optical metasurfaces has often used a basic form of inverse design, where first maps of transmission phase and amplitude as a function of the various free parameters of the meta-atom (unit cell) geometry are generated [75]. For example, if one desired to design a metasurface beam deflector composed of Si rods on SiO_2 substrate [76], [77], one would start by simulating the complex transmission (amplitude and phase) of the rods as function of height and rod radius, for a given size of the unit cell. Once that map is complete, specific geometries would be chosen based on what phase is needed at a given location on the device and the entire device would be populated this way.

In this example, we can see that the inverse-design approach is used, but operating under stronger assumptions and constraints than necessary. We have a desired field (for a beam deflector, this is constant amplitude with a periodic phase variation with the period depending on the desired deflection angle [78]) and we choose subwavelength geometries that can best scatter the light to approximate the desired field. However, the resulting geometries are limited. Adjoint optimization is a relatively recent optimization technique that helps solve the inverse design problem and can relax the constraints put on the geometry design space.

While adjoint optimization itself has been used for quite some time in mechanical engineering applications [79]–[83], it has only been within the past 10 years that it has seen development and use for photonic component engineering. Today, there are many good resources for understanding the derivation of the adjoint method for photonic applications. Of particular use to us have been the works by Sigmond Ole[84]–[88] and some of the publications from Eli Yablonovitch's group at the University of California at Berkeley. The thesis by Owen Miller [89]

contains a substantial derivation and commentary on the adjoint method applied to photonics and the interested reader can reference that document for further elucidation of the methods.

3.2 NV color centers

The remainder of this chapter focuses on our own implementation of the adjoint optimization method to design photonic devices that enhance the light extraction from nitrogen-vacancy (NV) defect color centers in diamond crystals. NV centers exist in diamond as two adjacent defects where one carbon atom sis replaced with a nitrogen atom which bonds to a neighboring atom-site vacancy defect [90], [91]. This pairing creates a distinctive quantum system that, due to the stable nature of the diamond crystal, remains stable (retaining relatively well-defined energy levels) up to room temperature [92]–[94]. The usefulness of this system is primarily due to the accessibility of the electron spin states via optical, magnetic and RF radiation[91]. This defect system is useful for quantum information processing as the electron spin information can be recorded for long periods of time into nuclear spins and then read out later[92], [95].

The second major application of color centers, and the application that we are primarily concerned with, is the ability of these systems to sense very weak and highly localized electric and magnetic fields and transduce them into optical signals [96], [97]. This is primarily due to the splitting of the electron states via the Zeeman or Stark effects [98], [99]. The splitting magnitude can be detected via optically detected magnetic resonance (ODMR) measurements in which the electrons are pumped to the excited via visible laser light, their spin is moved to the +/-1 state via microwave radiation and the non-radiative decay valley splits as a function of external DC field strength [100].

One challenge with this measurement technique that we seek to address in this chapter is the low signal-to-noise ratio (SNR) due to the limited collection of the light emitted by NV centers. Due to the high refractive index of the diamond, much of the emitted light from the NVs gets totally internally reflected [101], [102]. Both the sensitivity to external fields as well as spatial resolution depend on the ODMR SNR, so maximizing the signal output of the NVs is a primary concern.

And so, our engineering question becomes, what geometric structure can we design to extract the most light emitted by the NV defect? To answer this question, we turned toward adjoint optimization as we realized that the constraints imposed by tradition design techniques would not lead to high-performing structures but also that there really are no straightforward design techniques to answer this challenge.

The next section in this chapter will deal the initial implementation of our adjoint optimization algorithm applied to NV centers in [100] oriented diamond. This section outlines the specifics of our algorithm and the techniques we implemented in order to design fabricable nanoscale light extractors (NLEs) to enhance the emission from the NV centers. In the second section, we present an expanded array of NLEs using more adjoint figures of merit for various extraction geometries for [111] oriented diamond. Along with a stock extraction NLE, we present a geometry with a central void for molecular sensing, an NLE that minimizes excitation fields above the structure for backside illuminated NVs, and an NLE that splits the NV emission into two output modes based on the orientation of the NV emission dipole. This compilation of NLEs is meant to provide useful geometries for experimental NV sensing setups as well as demonstrate further capabilities of the adjoint optimization method.

3.3 Nanoscale light extractor for NV centers in (100) diamond

The following text is taken from our publication:

R. A. Wambold, Zhaoning Yu, Yuzhe Xiao, Benjamin Bachman, Gabriel Jaffe, Shimon Kolkowitz, Jennifer T. Choy, Mark A. Eriksson, Robert J. Hamers and Mikhail A. Kats, "Adjoint-optimized nanoscale light extractor for nitrogen-vacancy centers in diamond," Nanophotonics, vol. 10, no. 1, pp. 393–401, Nov. 2020, doi: 10.1515/nanoph-2020-0387. This work was performed in collaboration with the groups of Mark Eriksson, Bob Hamers as well as contributions from Shimon Kolkowitz and Jennifer Choy at UW Madison.

3.3.1 Introduction

Negatively charged nitrogen-vacancy (NV) centers in diamond are optical emitters whose level-structure is highly sensitive to external perturbations, which makes them excellent sensors of highly localized electric and magnetic fields, temperature, and strain [103]–[107]. NV centers are of great interest for quantum computing and communication [108]–[112] and the study of quantum phenomena such as quantum entanglement and superposition [113], [114]. However, efficiently extracting NV fluorescence is often challenging due to the high index of refraction in diamond (~2.4), which results in high reflectance at the diamond-air interfaces and total internal reflection for emission angles larger than the critical angle. Previous attempts to extract more light from bulk diamond primarily involved the etching of the diamond itself (a complicated fabrication process that can adversely affect NV properties such as spin coherence) [115]–[121] or fabricating structures that still required a high numerical aperture oil-immersion objective to efficiently collect

the emission (which adds system complexity and is detrimental to sensing applications) [122]–[125]. Furthermore, precision etching of diamond around NV centers can be a substantial challenge and can damage the surface of diamond, resulting in roughness and modifying the chemical termination [126], which can degrade the quantum properties of NV centers [127], [128].

Here, we design a silicon-based nanoscale light extractor (NLE) that sits on the top of a flat, unpatterned diamond surface and can enhance the optical output of near-surface NV emitters by more than 35× compared to the unpatterned case, directing the light into a narrow cone that can be easily collected with low-NA optical systems. Our NLE consists of a patterned silicon structure on top of the diamond surface (Figure 3.1), directly above a shallow NV center (< 300 nm below the surface). The proximity of a resonant high-index dielectric structure close to the emitter enables near-field coupling to and broadband Purcell enhancement of the emitter [125], [129]. We designed the silicon NLE using an adjoint-optimization method, optimizing for NV emission funneled into a narrow cone into the far field. This approach both increases collection efficiency and enhances the radiative emission rate of the NV center.

Our approach focuses on broadband emission enhancement of negatively charged NV centers close to the diamond surface, which is especially useful for sensing applications in which measurement sensitivity is limited by the number of collected photons across the entire NV emission spectrum.

3.3.2 Figure of merit for broadband NV emission extraction

We targeted our design towards negatively charged NV centers in [100] diamond. The NV axis is at an angle of 54.7° with respect to the normal to the diamond surface in this orientation [90], and the optical dipole moments are orthogonal to the NV axis [94]. At room temperature, the

emission from an NV center is expected to be unpolarized [130] and can therefore be simulated by the incoherent sum of emitted intensities from any two orthogonal linear dipoles that are also orthogonal to the NV axis. In bulk diamond, negatively charged NV centers have optical transitions at ~637 nm (the location of the zero-phonon line in the spectrum), but fluoresce over a bandwidth of >150 nm due to vibrational side bands in the diamond [90], [131]. Therefore, we define a spectrum-averaged figure of merit (FoM) to quantify the degree of light extraction:

$$FoM = \frac{\int I_{NV}(\lambda) \times \eta(\lambda)\, d\lambda}{\int I_{NV}(\lambda)\, d\lambda} \tag{3-1}$$

Here, $I_{NV}(\lambda)$ is the normalized emission spectrum of the NV centers in bulk diamond taken from ref. [132] and $\eta(\lambda)$ is the extraction efficiency of our NLE, which is defined as the number of photons emitted into free-space in the presence of the NLE divided by the number of photons emitted into free-space by that same NV (i.e., same depth and orientation) with no NLE present, at a wavelength λ. Note that $\eta(\lambda)$ includes both collection and Purcell enhancements. For example, an NLE that results in 10 times as many photons emitted into free space at every wavelength would have $\eta(\lambda) = 10$ and $FoM = 10$.

For all plots in this paper, the bounds of the integral in Eqn. (3-1) are set to 635 – 800 nm to cover the NV zero-phonon line and phonon sideband emission spectrum. Our NLE achieves broadband NV fluorescence enhancement, which is particularly useful for sensing applications using shallow NVs, but the design approach can readily be adapted to focus on more-narrow spectral ranges (e.g., the zero-phonon line alone).

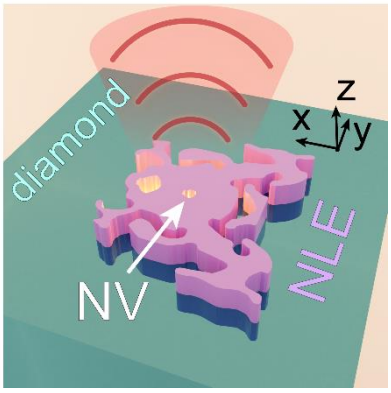

Figure 3.1 - Schematic of the nanoscale light extractor (NLE), which sits on the diamond surface above an NV center and directs fluorescence out of the diamond and into a narrow cone in the far field. The schematic is a render of the actual optimized structure reported below.

3.3.3 Adjoint-optimization method

To design a high-performance NLE, we used adjoint optimization—a design technique that has been used extensively in mechanical engineering and has recently been applied to the design of optical metasurfaces and other photonic structures [86], [133], [134]. For free-form structures that are allowed to evolve in three dimensions, the optimization process consists of the evolution of the structure, defined by a refractive-index profile $n(\boldsymbol{r}, \lambda)$, towards maximizing the overlap between a forward simulation field $[E_{\mathrm{fwd}}(\boldsymbol{r}, \lambda)]$ and a specific adjoint simulation field $[E_{\mathrm{adj}}(\boldsymbol{r}, \lambda)]$, where $\boldsymbol{r}$ is a position within the structure.

Due to the incoherent and unpolarized emission pattern of the NV center at room temperature, our approach was to maximize extraction from two orthogonal optical-dipole

orientations [94]. We used two pairs of forward/adjoint simulations for these two orientations. The first of our forward simulations is sourced by an electric dipole 10 nm below the diamond surface, with the dipole moment oriented in the XZ plane and tilted at an angle of 35.3° off the Z axis. The second was sourced by a dipole placed at the same position and oriented parallel to the Y axis. Note that NV centers in [100] diamond can be oriented in one of four directions along the (111) crystal axes. Thus to maximize performance, prior characterization can be done to align the NLE in the proper direction during fabrication [135]. Two adjoint Gaussian beams were used with orthogonal polarizations; one polarized along the X axis and the other polarized along the Y axis to pair with the forward dipoles in the XZ plane and along the Y axis, respectively.

We used finite-difference time-domain (FDTD) simulations (implemented in Lumerical FDTD [136]) to determine the forward and adjoint fields in the optimization region across the spectrum of interest for both optimization pairs. At each optimization generation, we calculated a figure-of-merit gradient similar to that of ref. [137], but averaged across the forward/adjoint simulation pairs:

$$G(\boldsymbol{r},\lambda) \propto \langle n(\boldsymbol{r},\lambda) \cdot Re\left[E_{\mathrm{fwd}}(\boldsymbol{r},\lambda) \cdot E_{adj}(\boldsymbol{r},\lambda)\right]\rangle_{pairs} \tag{3-2}$$

For the case of only one forward/adjoint pair, a positive value of $G(\boldsymbol{r},\lambda)$ indicates that a small increase in the refractive index $n(\boldsymbol{r},\lambda)$ at position $\boldsymbol{r}$ will result in a stronger overlap between the forward and adjoint fields. For more than one pair, like here, $G(\boldsymbol{r},\lambda)$ is a compromise quantity balancing the performance between all pairs. This causes the optimization algorithm to move in the direction that maximizes the gradient with respect to both pairs, and, while not necessarily reaching the same performance levels as with a single pair, is required in our case due to the nature

of the NV emission. Furthermore, because it is typically not possible to separately engineer the index at different wavelengths, we define a wavelength-weighted gradient:

$$G(\boldsymbol{r}) = \frac{1}{\Delta\lambda}\int I_{NV}(\lambda)G(\boldsymbol{r},\lambda)d\lambda \tag{3-3}$$

To design a structure that can realistically be fabricated using top-down techniques, there should ideally be no material variance in the vertical (Z) direction. We impose this constraint in our optimization by 1) forcing the index in a single column along Z to be constant, and 2) averaging $G(\boldsymbol{r})$ along each column during the update step such that a single G is applied to each column given by:

$$G(\bar{\boldsymbol{r}}) = \int_{z_{min}}^{z_{max}} \frac{G(\boldsymbol{r})}{z_{max} - z_{min}}\, dz \tag{3-4}$$

(similar to, e.g., ref. [137]), where $\bar{\boldsymbol{r}}$ is the two-dimensional position vector in the (x, y) plane.

Figure 3.2(A) shows the design process for our NLE. To start, the optimization is seeded with a random continuous distribution of complex refractive index, taking values between that of air (n_{air} =1) and crystalline silicon ($n_{\mathrm{Si}}(\lambda)$, from ref. [138]), calculated by $n(\bar{\boldsymbol{r}},\lambda) = p(\bar{\boldsymbol{r}}) \cdot n_{\mathrm{si}}(\lambda) + [1 - p(\bar{\boldsymbol{r}})] \cdot n_{\mathrm{air}}$ where $p(\bar{\boldsymbol{r}})$ ranges between 0 and 1, such that if $p = 1$, then $n = n_{\mathrm{si}}$. $G(\bar{\boldsymbol{r}})$ is then used to update the position-dependent refractive-index profile according to $p_{new}(\bar{\boldsymbol{r}}) = p_{old}(\bar{\boldsymbol{r}}) + c \cdot G(\bar{\boldsymbol{r}})$, where c is a normalization factor. In principle, this process can be iterated (while evolving c), until $p(\bar{\boldsymbol{r}})$ converges to yield an optimized index profile.

However, the optimization method described above yields an optimized profile $p(\bar{\boldsymbol{r}})$ corresponding to a continuum of index values $n(\bar{\boldsymbol{r}}, \lambda)$. To evolve the index distribution into a binary structure of silicon and air that can be made by lithography and etching, and also ensuring that the feature sizes are not too small, we followed the methods of Sigmund in ref. [84]. This approach applies a conically shaped blurring function to $p(\bar{\boldsymbol{r}})$ at each iteration to smooth the index distribution to remove features that are smaller than the cone radius, R. For our optimization we used a conical blurring function with $R = 40$ nm to result in structures that can be readily made with most electron-beam lithography systems. The new blurred profile $p_{\text{blur}}(\bar{\boldsymbol{r}})$, is created according to [84]:

$$p_{\text{blur}}(\bar{\boldsymbol{r}}) = \int \frac{w(\bar{\boldsymbol{r}}, \bar{\boldsymbol{r}}') \cdot p(\bar{\boldsymbol{r}})}{w(\bar{\boldsymbol{r}}, \bar{\boldsymbol{r}}')} \, d\bar{\boldsymbol{r}}' \tag{3-5}$$

$$w(\bar{\boldsymbol{r}}, \bar{\boldsymbol{r}}') = \begin{cases} 0, & \|\bar{\boldsymbol{r}}' - \bar{\boldsymbol{r}}\| > R \\ R - \|\bar{\boldsymbol{r}}' - \bar{\boldsymbol{r}}\|, & \|\bar{\boldsymbol{r}}' - \bar{\boldsymbol{r}}\| \leq R \end{cases} \tag{3-6}$$

where $w(\bar{\boldsymbol{r}}, \bar{\boldsymbol{r}}')$ is a weight function that becomes larger closer to position $\bar{\boldsymbol{r}}$.

In addition to blurring, a binary push function is required to finish the optimization with a fully binary structure (i.e., in our case, at each point in the optimization region the final material should be either air or silicon). The binarization method found in ref. [84] and [134] was the guide for our implementation. Specifically, this method modifies the blurred element matrix as follows:

$$p_{\mathrm{bin}}(\bar{\boldsymbol{r}}) = \begin{cases} \alpha e^{-\frac{\beta(\alpha - p_{\mathrm{blur}}(\bar{\boldsymbol{r}}))}{\alpha}} - (\alpha - p_{\mathrm{blur}}(\bar{\boldsymbol{r}}))e^{-\beta} \ , \ 0 \le p(\bar{\boldsymbol{r}}) \le \alpha \\ 1 - (1-\alpha)e^{-\frac{\beta(p_{\mathrm{blur}}(\bar{\boldsymbol{r}}) - \alpha)}{1-\alpha}} - \left(\alpha - p_{\mathrm{blur}}(\bar{\boldsymbol{r}})\right)e^{-\beta} \ , \ \alpha < p(\bar{\boldsymbol{r}}) \le 1 \end{cases} \quad (3\text{-}7)$$

where $p_{\mathrm{bin}}(\bar{\boldsymbol{r}})$ is the modified (new) value of p at position $\bar{\boldsymbol{r}}$. Here, α is a cutoff parameter that allows us to define dilated and eroded edges (to account for fabrication errors) but is set to the midpoint value of 0.5 during the optimization process. The parameter β controls the strength of binarization applied to the structure. A β-value of 0 gives no increased binarization to the structure leading to $p_{\mathrm{bin}} = p$. As β increases, values of $p < \alpha$ are pushed towards zero while values of $p > \alpha$ are pushed towards one. Like the blurring function, the binarization function is applied at every step of the optimization process, with β starting at 0 and increasing, until a binary structure is achieved at the end of the optimization [~ iteration 400 in Figure 3.2(C)]. We found that to further remove small features and move the structure away from sharp local maxima, it was beneficial to apply a second blurring function to $p(\bar{\boldsymbol{r}})$ once every 35 iterations late in the optimization process (after iteration 305 in Figure 3.2(C)).

We leveraged the inherently broadband nature of the FDTD method to simulate and optimize the performance of our device across the entire NV emission spectrum, simultaneously achieving a design that is robust to errors in fabrication. Previous works using adjoint optimization with a frequency-domain electromagnetic solver achieve fabrication robustness through simulating dilated and eroded devices each iteration, which corresponds to a three-fold increase in the number of simulations required [85], [134], [139], [140]. In our optimizations, fabrication robustness was built-in automatically due our requirement for broadband performance together with the scale invariance of Maxwell's equations (i.e., larger structures at longer wavelength

behave similarly to smaller structures at shorter wavelength). This correspondence between broadband performance and shape robustness has been observed previously for other inverse-design approaches [141], [142].

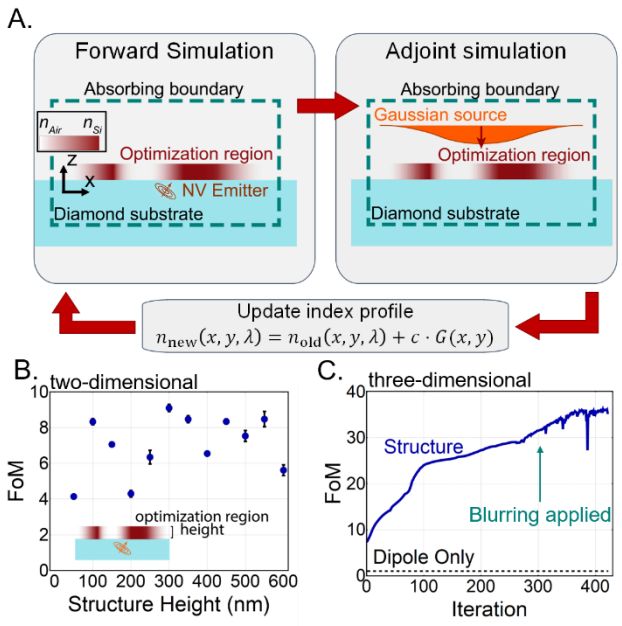

Figure 3.2 - A. Visualization of the optimization routine, which evolves an index profile situated above an NV emitter. In the forward simulation, the light source is a dipole at the location of the NV. The sources in the adjoint simulations are two orthogonally polarized Gaussian beams injected from free space toward the structure. The sensitivity gradient, $G(\bar{\boldsymbol{r}}) = G(x, y)$, is calculated and then used to evolve the index profile. The updated profile is then used in the next iteration. B. A sweep of the structure height (i.e., the thickness of the Si membrane), running five full 2D optimization cycles for each height. We found a height of approximately 300 nm to be optimal over the range of the sweep. The error bars represent the variance at each height. The variance is very small for some heights, so the error bars are not visible. C. FoM vs. the iteration number for a full 3D optimization run to generate our NLE. The dips are due to the application of the secondary blurring later in the optimization.

Note that our adjoint optimization setup that uses $G(\bar{\boldsymbol{r}})$ to evolve the structure does not directly maximize the FoM in Eqn. (3-1), because the formalism described above requires the selection of specific optical waveforms for the forward and adjoint fields, whereas the our FoM does not depend on the particular spatial mode(s) in which light escapes to free space, or the phase of that mode relative to the dipole source. Here, we selected our adjoint source to be a Gaussian beam with a diffraction angle of ±30° injected toward the diamond surface along the Z axis. To

relax the phase degree of freedom, we performed multiple optimization runs in which we varied the relative phases of the forward and adjoint sources from 0 to 2π in steps of $\pi/2$. Due to having two forward and two adjoint sources, this led to 16 total combinations of phases for a given seed. To reduce computation time, only the best 4 phase combinations were selected to finish after 30 generations. While not pursued here, the same type of approach can be taken for accounting for the polarization of the sources.

3.3.4 Designing the nanoscale light extractor

To find the ideal height, we first ran a series of a series of 2D optimizations, sweeping over a range of heights (Figure 3.2A,B). Like in the later 3D simulations, the 2D simulations were set up with the dipole embedded 10 nm into diamond, and we used two orthogonal dipole orientations in a plane perpendicular to the NV axis. We swept the height from 50 to 600 nm in steps of 50 nm and ran 5 optimization runs for each height. Figure 3.2(B) shows the average of our FoM for each height, with the oscillations reminiscent of Fabry-Perot fringes. Based on these results, we decided to run the full 3D optimization of the NLE for a height of 300 nm. This thickness is a common device-layer thickness in silicon-on-insulator (SOI) technology [143] and we note that SOI lends itself well to fabricating this structure as crystalline silicon device layers can be transferred to diamond using membrane-transfer techniques [144].

The full 3D optimization run can be seen in Figure 3.2(C). The FoM increases as the optimization progresses, with occasional dips due to the implementation of the secondary blurring function. The optimization terminates when the device is sufficiently binarized, such that the structure is entirely comprised of air and silicon, and the FoM is no longer improving significantly.

3.3.5 Results

The structure of our optimized NLE is shown in Figure 3.3(A), with the extraction efficiency in Figure 3.3(B). Though the NLE was optimized for an NV depth of 10 nm, we also show the results for depths of 5 and 15 nm but with the same NLE structure. The extraction efficiency increases for NVs closer to the structure due to increased coupling with the NLE, which leads to larger Purcell enhancement [145]. For an NV depth of 10 nm, the Purcell enhancement of our device is ~3, averaged across the emission spectrum. Here, we calculate the Purcell enhancement by dividing the power emitted by the dipole source in the presence of the interface and the NLE by the power emitted by the dipole in homogeneous diamond (no interface).

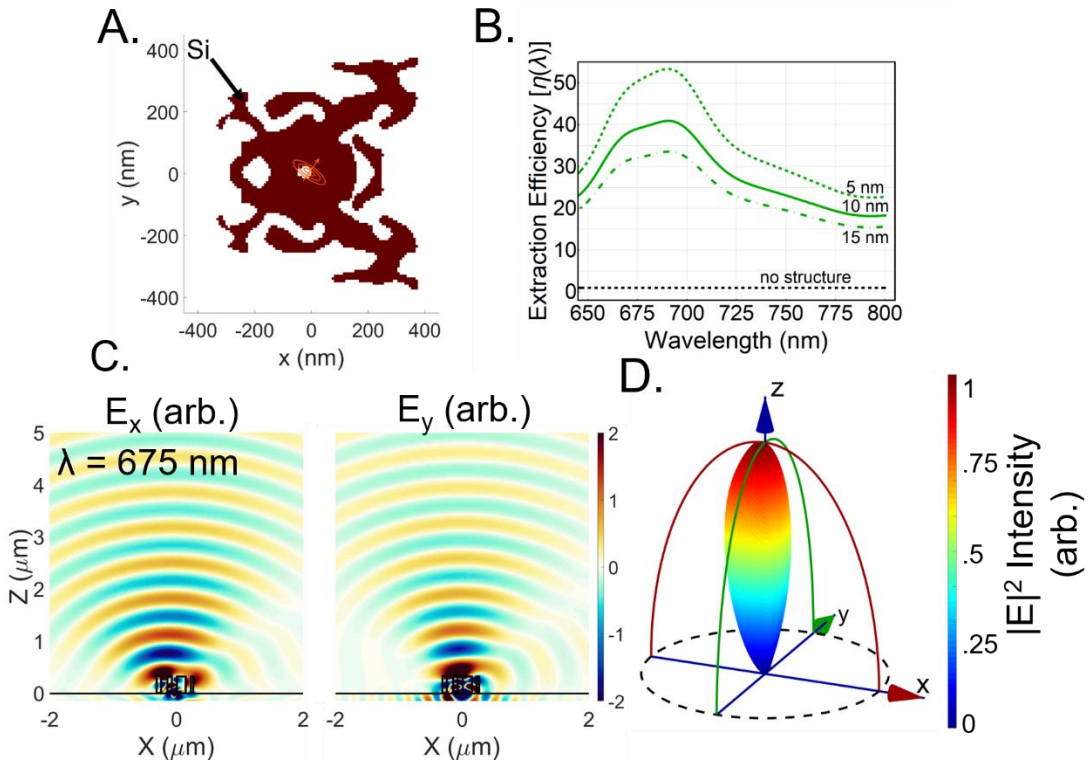

Figure 3.3. **A.** Top-down view of the final NLE optimized for an NV depth of 10 nm. **B.** Extraction efficiency $[\eta(\lambda)]$ of the NLE for NV depths of 5, 10, and 15 nm below the diamond/air interface. **C,D.** Snapshot of the near and meso-fields of the emitted electric field from (left) a dipole in the X-Z plane, and (right) a dipole in the Y direction. **D**. Intensity far-field averaged over the two dipole orientations. In C,D, the plots are at a wavelength of 675 nm. The bulk of the beamed power fits within a 60° cone in the far field.

Figure 3.3(C, D) shows the near- and far-field radiation patterns. In Figure 3.4 we show a comparison between the near fields with the NLE present compared to those without the NLE. The snapshot was taken along a slice in XZ plane at y=0. It can be seen that the strength of field above the structure increases dramatically with the addition of the NLE.

The NLE can shape the emitted fields from the dipole into a beam that is approximately Gaussian. The beam angle is slightly offset from the normal (~7.5°) but the bulk of the beamed power fits within a +/- 30° cone. The fields in Figure 3.3(C,D) are for a wavelength of 675 nm, but the beaming persists across the NV spectrum and the peak emission angle does not deviate by more than $\pm 5°$. A cross-sectional farfield plot taken at select frequencies can be found in Figure

3.5. For an objective with numerical aperture of 0.75, the collection efficiency, defined as the fraction of emitted optical power that can be collected, is around 40% in the 635 – 670 nm range, and above 25% across the entire 635 – 800 nm range. Figure 3.6 shows a full breakdown of the emitted power of the NV with the NLE present including absorption and collection efficiencies for various NA lenses.

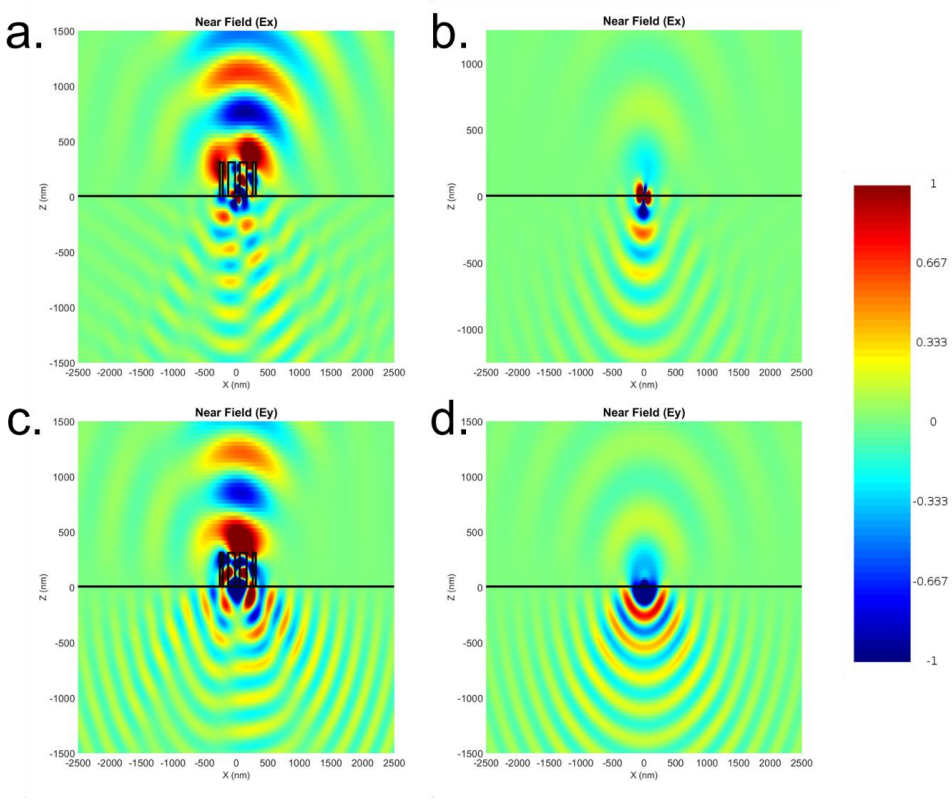

Figure 3.4 - A snapshot of the electric fields in the vicinity of the NV center taken along a slice in the XZ plane at y = 0 (the center of the device and the y-coordinate of the NV), at a wavelength of 675 nm. Here we directly compare the X components of the electric field for the dipole in the XZ plane (a) with and (b) without the NLE, and the Y components of the electric field for the dipole in the XY plane (c) with and (d) without the NLE. The NLE is outlined in black for reference.

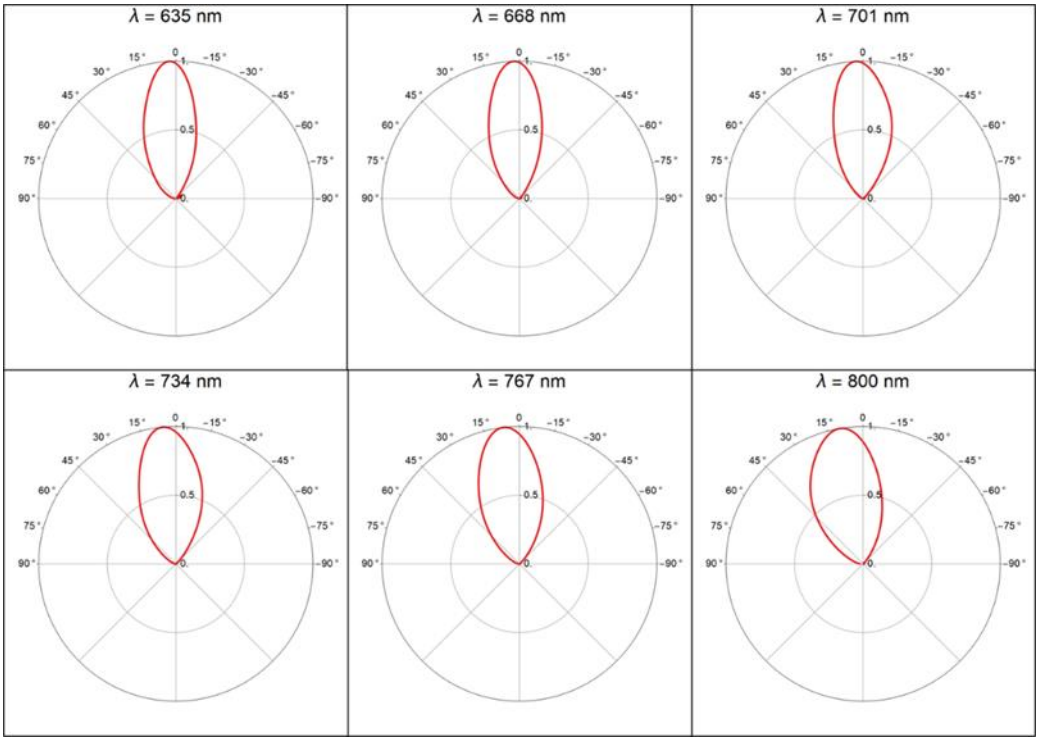

Figure 3.5 - To ensure the beaming capabilities of our device are maintained over the full 635 – 800 nm spectrum we simulated and calculated the far-field |E|2 intensities for multiple other wavelengths besides the one presented in Figure 3.3. Here we plot slices of the field profiles at an azimuthal angle of 0° and along a polar angle of ±90° for wavelengths 635, 668, 701, 734, 767, and 800 nm. The far-fields are normalized to the peak intensity over the full hemispherical projection (-90°≤θ≤90° and 0≤ϕ≤360°) so the peaks in some plots occur at another azimuthal angle. The peak intensity always stays within ±15° of the normal with only minor variations between the different wavelengths.

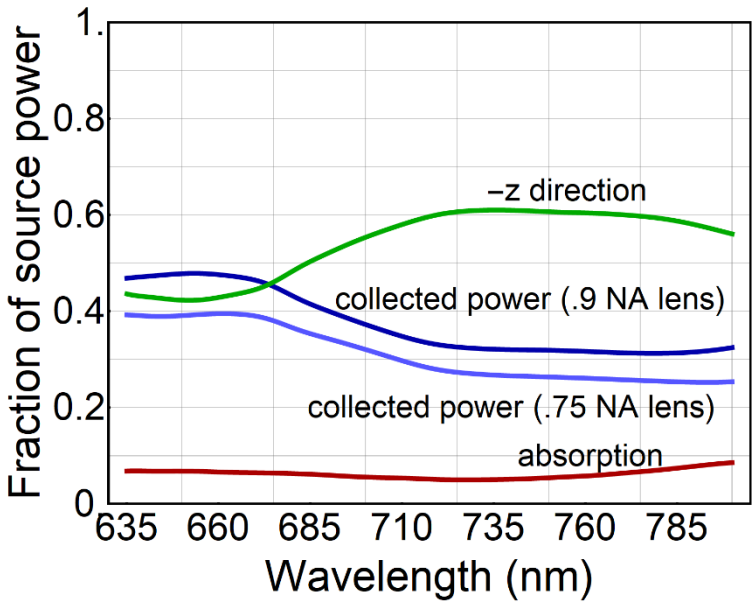

Figure 3.6 - Breakdown of light emitted by the NV center in the presence of the NLE. The collection efficiency is defined as the fraction of emitted power can be collected by a lens with a specific NA. The power recorded in the -z direction is defined as the power flow downwards through a plane placed 100 nm below the NV center and diamond interface. The absorption is calculated as 1 – (power going towards +z) – (power going towards -z). The power flow is normalized to the full, Purcell-enhanced output power of the dipole source and averaged between the two dipole orientations.

Due to the broadband spectrum incorporated into the optimization, our devices display considerable tolerance to various fabrication defects as well as robustness to alignment errors (Figure 3.7). Though our optimized NLE structure was based on a fixed NV depth of 10 nm, we also simulated its performance for a variety of depths [Figure 3.7(A)]. The NLE performs better as the dipole gets closer to the surface due to enhanced near-field coupling. The performance falls off by a factor of 2 at a depth of 40 nm, yet still maintains enhancements of about 15 times the emission of an NV with no NLE. Even down to a dipole depth of 300 nm, the NLE is able to increase the output of the NV by a factor of 3. Note that for NVs at depths substantially different than 10 nm, a more-effective design can very likely be found using the optimization method described above.

The NLE shows minimal performance loss from errors in rotational alignment in the range of -20° to +20° [Figure 3.7(B)]. The full 360° rotational plot can be found in Figure 3.8. We also tested the case where the NLE was offset by some amount from the central dipole position [Figure 3.7(D)] and found that the FoM remains above 25 for X offsets of +/-30 nm and Y offsets of +/-40 nm.

In practice, one can reasonably expect the fabrication process to cause deviations in edge locations, e.g., due to proximity effects in lithography [146]. Figure 3.7(C) shows the extraction efficiency of our eroded, optimized, and dilated structures with their index profiles shown in Figure

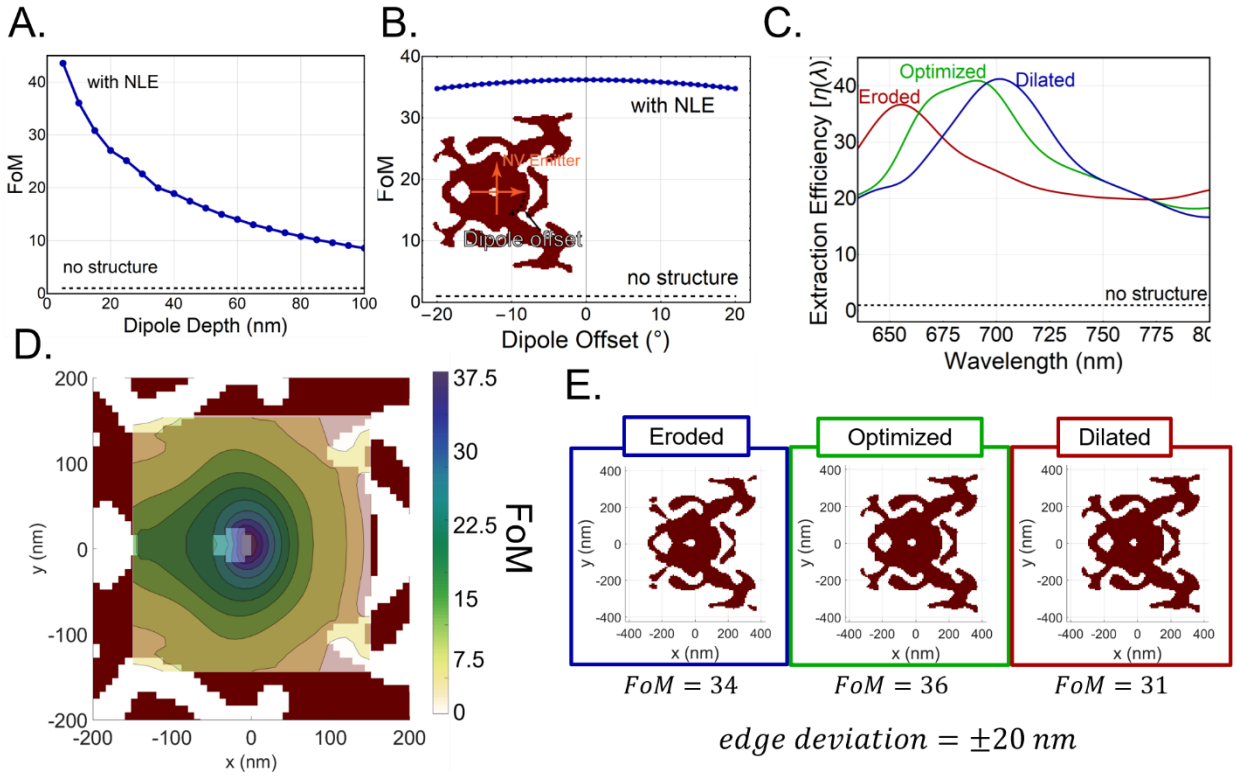

Figure 3.7 - A. The NLE maintains good performance through a range of depths, with increasing FoM for NVs closer to the surface. B. FoM dependence on the NV emitter angle emulating angular alignment errors of the NLE. C. Demonstration of the fabrication robustness of the optimized device for an NV depth of 10 nm. Eroded and dilated structures are based on the optimized structure with edge deviation of +/- 20 nm to represent fabrication under/over-etching, respectively. Due to the broadband nature of our optimization, the NLE shows strong tolerance to fabrication errors. D. Tolerance of the NLE to lateral offsets of the NV center. The FoM remains above 25 for X offsets of +/- 30 nm and Y offsets of +/- 40 nm. e. The geometries of the eroded, optimized, and dilated devices simulated in c. The edge deviation refers to how far the edges shifted inward for the eroded or outward for the dilated cases.

3.7(E). We calculated the deviated structures by applying a gaussian blurring filter across the optimized pattern, and then selecting cutoff points of the blurred edges to yield new binarized structures. Here, we selected the blur and cutoff to yield edge deviations of ±20 nm. The NLE maintains good performance across the spectrum despite erosion or dilation [Figure 3.2(B)].

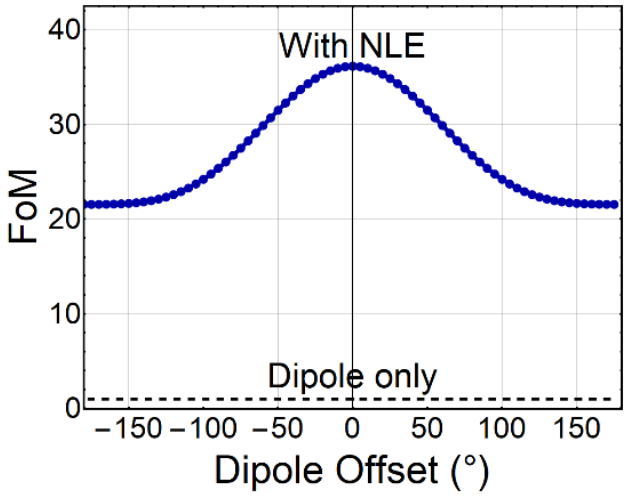

Figure 3.8 - The full rotational robustness of the NLE demonstrated by rotating the NV emitter a full 360° around the Z axis.

It is instructive to compare the FoM of our NLE (~35 for NVs at a depth of 10 nm) with some existing structures in the literature designed for broadband vertical outcoupling of light from a diamond slab, with the understanding that the FoM combines the distinct collection-enhancement and Purcell mechanisms. Dielectric structures comprising etched diamond typically do not provide much Purcell enhancement and include bullseye gratings [123], vertical nanowires [147], solid-immersion metalenses [121], and parabolic reflectors [115]; we estimate that these structures have calculated FoMs of approximately 10 to 20 due entirely to collection enhancement. Resonant geometries proposed thus far have primarily relied on plasmonic enhancement in structures such as metallic cavities [116], [148] and gratings [149], with FoMs of up to ~35. The theoretical FoM of resonant metal-dielectric structures in ref. [150] can be well above 100 due to large Purcell

enhancement, but the implementation would require elaborate fabrication including filling etched diamond apertures with metal, and the small mode volume limits the positioning tolerance of the NV center to within ~5 nm of the field maximum. Finally, note that unlike all of the aforementioned designs, our NLE does not require etching of the diamond, and can be fabricated by using Si membrane transfer techniques [144], [151] or direct CVD growth of Si on a diamond substrate. We do note that the exact effects of the NLE on the NV-center spin characteristics (e.g., coherence) due to the proximity of the NV to the NLE surface and material are unknown at this time. These effects are difficult to predict and are beyond the scope of this article, warranting future experimental investigations. However, our technique is likely to be less invasive than approaches that require etching the diamond near the NV, which inevitably introduces damage and additional defects.

We note that while the present manuscript was undergoing peer review, a preprint by Chakravarthi et al describing a similar approach to extracting light for NV centers, but for quantum information applications, was posted on arXiv [152].

3.3.6 Conclusion

We presented a nanoscale light extractor (NLE) designed using adjoint-optimization methods with time-domain simulations to enhance the broadband emission of NV centers in diamond. Our design not only enhances the fluorescence of the NV centers but also demonstrates exceptional beam-shaping qualities, enabling efficient collection with a low-NA lens in free space. Even given reasonable uncertainty in NV center localization and errors in device fabrication, simulations show the NLE maintains a high extraction efficiency. Our results suggest that such robustness to positioning and fabrication errors can be automatically achieved when optimizing

for a broadband figure of merit. The NLE can be fabricated with conventional electron-beam lithography techniques without etching of the diamond surface. Our approach can easily be extended to other color centers in diamond, as well as material systems where increased light extraction from defect centers is desired, such as silicon carbide or hexagonal boron nitride [153], [154].

3.4 Nanoscale light extractors for NV centers in (111) diamond

This work was inspired by and benefitted greatly from numerous discussions with Professor Shimon Kolkowitz and Professor Jennifer Choy at UW-Madison. The discussion on adjoint sources was inspired through conversations with Professor Zongfu Yu at UW-Madison.

3.4.1 Introduction

In the previous section (3.3) we used adjoint optimization to specifically design a photonic structure for (100) oriented diamond. In this section we turn our attention to looking at NV centers in (111) oriented diamond crystals. There are a few advantages to this orientation (described below) but here we also expand our adjoint optimization capabilities to alternative FoMs that enable diverse functionality.

Most commonly, bulk single-crystal diamond samples with nitrogen impurities are grown in the (100) crystal orientation whereas the primary alignment of the NV defect bond axis lies in the [111] direction[94], [155]–[157]. This leads to an oblique angle between the dipole emission direction of the NV center (orthogonal to the bond axis) and the diamond-air interface, further exacerbating the issue of total internal reflection. The NV defect embedded in (111) crystal diamond can be visualized in Figure 3.9(A). The emission of the NV centers can be maximized by

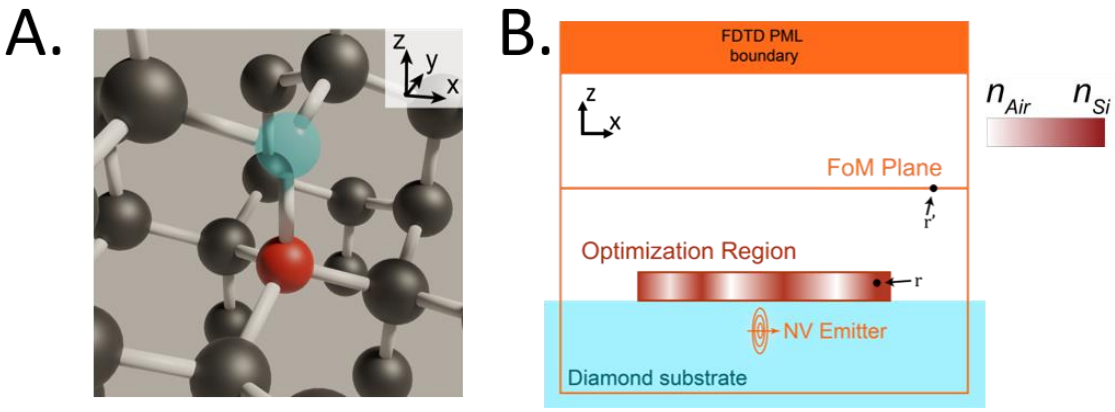

Figure 3.9 – A. Ball and stick representation of the NV defect in (111) diamond. The nitrogen (red) and vacancy (blue) bond in the Z direction leading to an emission dipole oriented in the XY plane. B. Schematic of the standard adjoint optimization region represented in 2D. The FoM plane is used to evaluate the performance of the NLE but also serves as the location to calculate the adjoint fields. The optimization region starts as a smooth gradient of refractive index varying from Si to air which eventually binarizes as the optimization completes.

growing substrates oriented in the (111) crystal orientation such that the NV emission dipole axes are aligned parallel to the air-diamond interface[158]–[161]. Traditionally, this crystallographic orientation was challenging to grow; however, there have been significant recent advancements in the growth of (111) diamond via plasma-assisted CVD that has led to high-quality (111) diamond crystals and well-aligned NV defects [157]–[160], [162]–[165].

Even with advances in growth techniques to enhance NV emission, the overall percentage of emitted light remains low. A shallowly implanted NV center ($\sim 20\ nm$ deep), modeled as a dipole emitter oriented parallel to the diamond-air interface, transmits only ~5% of its emitted

power out of the diamond into free space. Therefore, further photonic engineering solutions are needed if single NV sensors are to be made more widespread.

We seek to address this problem using adjoint optimization to design silicon-based photonic structures to enhance the light output of NV centers. Section 3.3 introduced the concept of using adjoint optimization to design a single nanoscale light extractor (NLE) for NVs embedded in (100) oriented crystalline diamond[166]. In this section we aim to expand on our previous method by establishing a range of NLEs with varying functions enabled through adjoint optimization. Seeing as (111) diamond may be more useful for both sensing and quantum information processing due to the advantageous NV alignment, we seek to provide structures that can 1) substantially increase the available SNR, 2) provide increased functionality by either directly controlling the emission direction or suppressing the excitation fields in a certain region of space, and 3) be easily used in both experimental setups as well as any NV-based sensing platforms.

3.4.2 Adjoint Optimization

To design these structures we turned to adjoint optimization following and improving on our approach outlined in section 3.3.3 [166]. In this section, we use a variety of adjoint simulations depending on the desired function of the NLE; the FoM and accompanying adjoint sources are described in each section. The general simulation setup, represented in 2D, can be seen in Figure 3.9(B). The optimization region consists of a spatial distribution of index of refraction that at each location varies from air (n_{air} =1) and crystalline silicon ($n_{\mathrm{Si}}(\lambda)$, from ref. [138]). The FoM plane is situated above the optimization region and is used to set the adjoint sources (details for each case below) along with evaluating the performance of the NLEs. In each case, the forward

simulations use a dipole source oriented parallel to the diamond-air interface and set to a depth of 20 nm.

In each case, we characterize the performance of our structure by normalizing the FoM to the performance of an NV center in the presence of the diamond/air interface alone (*i.e.,* no NLE structure present). In both cases we normalize to the amount of light collected into collection optics with an NA of 0.5. In our previous publication, we characterized and compared our NLE to the performance of other published design approaches and, while the FoMs in this case are slightly different, those comparisons are still useful for gauging the relative performance of each device. The comparable FoMs of interest are estimated to range from 10 to 35 [167]–[173].

To narrow down one of the free structural parameters, the thickness of the NLE, we performed a 2D height sweep similar to the procedure in 3.3.4 but using orthogonal dipoles laying in the XY plane. The results are show in Figure 3.10. Similar to Figure 3.2B, the peak FoM occurs at a thickness of approximately 300 nm, with additional “fringes” of locally maximum FoM occurring every 250 nm. Based on the results of this sweep, we choose a height of 300 nm for all NLE designs.

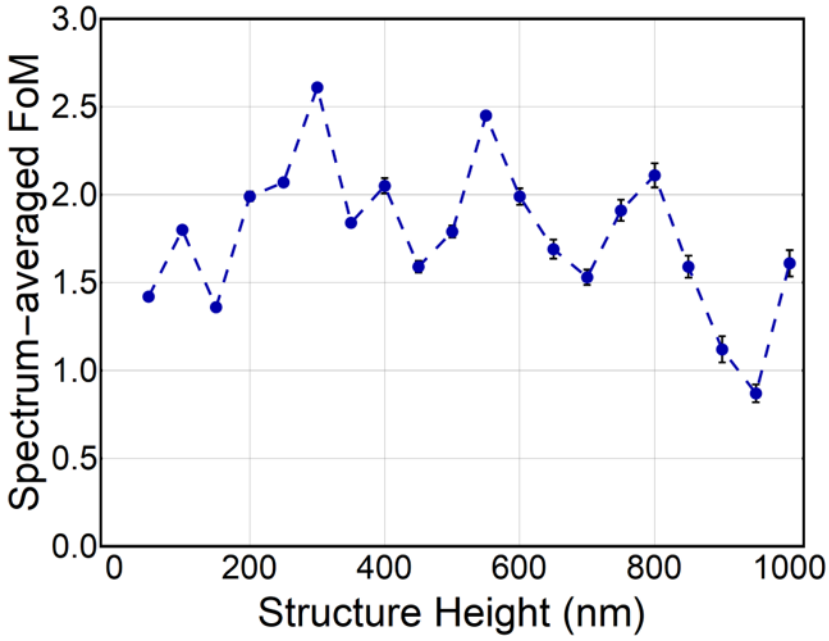

Figure 3.10 – Results from running a series of 2D optimizations to find the ideal NLE height. The height was swept from 50 to 1000 nm in steps of 50 nm with 8 optimization runs for each height. Error bars are present for all points, but the deviation was very small for smaller heights due to some of the runs converging to the same result. The peak FoM occurred at a height of 300 nm and fringes of high performance occur every 250 nm.

3.4.3 Stock NLE and Central Cavity NLE

The first structure we designed used the same basic optimization procedure as described in [166]. Our target FoM was to beam the NV emission into a $\pm 30°$ cone into the far field, in the direction normal to the diamond interface. This configuration makes it easy to collect the majority of emitted light using a modest-NA objective ($NA = .5$). Therefore, our adjoint FoM becomes the overlap integral of the output optical field with the desired free-space optical profile:

$$F(E_{dip}, H_{dip}) = \frac{\left(\int E_{dip} \times \bar{H}_{tgt} \cdot dS + \int \bar{E}_{tgt} \times H_{dip} \cdot dS\right)^2}{Re\left[\int E_{tgt} \times \bar{H}_{tgt} \cdot dS\right]} \tag{3-8}$$

Here E_{dip} and H_{dip} are the forward fields from our simulation which uses a dipole source and E_{tgt} and H_{tgt} are the target field profiles ($\bar{E}$ and $\bar{H}$ denoting complex conjugates). In our case, the target field profiles correspond to a Gaussian beam traveling in the direction normal to the diamond interface. Once our FoM is defined, dependent on our E_{dip} and H_{dip} fields, we can follow the derivation described in [89], [174] to obtain our adjoint source. The function determining our adjoint source is defined by:

$$E_{adj} = \int_S \left[G_{ji}^{EP}(r', r) \frac{\delta f}{\delta E_i(r)} - G_{ji}^{EM}(r', r) \frac{\delta f}{\delta H_i(r)} \right] dr \tag{3-9}$$

In this equation, G_{ji}^{EP} and G_{ji}^{EM} are the Green's functions mapping the fields due an electric and magnetic dipole respectively at r to r'. Here we define the location r' to be a location in our optimization region while r defines a location on our merit function surface. The indices i and j define the cartesian field components. This says that to define our adjoint source we need to populate the surface S with electric and magnetic dipoles that have a complex amplitude defined by $\delta F/\delta E_i\,(r)$ and $-(1/\mu_0) \cdot \delta F/\delta H_i\,(r)$, respectively (the factor of $1/\mu_0$ comes from the normalization of the Green's function, see Appendix A in [89]). Taking the appropriate derivatives of our merit function, we arrive at an adjoint equation of:

$$E_{adj} = \tilde{C} \int_S \left[G_{ji}^{EP}(r', r) \cdot \left(\bar{H}_{tgt} \times n \right) + G_{ji}^{EM}(r', r) \cdot \left(\bar{E}_{tgt} \times n \right) \right] \mathrm{d}r' \tag{3-10}$$

The dipoles described by the Green's functions along the surface form a Gaussian propagating downwards in the surface normal direction with a complex amplitude coefficient of $\tilde{C}$ defined by:

$$\tilde{C} = 2\frac{\int E_{dip} \times \bar{H}_{tgt} \cdot \mathrm{d}S + \int \bar{E}_{tgt} \times H_{dip} \cdot \mathrm{d}S}{Re\left[\int E_{tgt} \times \bar{H}_{tgt} \cdot \mathrm{d}S\right]} \tag{3-11}$$

Thus, to optimize an NLE that seeks to overlap the emitted dipole fields with our desired mode, we must calculate the adjoint gradient using a backwards-propagating Gaussian beam with the phase defined by $\tilde{C}$. It is important to note here that, while the phase of the adjoint source is defined relative to the desired mode [due to presence of E_{tgt} and H_{tgt} in Eq. (3-11)], there remains an unknown as to what desired mode phase gives the best performance. Therefore,, we performed multiple optimization branches using varying target-mode phases from 0 to 2π for a given seed. More details on this approach can be found in the section 3.4.5.

Using the procedure described above, we optimized a stock NLE (only binarization and vertical invariance constraints) and an NLE with a forced central opening with a radius of 25 nm. We are terming this structure the enhanced-proximity (EP) NLE as the central opening is useful for flow-chamber sensing configurations, allowing the target chemical to diffuse down to the diamond surface within close proximity of the NV center. The stock NLE does not always result in a large enough hole for this diffusion to take place and thus we created the EP NLE. Figure 3.11(A) and (B) shows the resulting structures after optimization, with the respective far-field $|E|^2$ profiles be seen in (C) and (D). The emitted light is beamed almost entirely into a cone with a half-angle of 30°. Each structure obtained a satisfactory FoM with the stock NLE reaching a FoM of

24 and the EP NLE obtaining a FoM of 22. The extraction efficiency as a function of wavelength is plotted in Figure 3.11E.

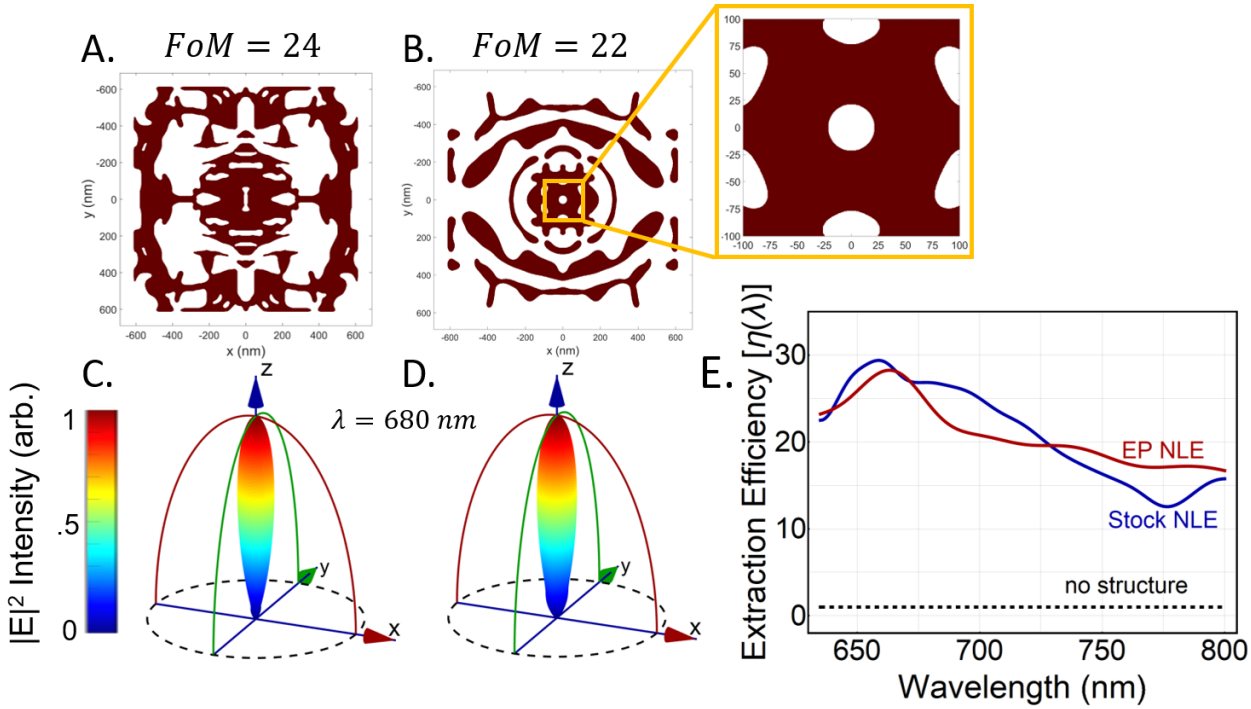

Figure 3.11 – A. & B. The refractive index profile of the standard NLE and NLE with a central cavity, respectively. The central cavity in B is designed to enable diffusion of an analyte to be detected to within close proximity of NV center while maintaining enhancement and beaming characteristics. C. & D. Far-field $|E|^2$ of the standard and cavity NLE, respectively, taken at the peak NV emission wavelength of 680 nm. E. Normalized extraction efficiency of each NLE. The spectra correspond to the amount of emitted light collected by a lens with $NA = .5$ and normalized to the collected emission of the dipole with no structure.

3.4.4 Dipole orientation dependent NLE

In the next example, we use the same adjoint FoM described in the previous section except that instead of using a Gaussian adjoint normal to the surface, we offset the angle of the Gaussian based on the orientation of the dipole moment of the NV. The resulting optimized NLE is able to split the emission of the photons based on which orientation the NV is in, which can be useful for quantum information processing applications [175]–[177]. The device put forward here has the benefit that it not only splits the emission of the NVs based on orientation, but it also acts on NVs

within a 60 x 60 nm region beneath it, enabling a single NLE to operate on two coupled NV centers as coherent NV center coupling has been shown on the order of 10 nm (we do note that the presence of the NLE will have some effect on the coupling, but that is beyond the scope of this section) [178]. Beyond polarization splitting, the NLE also retains the emission enhancement properties of the stock NLE. This NLE unites enhanced emission while replacing the need for external polarization splitting optics which can result in configurations that are more compact.

Because this type of NLE is likely most-useful for quantum information applications, we optimized the structure for operation only at the NV zero-phonon line (ZPL) (637 nm). Measurements of entanglement of NV emission are often performed at sub-10° K temperatures, to reduce the effect of phonon-side band transitions and spin mixing [179]–[181]. This causes the primary emission of the NVs to be at the ZPL and extends the spin relaxation time[177]. While our optimization was centered on the ZPL, we did employ broadband simulations in order to preserve fabrication robustness [141], [142], [166]. For room-temperature NVs, we simulate the full emission spectrum (635-800 nm) and weight the optimization gradient according to the normalized intensities of the spectrum. To maximize the performance at the ZPL, we center the simulation at a wavelength of 637 nm and weight the optimization gradients according to a Gaussian distribution defined in the frequency domain with a full-width at half-max of 13 nm.

The NV with a dipole moment oriented along the x-axis was targeted to be beamed at an angle $-25°$ from the normal, while the y-axis-oriented NV dipole was targeted to be beamed at $+25°$ to the normal. A schematic of the targeted performance can be seen in Figure 3.12A while the resulting index profile is represented in Figure 3.12(B). The beamed normalized $|E|^2$ farfields for the X and Y dipoles can be seen in (C) and (D) respectively. We defined a separate FoM for each beaming direction. This FoM is our standard FoM but with the collection cone being centered

at the target angle (25° for x-orientation, -25° for y-orientation). The cone half-angle is still 30° and normalized to the amount of light emitted by an isolated NV into that same direction. The FoM for the x-orientation was 6 and the FoM for the y-orientation was 11. While not as large as the stock NV, this device still amplifies the output of the NVs by an order of magnitude while beaming into the target direction. Figure 3.12(E) is a cross-section taken from (C) and (D) along the red hemisphere oriented in the XZ plane.

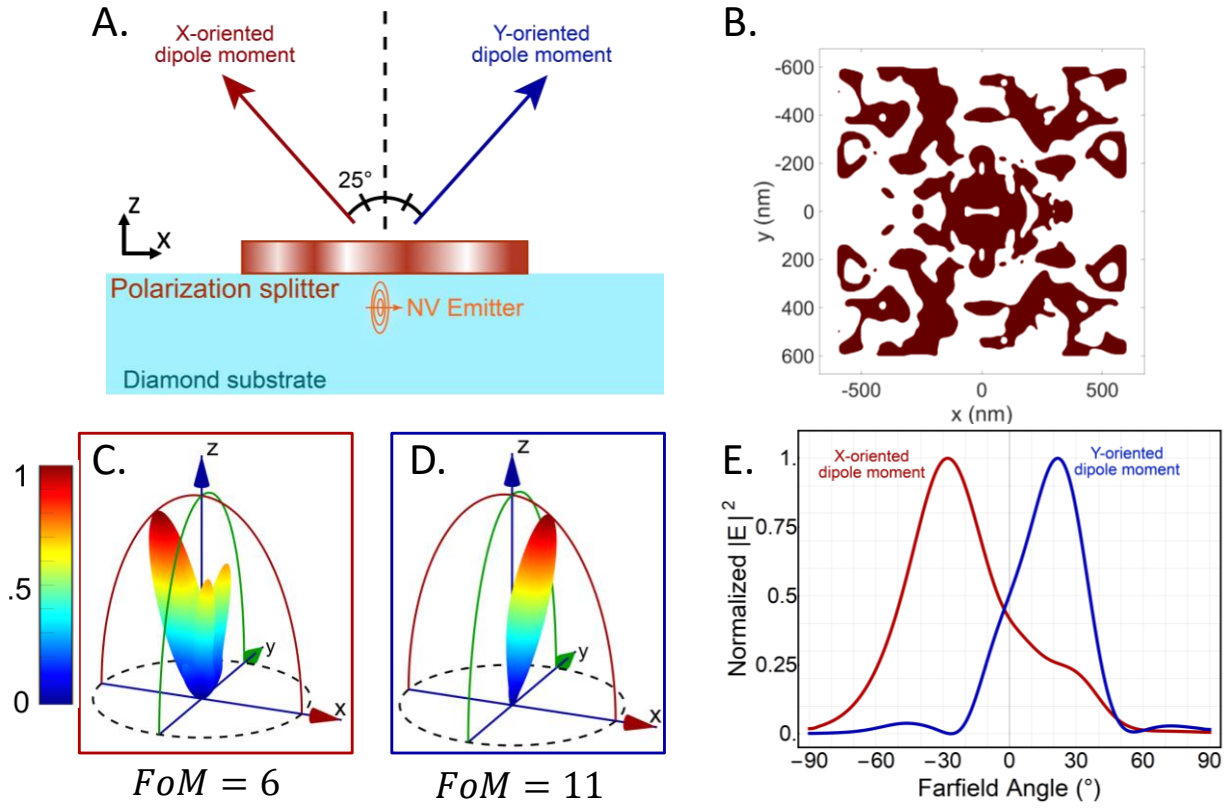

Figure 3.12 – A. Schematic outlining the targeted performance of the NLE. X-oriented dipole emission is sent at a $-25°$ angle with respect to the normal while the Y-oriented dipole emission is sent toward +25°. B. the resulting refractive index of the optimized device. C. and D. the farfield $|E|^2$ patterns for the x and y orientations, respectively, taken at 637 nm. E. a cross section of the farfield taken along the x-axis and sweeping along the hemispherical projection. The angle is measured off of the +Z axis.

There is some non-negligeable crosstalk between the target directions, however for some applications this can be compensated by using a polarizer on the respective detectors (although we do note that there can be some change in polarization as the light passes through the NLE though the effect is minimal here). The extinction ratio for the X-polarization is 6.2 dB at $\pm 25°$ while the extinction ratio for the Y-polarization is 28 dB.

3.4.5 Comments on the Phase Maximization Approach

In this section, we comment on the phase maximization approach we take to maximize the emitted NV power into a specific free-space mode. This approach was used in the stock NLE, the EP NLE and the polarization splitting NLE. It was noted in Section 3.4.3 in Eq. (3-11) that, while the adjoint phase is determined by the complex coefficient $\tilde{C}$, the coefficient itself is dependent on the phase of the forward fields as well as the phase of the target mode.

In the following Figure 3.13 we plot the first 50 generations of a 3D optimization of an NLE for (111) diamond for a single dipole orientation (electric field along the x axis). In these optimization runs, we use a Gaussian adjoint source set to the phase prescribed by Eq. (3-11). Each run used the same starting seed (*i.e.*, the random, slowly varying, and continuous starting refractive index distribution), but the phase for the target mode was set to a fixed value, stepped between 0 and $3\pi/2$. It is important to note that in these optimization runs, the adjoint source phase does vary from iteration to iteration due to the dependence of $\tilde{C}$ on the forward fields E_{dip} and H_{dip}, we only fix the phase of E_{tgt} and H_{tgt}. Thus, while a fixed target mode phase is chosen, the actual adjoint source is free to vary based on the forward fields.

From the figure, it can be seen that the original seed reaches different figures of merit depending on what phase is chosen. Thus, it cannot be known from the outset of the optimization what phase would be best to optimize towards.

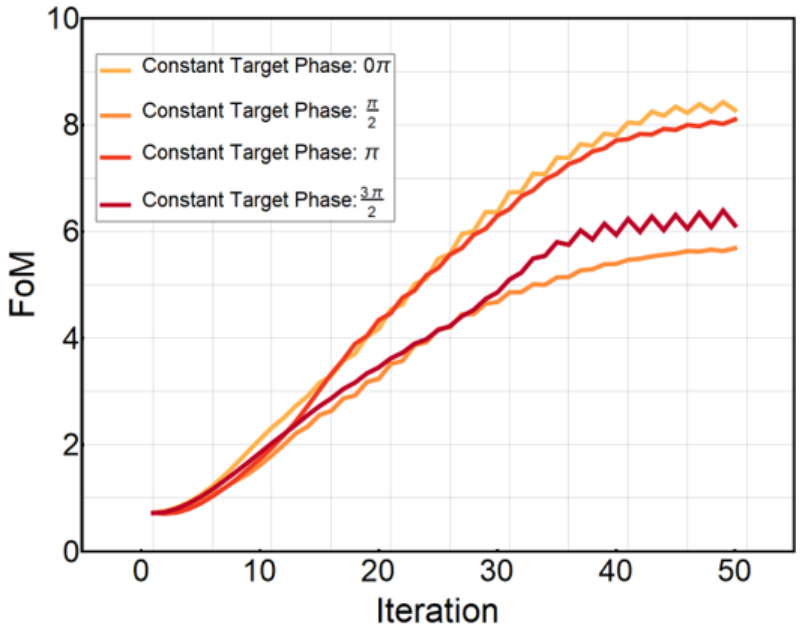

Figure 3.13 - FoM vs. iteration for the first 50 iterations for varying target mode phases optimized using the adjoint FoM in Eq. (3-11). The target FoM is the overlap of the dipole emission with a Gaussian beam propagating normal to the surface of the diamond. The optimization was performed using only one polarization of the dipole as a pedagogic example. The lower FoMs here compared to the main text are primarily due to smaller size (only 750 x 750 nm compared to 1220 x 1220 nm) and larger voxel resolution (25 nm compared to 6 nm). The same starting seed is used in each case.

One way that we attempted to overcome this limitation was to use an adaptive phase approach. In this approach, the target mode phase is set to the phase of the emitted light at the adjoint source plane. In this way, the optimization always steps toward a phase that is preferred by the current iteration of the NLE. The target phase is calculated by taking a weighted average of the phases at the source-plane monitor (weighted by the Gaussian intensity distribution).

$$\phi_{tgt} = \frac{\sum_{ij} w_{ij} \cdot \arg\left(E_{ij}^{fwd}\right)}{\sum_{ij} w_{ij}} \tag{3-12}$$

where w_{ij} is the gaussian weight at point (i,j) on the source-plane monitor and E_{ij}^{fwd} is the forward field recorded at the source-plane monitor at point (i,j).

Figure 3.14 shows how this approach compares using the same seed as the previous example.

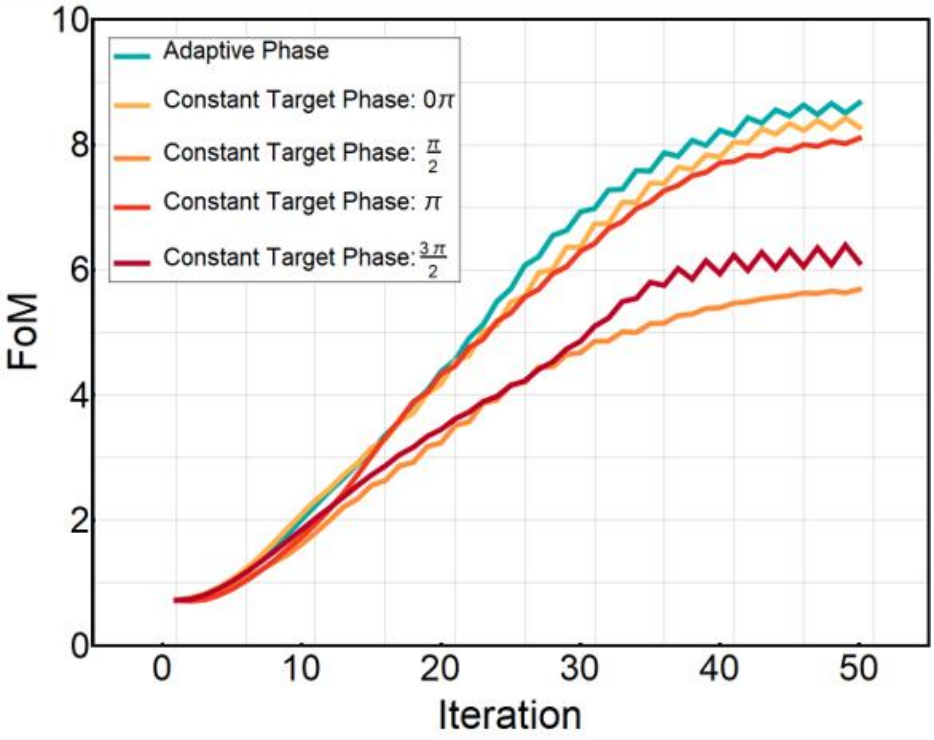

Figure 3.14 - FoM vs. iteration for the first 50 iterations of the adaptive phase approach compared to using a fixed target mode phase. The adjoint source phase was still set according to Eq. (3-11) but the target mode phase was determined from the forward simulation fields. The phase of forward fields at the FoM monitor were recorded and a weighted average was taken to determine the target mode phase. The average was weighted with a normalized Gaussian intensity distribution. Each run uses the same starting seed.

Here it can be seen that the adaptive phase approach is able to match and beat the performance of the best of the fixed phases. This also comes with the advantage that fewer

simulations are required as four concurrent optimizations are not needed, only one. The performance of the adaptive phase method consistently tied or beat the highest performing fixed phase optimization branch on multiple different starting seeds.

However, we still decided to forgo this approach in favor of using a fixed-phase adjoint source. In this case, $\tilde{C}$ is set to a constant value every iteration. Figure 3.15 shows the comparison between the adaptive phase approach and the fixed adjoint phase approach that we use in the main text. This comparison is again using the same starting seed for both optimization runs. Our phase-max approach uses a constant Gaussian adjoint phase each generation. In this case, the constant

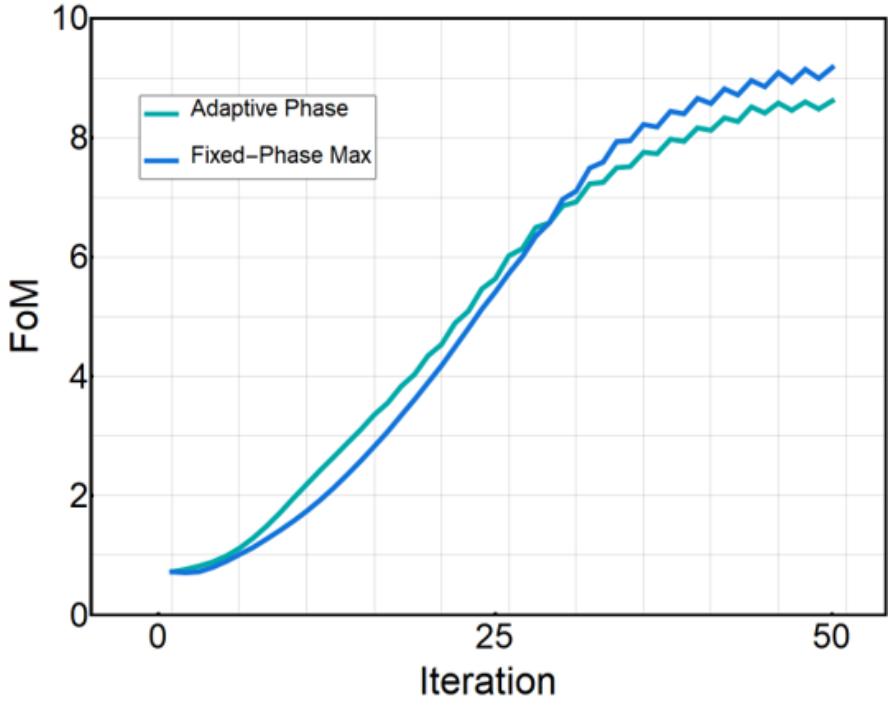

Figure 3.15 - FoM vs. Iteration for the first 50 iterations of the adaptive phase approach compared to our currently employed fixed-phase approach. In the fixed-phase approach, the phase of the adjoint source does not change for subsequent iterations but is fixed to a value between 0 and $3\pi/2$. The interval is broken down into 4 steps with the best performing phase being isolated and optimized to completion. The best phase for the given starting seed is shown.

phase was set to between 0 and $3\pi/2$ in steps of $\pi/2$ and the best performing phase is chosen to carry out the optimization. In the figure, only the best performing phase is plotted against the

adaptive phase method. It can be seen here that the constant phase source indeed outperforms even the adaptive phase method. This was also consistently seen using other seeds.

We believe this is due to the gradient-ascent nature of the adjoint optimization approach. Using the adaptive phase approach allows the algorithm to ascend to the local maximum of the FoM function around the starting position quite effectively. However, there is no way to escape that local max and even using the varying target mode phase, it is still unable to escape as the varying adjoint phase due to the forward fields keeps the optimization locked in that max. It can be seen in Figure 3.13 and Figure 3.14 that the FoM never decreases. However, when the fixed-phase maximization approach is used, the FoM does sometimes decrease in the early iterations and then, in later iterations, it can surpass the adaptive phase FoM. This early decrease in the FoM is a way of escaping a local max and moving towards a FoM that is larger and outside the initial search area.

While the fixed-phase maximization approach will consistently out-perform the adaptive phase approach, we find both approaches useful. The adaptive phase approach is particularly useful when computational resources are limited and it is not feasible to carry out N optimization runs for each mode that is targeted. Even only two targeted modes broken down into four phase steps requires 16 optimization branches which approaches the limit of feasibility; any smaller step size or more modes would quickly make this technique implausible. Therefore, the adaptive phase approach would be preferred as it can reach FoMs within $\sim 5-10\%$ of the fixed-phase method using significantly fewer simulations as only one optimization branch is required. However, for our case with only 16 optimization branches and because we have significant computational resources at our disposal (server with 128 CPU cores and 256 GB RAM), we were able to use the phase maximization approach here and in our previous work [166].

Finally, we note that this analysis is only applicable to the adjoint FoM of overlapping modes. This discussion has no bearing on the transmission maximizing FoM as there is no target mode phase present. This is explained in further detail in section 3.4.6 below.

3.4.6 Excitation Field Suppression

The final class of NLE presented here uses a FoM defined by the total power transmission through a surface above the NLE. This particular FoM is useful for a couple experimental configurations. First, if the setup can easily incorporate a high NA oil objective to collect a large angle of emitted light, this FoM would be preferable to design an NLE as more light can be extracted than using a mode overlap with a Gaussian beam. Secondly, and the particular configuration we are focusing on here, is that this FoM can be used to reduce the intensity of the NV pumping field in a certain region. In molecular sensing applications, the interaction of the excitation laser with the sample can cause degradation of the sample or artifacts to be observed in the emission spectrum measurement or, due to fluorescence of the sample of interest, can reduce the SNR [182]–[184]. Also, NV emission measurements often require optical components to reject or minimize the contributions from the excitation fields [177], [185].

To use adjoint optimization to minimize the transmitted field of the excitation laser, we first need to derive our adjoint source based on our FoM function, F. In this case, we seek to minimize the power flow through a surface, S, placed above the NLE and thus we maximize the negative of the power flow through that surface. The following is adapted to our application from Owen Miller in [89]. A diagram of the setup can be seen in Figure 3.16(A). Our merit function F becomes:

$$F(E_{Ex}, H_{Ex}) = -\frac{1}{2}\int_S Re(E_{Ex}(r) \times \bar{H}_{Ex}(r)) \cdot \hat{\mathbf{n}}\, \mathrm{d}S \tag{3-13}$$

with E_{Ex} and H_{Ex} being the excitation laser fields. Again, following the derivation described in Eq. (3-9), we first find the derivative of f (the statement in the integrand) with respect to the excitation fields:

$$\frac{\delta f}{\delta E(r)} = -\frac{\bar{H}_{Ex}(r) \times \hat{\mathbf{n}}}{2} \tag{3-14}$$

$$\frac{\delta f}{\delta H(r)} = -\frac{\bar{E}_{Ex}(r) \times \hat{\mathbf{n}}}{2} \tag{3-15}$$

We then insert Eq. (3-14) and Eq. (3-15) into Eq. (3-9). This gives us:

$$E_{adj} = \frac{1}{2}\int_S \left[G_{ji}^{EP}(r', r) \cdot (\bar{H}_{Ex}(r) \times \hat{\mathbf{n}}) + G_{ji}^{EM}(r', r) \cdot (\bar{E}_{Ex}(r) \times \hat{\mathbf{n}})\right] \mathrm{d}r' \tag{3-16}$$

which tells us is that our adjoint source for minimizing the excitation field becomes electric and magnetic dipoles that re-radiate the transmitted fields back toward the optimization region. This is specifically accomplished in Lumerical FDTD [39] by recording the transmitted fields from the forward dipole simulation, through a monitor placed above the optimization region, performing the operations in (3-14) and (3-15), and using those fields as a custom source object to inject them back into optimization region to obtain the adjoint field in our optimization region. We are able to

calculate a gradient for this figure of merit (G_{Ex}) which we then average with our gaussian enhancing gradient (G_{Gauss}) with a weight of $(.25 \cdot G_{Ex} + .75 G_{Gauss})/2$.

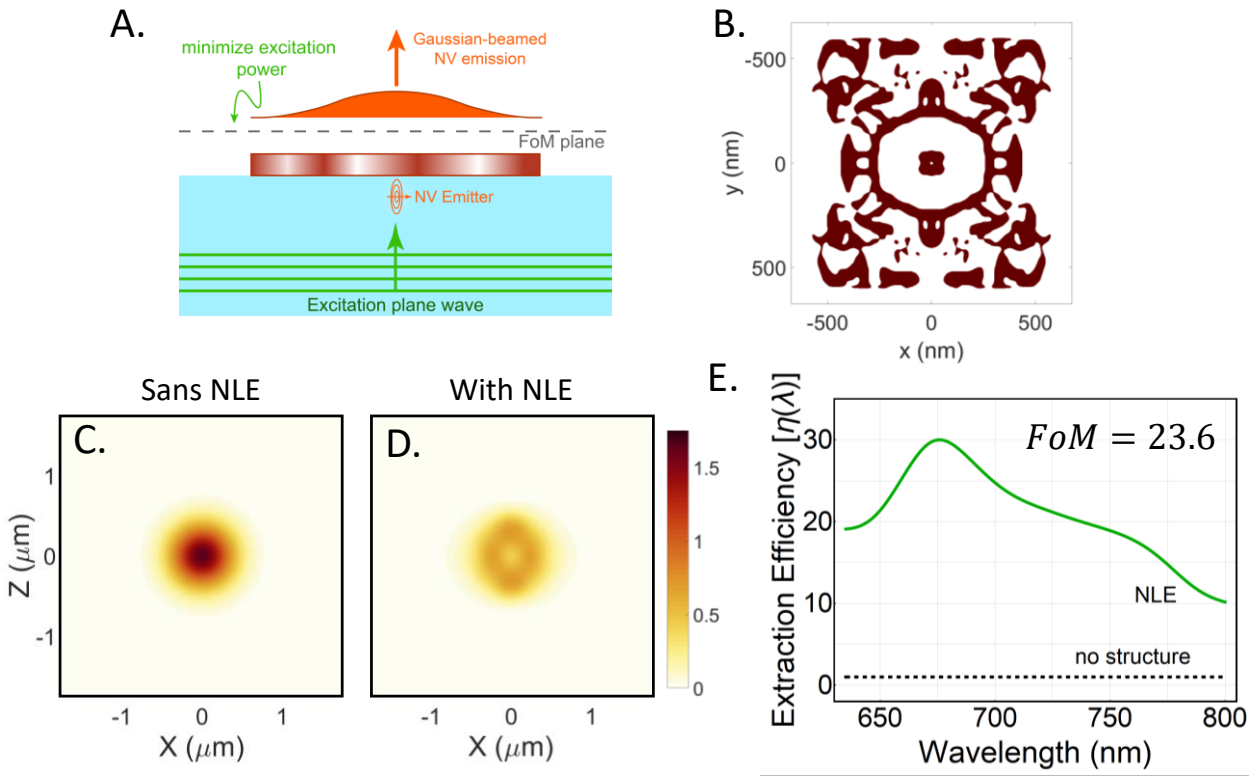

Figure 3.16 – A. schematic of the target adjoint optimization scheme. The adjoint is setup to minimize the power transmission of the excitation plane wave through the FoM plane while also incorporating the standard adjoint of enhancing and beaming the NV emission in the previous sections. B. Resulting index profile of the optimized NLE. C. and D. $|E|^2$ distribution in the FoM plane of the raw excitation source without and with the NLE. E. The normalized extraction efficiency of the optimized structured and final FoM.

For our demonstration, we paired the FoM for minimizing the transmission of the excitation fields with our Gaussian beam phase maximization approach used in section 3.4.2 [*i.e.* using two orthogonal dipoles oriented in the XY plane to model the (111) NV]. We weighted the optimization gradient contributions from the excitation minimization and beaming FoMs equally. The results of the optimization are shown in Figure 3.16 (B-E). In particular, shown in Figure 3.16(C,D) the peak electric field intensity has been reduced by 71% compared to the sans-NLE

case. Figure 3.16(E) demonstrates that there is little cost to this additional functionality as this NLE has a final FoM of 23.6 (compared to the stock NLE FoM of 24) and maintains good extraction efficiency across the entire spectrum (> 10). Figure 3.17(A) shows the resulting farfield $|E|^2$ which shows a slight compromise with the emission cone being slightly wider than the stock NLE case to enable the large reduction in excitation field intensity. Figure 3.17(B) shows the $|E|^2$ of the stock NLE [Figure 3.11(A)]. The stock NLE results in a peak field intensity of 3.4× that of the excitation beam alone [Figure 3.16(C)] and almost 8× larger than the excitation minimizing NLE.

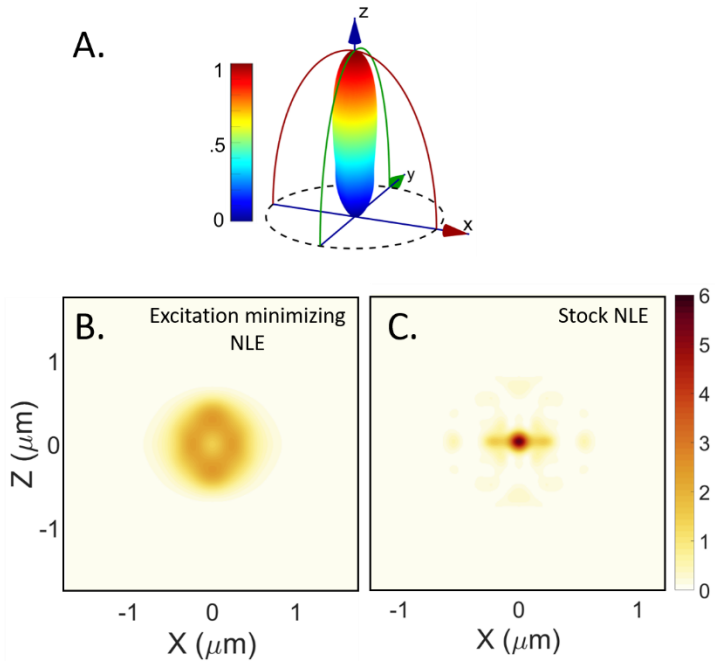

Figure 3.17 – A. Farfield $|E|^2$ of the excitation minimization NLE. B. and C. Side-by-side comparison of $|E|^2$ of the excitation minimizing NLE and stock NLE presented in Figure 3.11(A). The peak intensity reaches a value 340% larger than the excitation source without any structure [Figure 3.16(C)].

3.4.7 Conclusion

In conclusion, we present several NLE geometries optimized using adjoint optimization techniques with varying FoMs for use in enhancing and controlling the emission from shallow NV centers in (111) bulk diamond. Our results demonstrate significant enhancement of the NV emission on the order of comparable enhancement techniques presented elsewhere in the literature. Our techniques have the further advantage of fine control over the emission beam properties as well as preservation of the diamond surface as no etching of the surface is required. Furthermore, these advanced configurations can pave the way for creating compact, on-chip NLEs with bespoke functionalities. These geometries are useful for many experimental applications and configurations using NV emitters but the adjoint optimization techniques described can be generalized to other emission applications such as Si defects in diamond, point emitters in semiconductors or 2D emitters such as hexagonal boron-nitride [153], [154]. We believe further understanding and utilization of adjoint optimization techniques can pave the way to further high-performance nano-photonic device engineering in many more applications.

4 An intuitive model for reflectarray metasurfaces

This work benefitted significantly from initial investigations by Anthony Wang (former member of Mikhail Kats's group and Victor Brar's group) as well as guidance and contributions by Victor Brar.

4.1 Introduction

The ability to control and manipulate the wavefront of the light using optical components has been paramount to many technological innovations and achievements. Traditional optical components, such as mirrors, beam splitters and lenses, make use of macro-scale manipulation of the wavefront, i.e., they utilize reflection, refraction, and diffraction to achieve a desired beam pattern. Metasurfaces are optical components that are designed to manipulate light by selectively controlling its amplitude, phase, or polarization as it interacts with a thin engineered surface [71], [75], [186]–[188]. They are comprised of sub-wavelength structures that alter the phase and transmission of light on a local scale. Such metasurfaces can be made of dielectric or metallic materials and can operate in reflective or transmissive geometry. In the most-common type of metasurface, engineering of the wavefront is achieved through altering the resonance of the resonant structures such as metal antennas or dielectric Mie-type resonators, which is usually accomplished through changing their geometry (e.g., height, width, length, etc.). In this chapter, we will focus on a specific type of metasurface comprising metallic nanostrips placed periodically above a metallic ground plane separated by a dielectric spacer [189]–[194] and we seek to explain how the metasurface achieves a full 2π phase coverage of the scattered light in addition to what causes an anomalous phase response exhibited by

this structure. This type of metasurface has recently been used to construct tunable and dynamically responsive surfaces [195]–[197]. Previous examples of reflectarrays include metasurfaces designed for polarization conversion[198], vortex beam generation[199], anomalous reflection[200], and birefringent surfaces[189] among others.

In isolation, metallic nano-rod resonators can only scatter incident light with a maximum range of phase of π relative to the incident wave (Figure 4.1) [7]. The phase difference can be understood in the limit of the rod being much smaller than the incident wavelength. The charges in the rod, responsible for the scattered field, re-radiate light proportional to their acceleration. At max incident field, the charges are not accelerating and therefore the scattered field is at a minimum. When the incident field flips direction, the charges are at max acceleration and therefore maximum scattered field, thus a π shift in the scattered field phase. At very long rod lengths, the charges are able to accelerate with the incident field (*i.e.* they do not reach an end and can continue matching the field) and therefore scatter the light in phase with the incident field. Therefore, additional mechanisms are required to increase the phase coverage of the nanorods to 2π. Without 2π phase control, phase discontinuities will be prevalent resulting in unwanted interference and leading to poor beam propagation or an unusable profile. Therefore, being able to tune the scattered field phase a full 2π radians is important for metasurfaces to be able to create arbitrary wavefronts.

In the literature, the full phase coverage has been accomplished by coupled nanorod resonances [78], [201], [202], polarization effects (through the Pancharatnam–Berry phase) [203]–[205] or, in the case of reflectarrays, coupling the resonators to a ground plane. In this configuration, metallic nanorods are placed above a metallic ground plane with a dielectric spacer in between. It is generally understood that the increased phase coverage originates from the onset of gap plasmon resonances which are strongly localized plasmonic resonances that exist in small gaps between

metallic films [206]. The coupling to and subsequent radiation from these modes is responsible for creating the phase delay in the reflected light. One can then obtain the phase dependence as a function of specific geometric parameters of the layered structure which can then be used for designing optical metasurfaces. However, it has been observed that these metasurfaces exhibit an anomalous phase response where the range of phases attainable reduces from $\sim 2\pi$ to almost 0 for specific periods or substrate index/thickness values. One model for understanding this behavior was put forward by Park *et al.* in [198], using nanostrips on a substrate with an electrically tunable refractive index. In their work, they used coupled mode theory (CMT) to differentiate between the contributions of the resonant and non-resonant reflection channels owing to the reflection from the ground plane and the gap plasmon modes, respectively. We seek to expand upon previous explanations by offering an intuitive and pedagogic model of the origin of the metasurface phase coverage behavior. Our aim in this instructional chapter is twofold: 1) we seek to give an intuitive explanation and demonstration for how elements of reflectarray metasurfaces can cover a phase range of 2π, and 2) use this model to explain and visualize reduction in attainable scattered field phases.

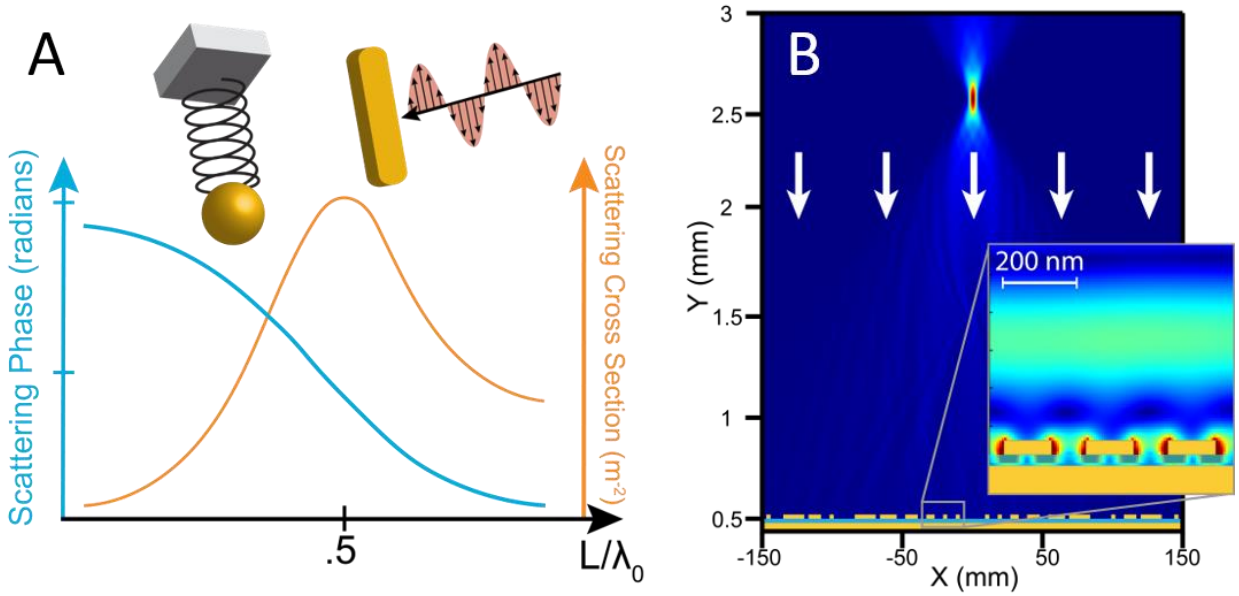

Figure 4.1 - A) Visualization of the general behavior of damped, driven oscillators. For a constant frequency (ω), there is a specific ratio of spring constant to mass ($\sqrt{k/m}$) or nano strip length (L) at which these oscillators see a maximum in their amplitude (orange curve) while the phase (blue curve) of oscillation undergoes a shift of $\pi/2$ relative to the driving force. B) Example simulation of a focusing reflective metasurface. Plotted here is the $|E|^2$ after an incident plane wave reflects from the metasurface. The structure at the bottom is a cartoon representation of the actual structure, not to scale. The inset shows the local $|E|^2$ around a subset of three elements that comprise the device.l

4.2 Intuitive Model

The interaction of light with a metallic nanorod can be understood through the lens of a damped, driven harmonic oscillator. The electron sea in the metal oscillates with the incident fields and, in the case of a finite rod, the negative charge builds in one end of the rod giving rise to a collection of positive charges on the opposing end of the rod which acts as a restoring force for the previously displaced electrons. The restoring force attracts the electrons which causes them to oscillate towards the positive side and thus the cycle continues. The charge-force system is analogous to a damped driven mass-on-a-spring harmonic system in classical physics where the mass is the mass of the displaced electrons, the Coulombic force between the charged ends of the rod act as the spring,

the damping force arises from electron-lattice collisions as well as radiation and the driving force is the incident electromagnetic wave [7]. The entire collective oscillation of the free electrons being driven is termed a plasmon [207]. A cartoon of the scattering field response of the charge-force damped driven harmonic oscillator is shown in Figure 4.1. Because the incident light acts as a driving force of the resonator, the frequency (wavelength) of the light is the frequency of the force. As the driving frequency approaches the resonant frequency of the oscillator (ω_0), the scattering cross section reaches its peak value while the scattered field phase is exactly $\pi/2$ out of phase with the driving force[78]. At frequencies above the resonant frequency, the scattered phase reaches a peak of π radians out of phase with the driving force.

While it is widely known in the field that coupling the resonators to a ground plane increases the phase coverage of the scattered fields, an intuitive and pedagogic understanding of this mechanism is mostly absent. Previous explanations for the behavior of these metasurfaces have attributed the field response to magnetic dipole radiation [208], gap plasmons [209] or have used coupled-mode theory [210], [211]. While these models can be used to explain the behavior of the coupled nanostrips, we seek to put forward an intuitive and easily visualized model of this behavior. Here, we consider scattered field from the ground-plane-mounted nanostrip (3D nanorod reduced to 2D) as the sum of the scattered field from the nanostrip itself and the scattered field from the substrate (metallic ground plane plus the dielectric spacer).

We are able to isolate the scattered fields using a simulation formalism called a Total-Field Scattered-Field (TFSF) [212]–[214]. A TFSF simulation is comprised of two regions: the total field region and the scattered field region. The total field region contains all simulation fields, namely, the injected source fields (a plane wave) as well as the scattered fields from the enclosed particles. The scattered field region is the interesting feature of the TFSF element; it consists of all areas outside of

the total-field region. At the element boundary, the same incident plane wave that was injected into the simulation is subtracted out from the fields passing through which causes only the fields scattered by objects within the TFSF region to pass while fields that are unaffected by the enclosed elements are removed. If there is a substrate that extends through the TFSF boundary, as in this case, the fields scattered by the substrate (i.e., reflected from a metal substrate) are subtracted out and do not contribute to the fields in the scattered-field region. The TFSF source is primarily used for scattering and absorption cross-sections of small particles but can be used to isolate the scattering response of any sort of object.

We use three basic configurations in our simulations of the metasurface in order to isolate the various scattering components and these are outlined in Figure 4.2. The first configuration (config. 1) isolates the scattered phase of the nanostrips themselves. We use periodic boundary conditions to enable interaction of the strips with neighbors and the reflected field is recorded at the e-field monitor placed above the strips. The recorded field at the monitor is then projected 1 m into the farfield where we then can take the phase and amplitude. The second configuration (config. 2) we used removes the nanostrip and truncates the substrate. The substrate width is set to the same period as used in config. 1 and the FDTD boundaries are switched to absorbing boundaries. In this way, the TFSF surrounds the substrate and the reflected fields at the monitor are due to the substrate alone. The scattered field amplitude and phase are determined the same as config. 1. The third configuration (config. 3) retains the truncated substrate and absorbing boundaries but adds the nanostrip back in resulting in the full scattered response of both the substrate and nanostrip together in a non-periodic arrangement (with the scattered field and amplitude calculated via the farfield transform).

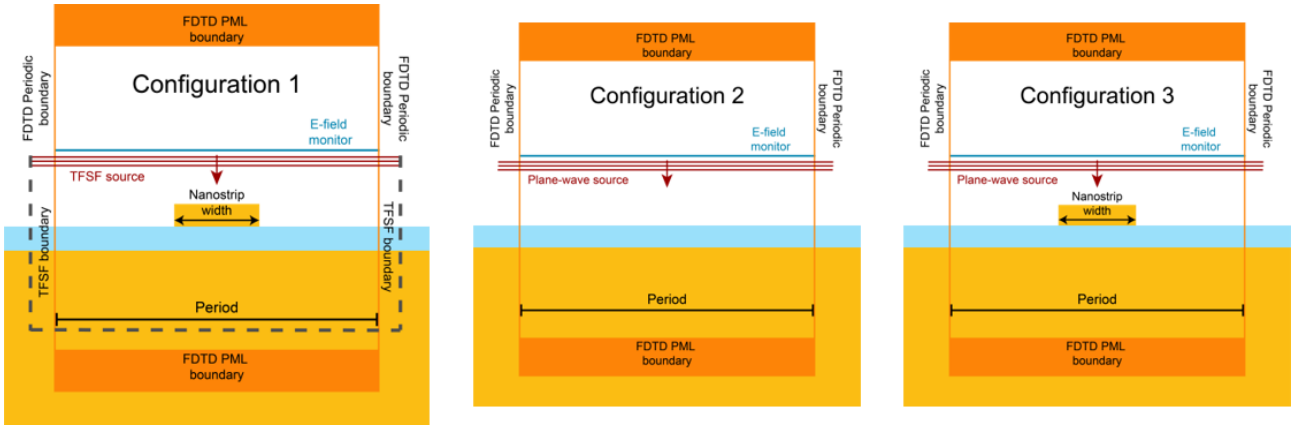

Figure 4.2 - Schematic showing the simulation configurations to isolate the individual reflectarray element contributions to the scattered field. Config. 1 isolates the nanostrip, Config. 2 isolates the substrate contribution, and Config. 3 gives the combined response.

With these simulation configurations we are able to show the contributions of each element of the metasurface and demonstrate that the scattered field of the entire structure is comprised of the response of each individual part. This concept is represented in Figure 4.3A. The example metasurface seen at the top of Figure 4.3A is comprised of individual scattering elements seen in the central red box. The scattering response of the ground plane (yellow box) along with the scattering response of the nanostrips (blue box, represented by a harmonic oscillator) constitutes the total scattering response of the element. The combined scattered field effect is elucidated through full-wave simulations of the individual components of this metasurface. From these simulations, the scattered fields can be broken down and isolated between the individual contributions of the nanostrip and of the piece of substrate (all simulations were performed with Lumerical FDTD [136]).

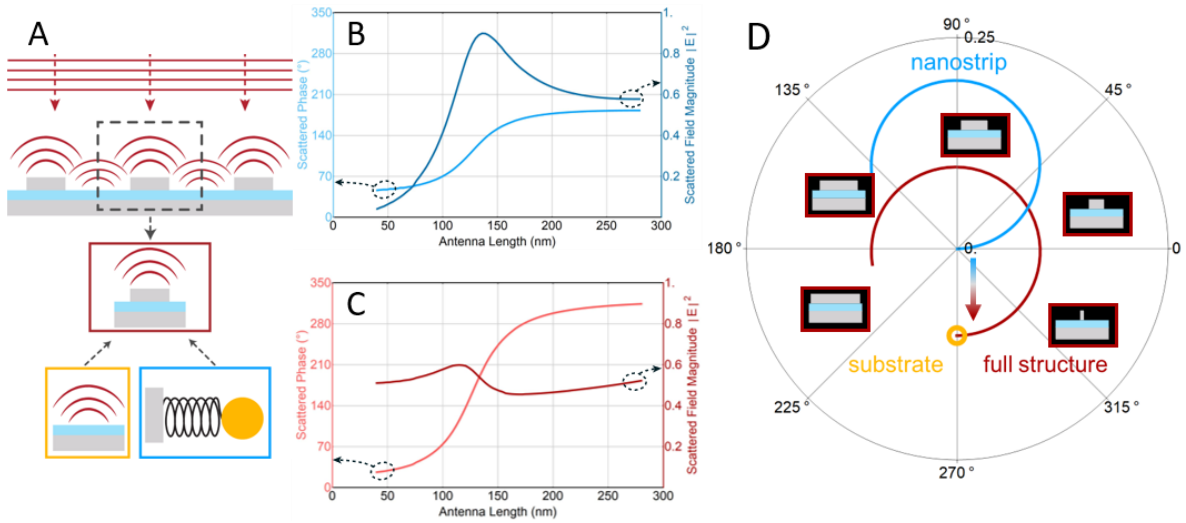

Figure 4.3 - A) Cartoon of the components that comprise a reflective metasurface. The entire metasurface is comprised of individual resonators (red box) which scatter incident light with a phase delay that can be engineered via the geometry of the unit cell. The response of each individual resonator can be broken down into a substrate scattering component (yellow box) and a nanostrip scattering component (blue box). B) Plot of the phase and scattering amplitude of a single nanostrip (the response of the substrate has been subtracted out). Note the scattered field amplitude has a strong peak and the phase response on varies from 0 to approximately π, as would be expected for a damped, driven harmonic oscillator. C) Phase and scattering amplitude of a full metasurface element (combination of the substrate response and the nanostrip response). The blue circles and red dots in the plot correspond to identical antenna lengths in both configurations. The cartoons along the red curve depict the structure at each circle. Note the scattering-field amplitude is much flatter and the phase covered goes from 0 to almost 2π. D) The scattered field response of only the nanostrip (solid blue curve) and the full element (solid red curve) in polar form with increasing nanostrip length in the counterclockwise direction. The substrate scattered field and phase is marked by a yellow circle. When the nanostrip response is added to the substrate response, the resultant curve is the full element simulation.

Employing the TFSF source to isolate the scattered fields, we performed 2D simulations using gold nanostrips (really, patches, because in these 2D simulations the gold structures have infinite extend in the out-of-page direction), on an MgF_2 dielectric layer with a thickness of 50 nm (n = 1.373) [215], [216], and a gold ground plane. In Figure 4.3B we plot the scattered field magnitude and phase of the isolated nanostrips as function of strip width (config. 1). In each case, the nanostrip width is swept from 40 to 280 nm, the period was fixed to 300 nm, and our plane wave source was set to a fixed free-space wavelength of 850 nm. Consistent with our previous description, Figure 4.3 shows the range of scattered field phase of the nanostrip covers fewer than π radians as the length is swept, and the peak scattering amplitude occurs at the plasmon-resonant strip width (~140 nm). We contrast

this response with that shown in Figure 4.3C. He we show the scattered field from the substrate and the nanostrip together (config. 3) and the phase coverage has increased to reach a full 2π while the scattering amplitude stays relatively constant.

The key insight to the behavior is gained when we plot the phase and amplitude together in polar form (Figure 4.4D). The response of the nanostrip (config. 1) is seen in the solid blue curve. As the length of the nanostrip is increased, the phase and amplitude response can be traced counterclockwise around the curve. It should be noted here that this is the scattered field of the nanostrip in the presence of the substrate. That is, this strip resonance curve is due to interactions with the substrate and is different from the scattered field of the strip in free space. The curve being situated solely in the upper half of the plot shows only 180° scattered phase coverage. Even though this is the field of the nanorod in the presence of the substrate, there are no geometric configurations of the nanostrip array (*e.g.* thickness, length, periodicity, etc…) that will give rise to full 2π phase coverage. The full coverage is only obtained by adding the response of the ground plane and MgF_2 spacer (config. 2), which is situated in the fourth quadrant. The location of this point is consistent with the case where the strip width in config. 3 approaches 0 and is primarily determined by the properties of the dielectric spacer (*e.g.* thickness, dielectric constant, width etc…). We then add the response of the nanostrip with the response of the ground plane to illustrate that the total response of the reflectarray nanostrip as a whole is the combined response of these two elements resulting in the red curve in Figure 4.3D. Now it can be seen that the center of the circle is shifted through the origin, which explains the larger phase coverage along with the compression of the scattered amplitude seen in Figure 4.3C. Again, the red curve is the full structure (config. 3) which is the added response of the isolated nanostrip (config. 1) with the substrate response (config. 2). Here we see the full phase coverage as the scattered field traces nearly a full circle around the polar axis.

4.3 Anomalous Phase Behavior

One interesting effect that can be seen from reflectarray metasurfaces but is hard to explain without the model we present here is the significant reduction in the range of scattered phases for particular unit-cell geometries. This behavior is demonstrated in Figure 4.4A and B. For these simulations, we isolate the phase response of an array of nanostrips (configuration 1 described previously) but focus on the effects of changing the scattered field of the substrate. The anomalous phase response occurs in this configuration with increasing periodicity of the metasurface, corresponding to a larger unit cell, effectively increasing the length of the ground-plane substrate. Each curve in Figure 4.4A and B corresponds to a different periodicity, and in all cases the nanostrip width is swept from 0 to 280 nm with an incident free-space wavelength of 850 nm. With increasing ground-plane length, the scattered phase shift as a function of strip width becomes increasingly abrupt and the scattered field amplitude is reduced until, at the largest periodicity, the maximum phase coverage drops to only 80°. As the phase coverage is reduced in Figure 4.4B, the corresponding scattered field amplitudes can be seen to reach a minimum as well in Figure 4.4A. The minimum in scattered field corresponds to the point where the scattered field from the nanostrips destructively interfere with the scattered field from the substrate and thus the total combined scattered field is negated to an evanescent field.

The reduction in scattered field amplitude has been seen and utilized in a number of literature results. Early observations lead to the development of metasurfaces used for perfect absorption because of the reduction in scattering amplitude [217]–[219]. Recent metasurface designs take

advantage of this pass through point (where the scattered field passes through the origin of the polar plot) to create active, tunable metasurfaces that use electro-optical materials to utilize this effect [192], [198],[220], [221]. The main tunability knob comes through electrically changing the index of refraction of the dielectric spacer layer which can be accomplished by using electrically connected conductive oxides such as indium tin oxide (ITO) [222].

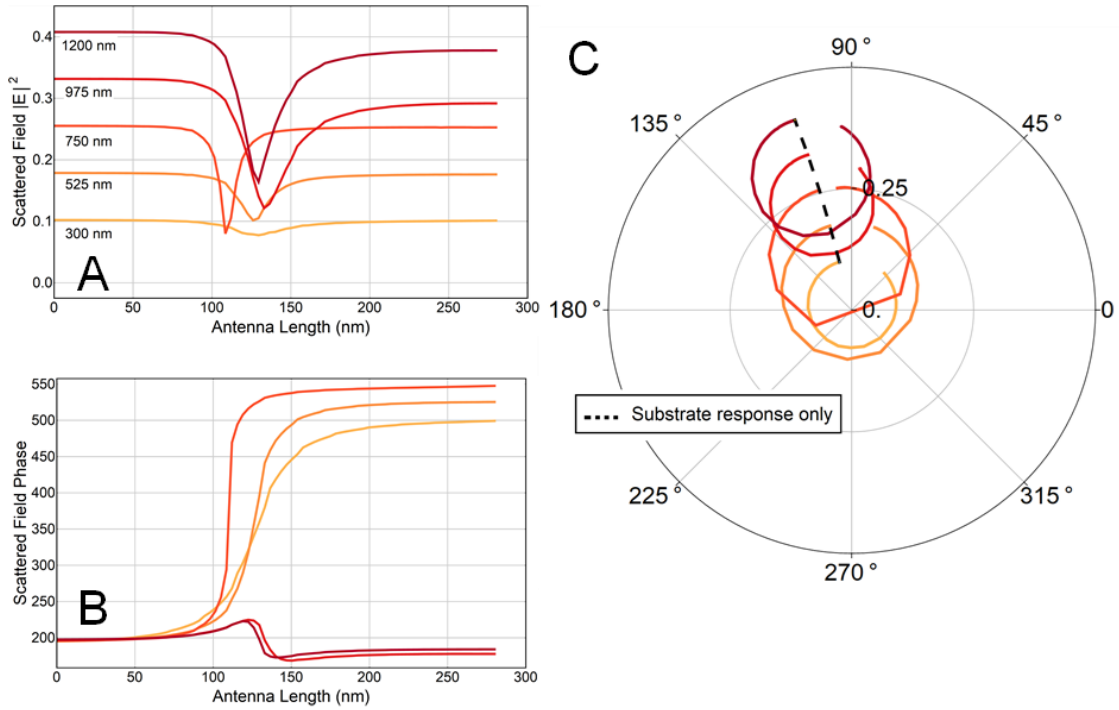

Figure 4.4 - A demonstration and visual explanation of the collapsing phase behavior exhibited by reflective metasurfaces as the width of the unit cell is varied. A) The simulated scattered field amplitude and phase B) of the full metasurface element as the width of the substrate is increased. The labels indicate unit cell width. For both A and B, the nanostrip length is swept between 40 and 280 nm for each substrate width. C) The scattered fields from A and B in polar form (solid red-orange lines corresponding to periodicity/width). As the substrate width is increased, the response curve sweeps across the origin of the plot leading to anomalous scattered phase coverage. The dashed black line plots the substrate scattered field for increasing width.

Although not shown here, changing the refractive index of the substrate (as done in active reflectarray metasurfaces) or changing the thickness of the substrate results in similar changes to the scattered field as changing the periodicity due changing the scattered field of the substrate and can be

illustrated by looking at the polar plot of the scattered fields in Figure 4.4C. The black dashed line in this plot corresponds to the substrate scattered field and was obtained by using the previously described configuration 2. The nanostrip is removed and the substrate width (MgF_2 film and Au ground plane) is truncated to the width of the unit cells in Figure 4.4A and B. We then simulate and record the scattered field phase and amplitude. Here, we can see the response of the isolated substrate (dotted black line), which increases in amplitude radially outward while staying relatively constant in phase as the unit cell width is increased. This behavior is expected as changing the width of the substrate should have a negligeable effect on the scattered phase (as it is just reflecting the incident wave and is a non-resonant structure) and increase in amplitude as its surface area becomes larger and more of the beam is reflected. The polar phase arc is essentially offset and translated through the origin. Indeed, it is the same effect as in Figure 4.3D, where the substrate enables a full 2π phase coverage by translating the response of the nanostrip through the origin. Therefore, one can see from this example that the phase response is strongly dictated by the intrinsic scattered field of the substrate. Any effect that changes the substrate's scattering properties (*e.g.*, changing the index of refraction or the thickness of the dielectric spacer) can strongly impact the phase of the scattered fields of the combined structure.

Finally, we show in Figure 4.5 that this model not only works for two dimensional nanostrips but extends to three dimensional nanorod structures as well. Figure 4.5A shows the polar plot of the nanorod scattered field (3D config. 1), substrate scattered field (3D config. 2), the additive response of the nanostrip plus the substrate (calculated from 1 and 2), and the full structure simulation for comparison (3D config. 3). The rods simulated have a width of 75 nm, a thickness of 30 nm and are situated on a 50 nm thick MgF_2 spacer layer. The periodicity is set to 300×300 nm. The inset is a

cartoon of the full 3D structure with the TFSF source represented as the dashed box and the plane of incidence is highlighted in red.

Leveraging the full phase coverage ability, which we have demonstrated is due to the addition of the substrate scattered field, we are able to construct a full 3D metasurface: here, an optical vortex beam generator. An optical vortex beam is an free-space optical beam with a phase front that is twisted around the azimuthal angle, normal to the direction of propagation [223]. The corkscrew phase profile creates a central minimum of field intensity as the fields in this location destructively interfere. Optical vortex beams can carry orbital angular momentum and, in turn, can produce a torque on an electric dipole [224]. These beams can be used as optical tweezers (holding a particle in place) or as an information carrier in a photonic-based communication systems [225].

In this example, we used the recorded scattered field phases from the nanorods to engineer a 3D vortex beam generator. For this metasurface we needed the combined response of both the substrate and nanorods, ideally finding the scattered field as a function of rod length. Therefore, we used a periodic simulation with a standard plane wave source (config. 3 in 3D). We similarly recorded the field at the monitor and use the farfield projection to determine the field phase and amplitude. Knowing the scattered field phase as a function of rod length, we can then construct the metasurface to reflect the desired azimuthally varying phase profile. In this case, we discretized the phase into 6 slices to provide better field continuity.

The phase profile is demonstrated in Figure 4.5A which shows a cartoon of the structure showing that the nanostrip length is chosen to give an azimuthally varying phase profile from 0 to 2π around the circumference of the metasurface. The resulting field magnitude is shown in Figure 4.5B. While the beam quality is not perfect, there is a well-defined central cavity surrounded by a higher intensity donut. Further, the azimuthally varying phase is clearly visible (Figure 4.5C).

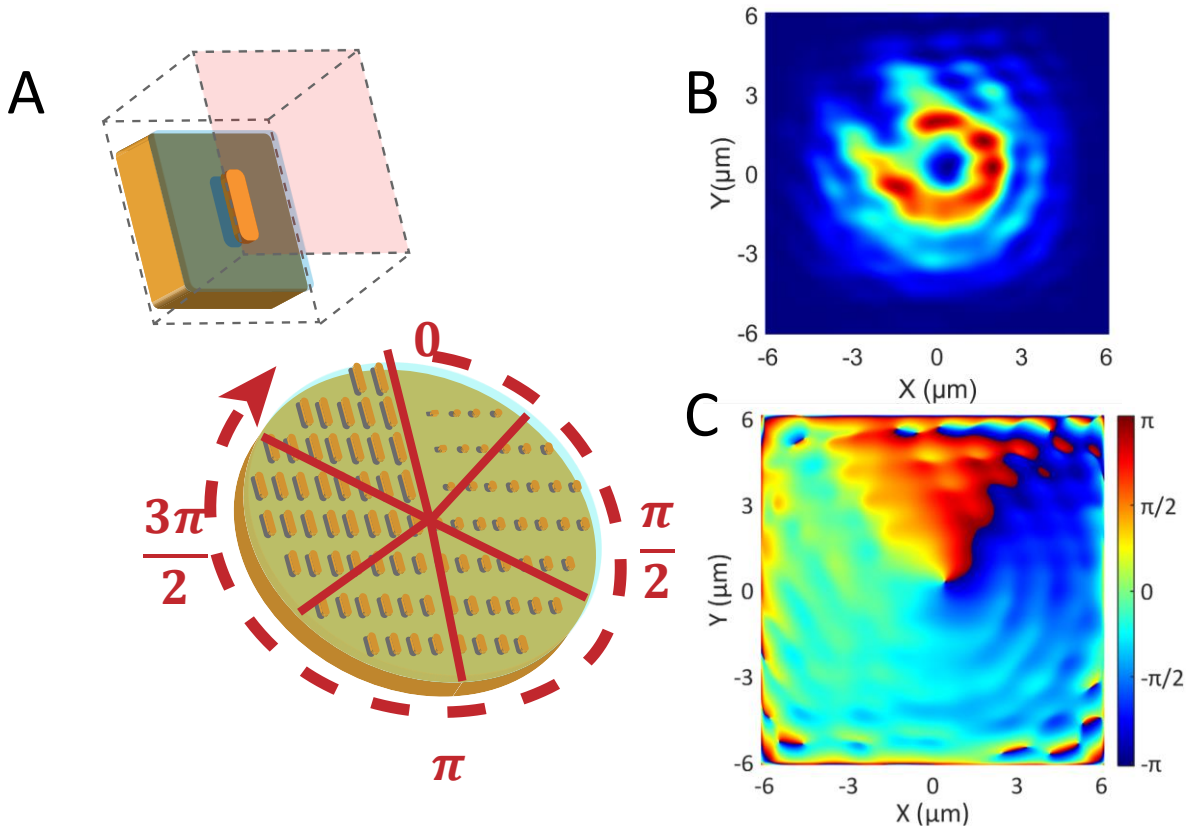

Figure 4.5 - A simulated demonstration of a 3D vortex beam generator using a reflectarray metasurface. A) A schematic of the individual nanorod metasurface element along a schematic of the phase breakdown of the vortex generator. The phases are broken into six discrete chunks from 0 to 2π which increases in the azimuthal direction. B) and C) The scattered field intensity and phase, respectively, of a reflectarray metasurface designed to generate a vortex beam. The beam has a donut shaped intensity profile encircling a point of undetermined phase as can be seen in the above figures.

4.4 Conclusion

In summary, we present an intuitive model for understanding and engineering of reflectarray metasurfaces, and a clear visualization of the anomalous phase response displayed by these structures. Determination of the phase response can be thought of as the combination of the scattering response of a single nanostrip with the scattered field response of the substrate. Using this model, we explain the importance of the substrate scattered field that drives both the anomalous phase response of these metasurfaces as well as enables full 2π phase coverage of the scattered field of the nanorods. Factors

that change or modify the substrate scattering (*i.e.*, dielectric spacer thickness or index of refraction) gives strong control over the behavior of the metasurface. Our model can be used to simplify the engineering of reflectarray metasurfaces and give an intuitive approach for those designing such devices. Primarily, understanding and engineering the substrate scattered field can lead to significant variations of the response of the entire metasurface. Therefore, each application can find appropriate substrate properties to give the desired response. We also show some capabilities of the reflectarray metasurface by demonstrating a metasurface that generates an optical vortex beam. We hope that our intuitive model will be helpful for future design of reflectarray and active tunable metasurfaces.

References

[1] A. Diaspro, *Optical Fluorescence Microscopy: From the Spectral to the Nano Dimension*. Berlin, Heidelberg, GERMANY: Springer Berlin / Heidelberg, 2010.

[2] M. L. Rubin, "Spectacles: Past, present, and future," *Survey of Ophthalmology*, vol. 30, no. 5, pp. 321–327, Mar. 1986, doi: 10.1016/0039-6257(86)90064-0.

[3] W. D. Callister and D. G. Rethwisch, *Fundamentals of Materials Science and Engineering: An Integrated Approach*. John Wiley & Sons, 2012.

[4] D. J. Griffiths, *Introduction to Electrodynamics*. Cambridge University Press, 2017.

[5] N. Ashcroft and N. Mermin, *Solid State Physics*. Cengage Learning, Incorporated, 2021.

[6] C. Kittel, *Introduction to Solid State Physics*. John Wiley & Sons, 2015.

[7] Lukas. Novotny and B. Hecht, *Principles of nano-optics*. Cambridge University Press, 2012.

[8] S. Basu, B. J. Lee, and Z. M. Zhang, "Infrared Radiative Properties of Heavily Doped Silicon at Room Temperature," *Journal of Heat Transfer*, vol. 132, no. 023301, Nov. 2009, doi: 10.1115/1.4000171.

[9] J. D. Plummer, M. D. Deal, and P. B. Griffin, *Silicon VLSI technology : fundamentals, practice, and modeling*. Prentice Hall, 2000.

[10] U. Guler, V. M. Shalaev, and A. Boltasseva, "Nanoparticle plasmonics: going practical with transition metal nitrides," *Materials Today*, vol. 18, no. 4, pp. 227–237, May 2015, doi: 10.1016/j.mattod.2014.10.039.

[11] G. V. Naik, V. M. Shalaev, and A. Boltasseva, “Alternative Plasmonic Materials: Beyond Gold and Silver,” *Advanced Materials*, vol. 25, no. 24, pp. 3264–3294, 2013, doi: https://doi.org/10.1002/adma.201205076.

[12] C. M. Zgrabik and E. L. Hu, “Optimization of sputtered titanium nitride as a tunable metal for plasmonic applications,” *Opt. Mater. Express, OME*, vol. 5, no. 12, pp. 2786–2797, Dec. 2015, doi: 10.1364/OME.5.002786.

[13] G. V. Naik, J. L. Schroeder, X. Ni, A. V. Kildishev, T. D. Sands, and A. Boltasseva, “Titanium nitride as a plasmonic material for visible and near-infrared wavelengths,” *Opt. Mater. Express, OME*, vol. 2, no. 4, pp. 478–489, Apr. 2012, doi: 10.1364/OME.2.000478.

[14] R. Soref, J. Hendrickson, and J. W. Cleary, “Mid- to long-wavelength infrared plasmonic-photonics using heavily doped n-Ge/Ge and n-GeSn/GeSn heterostructures,” *Opt. Express, OE*, vol. 20, no. 4, pp. 3814–3824, Feb. 2012, doi: 10.1364/OE.20.003814.

[15] L. Baldassarre *et al.*, “Midinfrared Plasmon-Enhanced Spectroscopy with Germanium Antennas on Silicon Substrates,” *Nano Lett.*, vol. 15, no. 11, pp. 7225–7231, Nov. 2015, doi: 10.1021/acs.nanolett.5b03247.

[16] S. Prucnal *et al.*, “Ultra-doped n -type germanium thin films for sensing in the mid-infrared,” *Scientific Reports*, vol. 6, no. 1, Art. no. 1, Jun. 2016, doi: 10.1038/srep27643.

[17] Y. B. Chen and Z. M. Zhang, “Heavily doped silicon complex gratings as wavelength-selective absorbing surfaces,” *J. Phys. D: Appl. Phys.*, vol. 41, no. 9, p. 095406, Apr. 2008, doi: 10.1088/0022-3727/41/9/095406.

[18] J. C. Ginn, R. L. Jarecki, E. A. Shaner, and P. S. Davids, “Infrared plasmons on heavily-doped silicon,” *Journal of Applied Physics*, vol. 110, no. 4, p. 043110, Aug. 2011, doi: 10.1063/1.3626050.

[19] M. Shahzad *et al.*, “Infrared surface plasmons on heavily doped silicon,” *Journal of Applied Physics*, vol. 110, no. 12, p. 123105, Dec. 2011, doi: 10.1063/1.3672738.

[20] F. Marquier, K. Joulain, J. P. Mulet, R. Carminati, and J. J. Greffet, “Engineering infrared emission properties of silicon in the near field and the far field,” *Optics Communications*, vol. 237, no. 4, pp. 379–388, Jul. 2004, doi: 10.1016/j.optcom.2004.04.024.

[21] A. Rosenberg *et al.*, “Flat mid-infrared composite plasmonic materials using lateral doping-patterned semiconductors,” *J. Opt.*, vol. 16, no. 9, p. 094012, Sep. 2014, doi: 10.1088/2040-8978/16/9/094012.

[22] F. Marquier, M. Laroche, R. Carminati, and J.-J. Greffet, “Anisotropic Polarized Emission of a Doped Silicon Lamellar Grating,” *Journal of Heat Transfer*, vol. 129, no. 1, pp. 11–16, Jun. 2006, doi: 10.1115/1.2360594.

[23] W. Streyer, S. Law, G. Rooney, T. Jacobs, and D. Wasserman, “Strong absorption and selective emission from engineered metals with dielectric coatings,” *Opt. Express, OE*, vol. 21, no. 7, pp. 9113–9122, Apr. 2013, doi: 10.1364/OE.21.009113.

[24] S. Law, D. C. Adams, A. M. Taylor, and D. Wasserman, “Mid-infrared designer metals,” *Opt. Express, OE*, vol. 20, no. 11, pp. 12155–12165, May 2012, doi: 10.1364/OE.20.012155.

[25] E. Tokumitsu, “Correlation between Fermi Level Stabilization Positions and Maximum Free Carrier Concentrations in III-V Compound Semiconductors,” *Japanese Journal of Applied Physics*, vol. 29, no. Part 2, No. 5, pp. L698–L701, May 1990, doi: 10.1143/JJAP.29.L698.

[26] S. Law, L. Yu, and D. Wasserman, “Epitaxial growth of engineered metals for mid-infrared plasmonics,” *J. Vac. Sci. Technol. B*, vol. 31, no. 3, p. 03C121, May 2013, doi: 10.1116/1.4797487.

[27] S. Law, L. Yu, A. Rosenberg, and D. Wasserman, “All-Semiconductor Plasmonic Nanoantennas for Infrared Sensing,” *Nano Lett.*, vol. 13, no. 9, pp. 4569–4574, Sep. 2013, doi: 10.1021/nl402766t.

[28] G. V. Naik, J. Kim, and A. Boltasseva, “Oxides and nitrides as alternative plasmonic materials in the optical range [Invited],” *Opt. Mater. Express, OME*, vol. 1, no. 6, pp. 1090–1099, Oct. 2011, doi: 10.1364/OME.1.001090.

[29] A. Agrawal, R. W. Johns, and D. J. Milliron, “Control of Localized Surface Plasmon Resonances in Metal Oxide Nanocrystals,” *Annu. Rev. Mater. Res.*, vol. 47, no. 1, pp. 1–31, Jul. 2017, doi: 10.1146/annurev-matsci-070616-124259.

[30] H. Zhang *et al.*, “Doped Silicon Nanocrystal Plasmonics,” *ACS Photonics*, vol. 4, no. 4, pp. 963–970, Apr. 2017, doi: 10.1021/acsphotonics.7b00026.

[31] P. Zijlstra, J. W. M. Chon, and M. Gu, “White light scattering spectroscopy and electron microscopy of laser induced melting in single gold nanorods,” *Physical Chemistry Chemical Physics*, vol. 11, no. 28, p. 5915, Jul. 2009, doi: 10.1039/b905203h.

[32] N. Pfullmann *et al.*, “Bow-tie nano-antenna assisted generation of extreme ultraviolet radiation,” *New Journal of Physics*, vol. 15, no. 9, p. 093027, Sep. 2013, doi: 10.1088/1367-2630/15/9/093027.

[33] P. Biagioni, J.-S. Huang, and B. Hecht, “Nanoantennas for visible and infrared radiation,” *Rep. Prog. Phys.*, vol. 75, no. 2, p. 024402, Jan. 2012, doi: 10.1088/0034-4885/75/2/024402.

[34] V. Giannini, A. I. Fernández-Domínguez, S. C. Heck, and S. A. Maier, “Plasmonic Nanoantennas: Fundamentals and Their Use in Controlling the Radiative Properties of Nanoemitters,” *Chem. Rev.*, vol. 111, no. 6, pp. 3888–3912, Jun. 2011, doi: 10.1021/cr1002672.

[35] Y. Zhong, S. D. Malagari, T. Hamilton, and D. M. Wasserman, “Review of mid-infrared plasmonic materials,” *JNP*, vol. 9, no. 1, p. 093791, Feb. 2015, doi: 10.1117/1.JNP.9.093791.

[36] N. E. Dickenson, E. S. Erickson, O. L. Mooren, and R. C. Dunn, “Characterization of power induced heating and damage in fiber optic probes for near-field scanning optical microscopy,” *Review of Scientific Instruments*, vol. 78, no. 5, p. 053712, May 2007, doi: 10.1063/1.2740133.

[37] G. Masetti, M. Severi, and S. Solmi, “Modeling of carrier mobility against carrier concentration in arsenic-, phosphorus-, and boron-doped silicon,” *IEEE Transactions on Electron Devices*, vol. 30, no. 7, pp. 764–769, Jul. 1983, doi: 10.1109/T-ED.1983.21207.

[38] E. Hecht, *Optics (Fifth Edition)*. Pearson Education, New Jersey, United States, 2015.

[39] *Lumerical Inc.* .

[40] D. L. Losee, J. P. Lavine, E. A. Trabka, S. -T. Lee, and C. M. Jarman, “Phosphorus diffusion in polycrystalline silicon,” *Journal of Applied Physics*, vol. 55, no. 4, pp. 1218–1220, Feb. 1984, doi: 10.1063/1.333166.

[41] E. Di Fabrizio *et al.*, “High-efficiency multilevel zone plates for keV X-rays,” *Nature*, vol. 401, no. 6756, Art. no. 6756, Oct. 1999, doi: 10.1038/44791.

[42] I. Mohacsi *et al.*, “High-efficiency zone-plate optics for multi-keV X-ray focusing,” *Journal of Synchrotron Radiation*, vol. 21, no. 3, pp. 497–501, May 2014, doi: 10.1107/S1600577514003403.

[43] M. H. Horman and H. H. M. Chau, “Zone Plate Theory Based on Holography,” *Appl. Opt., AO*, vol. 6, no. 2, pp. 317–322, Feb. 1967, doi: 10.1364/AO.6.000317.

[44] D. Attwood, *Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Applications*. Cambridge University Press, 2007.

[45] Q. Cao and J. Jahns, "Comprehensive focusing analysis of various Fresnel zone plates," *J. Opt. Soc. Am. A, JOSAA*, vol. 21, no. 4, pp. 561–571, Apr. 2004, doi: 10.1364/JOSAA.21.000561.

[46] B. A. Munk, *Frequency Selective Surfaces: Theory and Design*. John Wiley & Sons, 2005.

[47] L. Novotny, "Effective Wavelength Scaling for Optical Antennas," *Phys. Rev. Lett.*, vol. 98, no. 26, p. 266802, Jun. 2007, doi: 10.1103/PhysRevLett.98.266802.

[48] M. A. Kats, N. Yu, P. Genevet, Z. Gaburro, and F. Capasso, "Effect of radiation damping on the spectral response of plasmonic components," *Opt. Express, OE*, vol. 19, no. 22, pp. 21748–21753, Oct. 2011, doi: 10.1364/OE.19.021748.

[49] M. Miyao, T. Motooka, N. Natsuaki, and T. Tokuyama, "Change of the electron effective mass in extremely heavily doped n-type Si obtained by ion implantation and laser annealing," *Solid State Communications*, vol. 37, no. 7, pp. 605–608, Feb. 1981, doi: 10.1016/0038-1098(81)90144-7.

[50] W. Wesch, "Ion implantation in III–V compounds," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 68, no. 1, pp. 342–354, May 1992, doi: 10.1016/0168-583X(92)96105-8.

[51] S. J. Pearton, A. R. Von Neida, J. M. Brown, K. T. Short, L. J. Oster, and U. K. Chakrabarti, "Ion implantation damage and annealing in InAs, GaSb, and GaP," *Journal of Applied Physics*, vol. 64, no. 2, pp. 629–636, Jul. 1988, doi: 10.1063/1.341952.

[52] I. Sandall, C. H. Tan, A. Smith, and R. Gwilliam, "Planar InAs photodiodes fabricated using He ion implantation," *Opt. Express, OE*, vol. 20, no. 8, pp. 8575–8583, Apr. 2012, doi: 10.1364/OE.20.008575.

[53] M. Sotoodeh, A. H. Khalid, and A. A. Rezazadeh, “Empirical low-field mobility model for III–V compounds applicable in device simulation codes,” *Journal of Applied Physics*, vol. 87, no. 6, p. 2890, Feb. 2000, doi: 10.1063/1.372274.

[54] “Handbook of Chemistry and Physics 101st Edition.” http://hbcponline.com/faces/contents/ContentsSearch.xhtml;jsessionid=30E5E2A9765FA632F1A1AF2DED079747 (accessed Apr. 10, 2021).

[55] A. F. 723-88 Committee F-1 on Electronics, *Standard Practice for Conversion Between Resistivity and Dopant Density for Boron-Doped and Phosphorus-Doped Silicon*. West Conshohocken, PA: American Society for Testing and Materials, 1995.

[56] N. Arnold, R. Schmitt, and K. Heime, “Diffusion in III-V semiconductors,” *Journal of Physics D: Applied Physics*, vol. 17, no. 17, pp. 443–474, 1984, Accessed: Mar. 21, 2018. [Online].

[57] T. Aoyama, H. Tashiro, and K. Suzuki, “Diffusion of Boron, Phosphorus, Arsenic, and Antimony in Thermally Grown Silicon Dioxide,” *Journal of The Electrochemical Society*, vol. 146, no. 5, pp. 1879–1883, 1999, Accessed: Jul. 17, 2018. [Online].

[58] D. T. Moore, “Gradient-index optics: a review,” *Applied Optics*, vol. 19, no. 7, p. 1035, Apr. 1980, doi: 10.1364/AO.19.001035.

[59] T. Vasara, “Zn diffusion doping of III-V semiconductors for new optoelectronics applications Title: Zn diffusion doping of III-V semiconductors for new optoelectronics applications,” *Espoo*, vol. 199, 2017, Accessed: Mar. 29, 2018. [Online].

[60] M. Katayama, Y. Tokuda, Y. Inoue, A. Usami, and T. Wada, “Ga out-diffusion in rapid-thermal-processed GaAs with SiO 2 encapsulants,” *Journal of Applied Physics*, vol. 69, p. 3541, 1991, doi: 10.1063/1.348496.

[61] P. M. Campbell, O. Aina, and B. J. Baliga, “Enhanced protection of GaAs against thermal surface degradation by encapsulated annealing in an arsine ambient,” *Applied Physics Letters*, vol. 45, no. 1, pp. 95–97, Jul. 1984, doi: 10.1063/1.94982.

[62] Y. I. Nissim, J. F. Gibbons, C. A. Evans, V. R. Deline, and J. C. Norberg, “Thermal diffusion of tin in GaAs from a spin-on SnO2/SiO2 source,” *Applied Physics Letters*, vol. 37, no. 1, pp. 89–91, Jul. 1980, doi: 10.1063/1.91714.

[63] J. Crank, “the Mathematics of Diffusion,” 1975, doi: 10.1016/0306-4549(77)90072-X.

[64] D. Zwillinger, *Handbook of Differential Equations*. Academic Press, 2014.

[65] J. Robinson and Y. Rahmat-Samii, “Particle swarm optimization in electromagnetics,” *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 2, pp. 397–407, 2004, doi: 10.1109/TAP.2004.823969.

[66] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*. John Wiley & Sons, 2013.

[67] Y. Zhang, S. Wang, and G. Ji, “A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications,” *Mathematical Problems in Engineering*, vol. 2015, p. e931256, Oct. 2015, doi: 10.1155/2015/931256.

[68] M. Mitchell, *An Introduction to Genetic Algorithms*. MIT Press, 1998.

[69] A. Scollen and T. Hargraves, *Simulated Annealing: Introduction, Applications and Theory*. Nova Science Publishers, 2018.

[70] M. Erik, H. Pedersen, and M. E. H. Pedersen, “Good parameters for particle swarm optimization,” *Technical Report HL1001, Hvass Laboratories*, vol. HL1001, pp. 1–12, 2010.

[71] N. Yu and F. Capasso, “Flat optics with designer metasurfaces,” *Nature Materials*, vol. 13, no. 2, Art. no. 2, Feb. 2014, doi: 10.1038/nmat3839.

[72] D. Wang, Q. Fan, J. Wang, Z. Zhang, Y. Liang, and T. Xu, “All-dielectric metasurface beam deflector at the visible frequencies,” *Opto-electron. Eng., OEE*, vol. 44, no. 01, Art. no. 01, Jan. 2017, doi: 10.3969/j.issn.1003-501X.2017.01.012.

[73] T. Xu, C. Wang, C. Du, and X. Luo, “Plasmonic beam deflector,” *Opt. Express, OE*, vol. 16, no. 7, pp. 4753–4759, Mar. 2008, doi: 10.1364/OE.16.004753.

[74] D. Lin, P. Fan, E. Hasman, and M. L. Brongersma, “Dielectric gradient metasurface optical elements,” *Science*, vol. 345, no. 6194, pp. 298–302, Jul. 2014, doi: 10.1126/science.1253213.

[75] H.-T. Chen, A. J. Taylor, and N. Yu, “A review of metasurfaces: physics and applications,” *Reports on Progress in Physics*, vol. 79, no. 7, p. 076401, Jul. 2016, doi: 10.1088/0034-4885/79/7/076401.

[76] N. Li *et al.*, “Large-area pixelated metasurface beam deflector on a 12-inch glass wafer for random point generation,” *Nanophotonics*, vol. 8, no. 10, pp. 1855–1861, Oct. 2019, doi: 10.1515/nanoph-2019-0208.

[77] Z. Zhou *et al.*, “Efficient Silicon Metasurfaces for Visible Light,” *ACS Photonics*, vol. 4, no. 3, pp. 544–551, Mar. 2017, doi: 10.1021/acsphotonics.6b00740.

[78] N. Yu *et al.*, “Light Propagation with Phase Discontinuities: Generalized Laws of Reflection and Refraction,” *Science*, vol. 334, no. 6054, pp. 333–337, Oct. 2011, doi: 10.1126/science.1210713.

[79] M. P. Bendsoe and O. Sigmund, *Topology Optimization: Theory, Methods, and Applications*, 2nd ed. Berlin Heidelberg: Springer-Verlag, 2004.

[80] 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, no. 2, pp. 197–224, Nov. 1988, doi: 10.1016/0045-7825(88)90086-2.

[81] A. Ben-Tal and A. Nemirovski, “Robust Truss Topology Design via Semidefinite Programming,” *SIAM J. Optim.*, vol. 7, no. 4, pp. 991–1016, Nov. 1997, doi: 10.1137/S1052623495291951.

[82] O. Sigmund, “Design of multiphysics actuators using topology optimization – Part I: One-material structures,” *Computer Methods in Applied Mechanics and Engineering*, vol. 190, no. 49, pp. 6577–6604, Oct. 2001, doi: 10.1016/S0045-7825(01)00251-1.

[83] C. B. W. Pedersen, T. Buhl, and O. Sigmund, “Topology synthesis of large-displacement compliant mechanisms,” *International Journal for Numerical Methods in Engineering*, vol. 50, no. 12, pp. 2683–2705, 2001, doi: https://doi.org/10.1002/nme.148.

[84] O. Sigmund, “Manufacturing tolerant topology optimization,” *Acta Mechanica Sinica*, vol. 25, no. 2, pp. 227–239, Apr. 2009, doi: 10.1007/s10409-009-0240-z.

[85] J. Andkjær, V. E. Johansen, K. S. Friis, and O. Sigmund, “Inverse design of nanostructured surfaces for color effects,” *Journal of the Optical Society of America B*, vol. 31, no. 1, p. 164, Jan. 2014, doi: 10.1364/josab.31.000164.

[86] J. S. Jensen and O. Sigmund, “Topology optimization for nano-photonics,” *Laser & Photonics Reviews*, vol. 5, no. 2, pp. 308–321, Mar. 2011, doi: 10.1002/lpor.201000014.

[87] J. S. Jensen and O. Sigmund, “Topology optimization for nano-photonics,” *Laser & Photonics Reviews*, vol. 5, no. 2, pp. 308–321, Mar. 2011, doi: 10.1002/lpor.201000014.

[88] F. Wang, J. S. Jensen, and O. Sigmund, “Robust topology optimization of photonic crystal waveguides with tailored dispersion properties,” 2011, Accessed: Jun. 18, 2018. [Online].

[89] O. D. Miller, “Photonic Design: From Fundamental Solar Cell Physics to Computational Inverse Design,” 2012. Accessed: Aug. 15, 2019. [Online].

[90] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. L. Hollenberg, “The nitrogen-vacancy colour centre in diamond,” *Physics Reports*, vol. 528, no. 1. North-Holland, pp. 1–45, Jul. 01, 2013, doi: 10.1016/j.physrep.2013.02.001.

[91] A. Gruber, A. Dräbenstedt, C. Tietz, L. Fleury, J. Wrachtrup, and C. von Borczyskowski, “Scanning Confocal Optical Microscopy and Magnetic Resonance on Single Defect Centers,” *Science*, vol. 276, no. 5321, pp. 2012–2014, Jun. 1997, doi: 10.1126/science.276.5321.2012.

[92] M. V. G. Dutt *et al.*, “Quantum Register Based on Individual Electronic and Nuclear Spin Qubits in Diamond,” *Science*, vol. 316, no. 5829, pp. 1312–1316, Jun. 2007, doi: 10.1126/science.1139831.

[93] P. Neumann *et al.*, “Multipartite Entanglement Among Single Spins in Diamond,” *Science*, vol. 320, no. 5881, pp. 1326–1329, Jun. 2008, doi: 10.1126/science.1157233.

[94] R. J. Epstein, F. M. Mendoza, Y. K. Kato, and D. D. Awschalom, “Anisotropic interactions of a single spin and dark-spin spectroscopy in diamond,” *Nature Physics*, vol. 1, no. 2, Art. no. 2, Nov. 2005, doi: 10.1038/nphys141.

[95] P. Neumann *et al.*, “Single-Shot Readout of a Single Nuclear Spin,” *Science*, vol. 329, no. 5991, pp. 542–544, Jul. 2010, doi: 10.1126/science.1189075.

[96] F. Dolde *et al.*, “Electric-field sensing using single diamond spins,” *Nature Physics*, vol. 7, no. 6, Art. no. 6, Jun. 2011, doi: 10.1038/nphys1969.

[97] S. Steinert *et al.*, “High sensitivity magnetic imaging using an array of spins in diamond,” *Review of Scientific Instruments*, vol. 81, no. 4, p. 043705, Apr. 2010, doi: 10.1063/1.3385689.

[98] Ph. Tamarat *et al.*, “Stark Shift Control of Single Optical Centers in Diamond,” *Phys. Rev. Lett.*, vol. 97, no. 8, p. 083002, Aug. 2006, doi: 10.1103/PhysRevLett.97.083002.

[99] G. Balasubramanian *et al.*, “Nanoscale imaging magnetometry with diamond spins under ambient conditions,” *Nature*, vol. 455, no. 7213, Art. no. 7213, Oct. 2008, doi: 10.1038/nature07278.

[100] E. van Oort, N. B. Manson, and M. Glasbeek, “Optically detected spin coherence of the diamond N-V centre in its triplet ground state,” *J. Phys. C: Solid State Phys.*, vol. 21, no. 23, pp. 4385–4391, Aug. 1988, doi: 10.1088/0022-3719/21/23/020.

[101] H. R. Phillip and E. A. Taft, “Kramers-Kronig Analysis of Reflectance Data for Diamond,” *Phys. Rev.*, vol. 136, no. 5A, pp. A1445–A1448, Nov. 1964, doi: 10.1103/PhysRev.136.A1445.

[102] D. Le Sage *et al.*, “Efficient photon detection from color centers in a diamond optical waveguide,” *Phys. Rev. B*, vol. 85, no. 12, p. 121202, Mar. 2012, doi: 10.1103/PhysRevB.85.121202.

[103] F. Jelezko and J. Wrachtrup, “Single defect centres in diamond: A review,” *physica status solidi (a)*, vol. 203, no. 13, pp. 3207–3225, Oct. 2006, doi: 10.1002/pssa.200671403.

[104] R. Schirhagl, K. Chang, M. Loretz, and C. L. Degen, “Nitrogen-Vacancy Centers in Diamond: Nanoscale Sensors for Physics and Biology,” *Annual Review of Physical Chemistry*, vol. 65, no. 1, pp. 83–105, Apr. 2014, doi: 10.1146/annurev-physchem-040513-103659.

[105] I. Aharonovich and E. Neu, “Diamond Nanophotonics,” *Advanced Optical Materials*, vol. 2, no. 10, pp. 911–928, Oct. 2014, doi: 10.1002/adom.201400189.

[106] L. Rondin, J.-P. Tetienne, T. Hingant, J.-F. Roch, P. Maletinsky, and V. Jacques, “Magnetometry with nitrogen-vacancy defects in diamond,” *Reports on Progress in Physics*, vol. 77, no. 5, p. 056503, May 2014, doi: 10.1088/0034-4885/77/5/056503.

[107] J. F. Barry *et al.*, “Sensitivity optimization for NV-diamond magnetometry,” *Reviews of Modern Physics*, vol. 92, no. 1, p. 015004, Mar. 2020, doi: 10.1103/RevModPhys.92.015004.

[108] L. Childress and R. Hanson, “Diamond NV centers for quantum computing and quantum networks,” *MRS Bulletin*, vol. 38, no. 2. Cambridge University Press, pp. 134–138, Feb. 2013, doi: 10.1557/mrs.2013.20.

[109] A. Russo, E. Barnes, and S. E. Economou, “Photonic graph state generation from quantum dots and color centers for quantum communications,” *Physical Review B*, vol. 98, no. 8, p. 085303, Aug. 2018, doi: 10.1103/PhysRevB.98.085303.

[110] A. Beveratos, R. Brouri, T. Gacoin, A. Villing, J. P. Poizat, and P. Grangier, “Single Photon Quantum Cryptography,” *Physical Review Letters*, vol. 89, no. 18, p. 187901, 2002, doi: 10.1103/PhysRevLett.89.187901.

[111] F. Jelezko, T. Gaebel, I. Popa, M. Domhan, A. Gruber, and J. Wrachtrup, “Observation of coherent oscillation of a single nuclear spin and realization of a two-qubit conditional quantum gate,” *Physical Review Letters*, vol. 93, no. 13, p. 130501, Sep. 2004, doi: 10.1103/PhysRevLett.93.130501.

[112] M. V. Gurudev Dutt *et al.*, “Quantum register based on individual electronic and nuclear spin qubits in diamond,” *Science*, vol. 316, no. 5829, pp. 1312–1316, Jun. 2007, doi: 10.1126/science.1139831.

[113] B.-C. Ren and F.-G. Deng, “Hyperentanglement purification and concentration assisted by diamond NV centers inside photonic crystal cavities,” *Laser Physics Letters*, vol. 10, no. 11, p. 115201, Oct. 2013, doi: 10.1088/1612-2011/10/11/115201.

[114] K. Fang *et al.*, “High-sensitivity magnetometry based on quantum beats in diamond nitrogen-vacancy centers,” *Physical Review Letters*, vol. 110, no. 13, p. 130802, Mar. 2013, doi: 10.1103/PhysRevLett.110.130802.

[115] N. H. Wan *et al.*, “Efficient Extraction of Light from a Nitrogen-Vacancy Center in a Diamond Parabolic Reflector,” *Nano Letters*, vol. 18, no. 5, pp. 2787–2793, May 2018, doi: 10.1021/acs.nanolett.7b04684.

[116] J. T. Choy *et al.*, “Enhanced single-photon emission from a diamond–silver aperture,” *Nature Photonics*, vol. 5, no. 12, pp. 738–743, Dec. 2011, doi: 10.1038/nphoton.2011.249.

[117] T. Schröder, F. Gädeke, M. J. Banholzer, and O. Benson, “Ultrabright and efficient single-photon generation based on nitrogen-vacancy centres in nanodiamonds on a solid immersion lens,” *New Journal of Physics*, vol. 13, 2011, doi: 10.1088/1367-2630/13/5/055017.

[118] P. Siyushev *et al.*, “Monolithic diamond optics for single photon detection,” *Applied Physics Letters*, vol. 97, no. 24, p. 241902, Dec. 2010, doi: 10.1063/1.3519849.

[119] A. Faraon, P. E. Barclay, C. Santori, K.-M. C. Fu, and R. G. Beausoleil, “Resonant enhancement of the zero-phonon emission from a colour centre in a diamond cavity,” *Nature Photonics*, vol. 5, no. 5, pp. 301–305, May 2011, doi: 10.1038/nphoton.2011.52.

[120] J. P. Hadden *et al.*, “Strongly enhanced photon collection from diamond defect centers under microfabricated integrated solid immersion lenses,” *Applied Physics Letters*, vol. 97, no. 24, p. 241901, Dec. 2010, doi: 10.1063/1.3519847.

[121] T.-Y. Huang *et al.*, “A monolithic immersion metalens for imaging solid-state quantum emitters,” *Nature Communications*, vol. 10, no. 1, p. 2392, Dec. 2019, doi: 10.1038/s41467-019-10238-5.

[122] J. Zheng, A. C. Liapis, E. H. Chen, C. T. Black, and D. Englund, “Chirped circular dielectric gratings for near-unity collection efficiency from quantum emitters in bulk diamond,” *Optics Express*, vol. 25, no. 26, p. 32420, Dec. 2017, doi: 10.1364/oe.25.032420.

[123] L. Li *et al.*, “Efficient Photon Collection from a Nitrogen Vacancy Center in a Circular Bullseye Grating,” *Nano Letters*, vol. 15, no. 3, pp. 1493–1497, Mar. 2015, doi: 10.1021/nl503451j.

[124] S. Schietinger, M. Barth, T. Aichele, and O. Benson, “Plasmon-enhanced single photon emission from a nanoassembled metal - Diamond hybrid structure at room temperature,” *Nano Letters*, vol. 9, no. 4, pp. 1694–1698, Apr. 2009, doi: 10.1021/nl900384c.

[125] M. Y. Shalaginov *et al.*, “Enhancement of single-photon emission from nitrogen-vacancy centers with TiN/(Al,Sc)N hyperbolic metamaterial,” *Laser & Photonics Reviews*, vol. 9, no. 1, pp. 120–127, Jan. 2015, doi: 10.1002/lpor.201400185.

[126] P. Latawiec, M. J. Burek, Y.-I. Sohn, and M. Lončar, “Faraday cage angled-etching of nanostructures in bulk dielectrics,” *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 34, no. 4, p. 041801, Jul. 2016, doi: 10.1116/1.4944854.

[127] S. Sangtawesin *et al.*, “Origins of Diamond Surface Noise Probed by Correlating Single-Spin Measurements with Surface Spectroscopy,” *Physical Review X*, vol. 9, no. 3, p. 031052, Sep. 2019, doi: 10.1103/PhysRevX.9.031052.

[128] M. Kaviani, P. Deák, B. Aradi, T. Frauenheim, J. P. Chou, and A. Gali, “Proper surface termination for luminescent near-surface NV centers in diamond,” *Nano Letters*, vol. 14, no. 8, pp. 4772–4777, Aug. 2014, doi: 10.1021/nl501927y.

[129] J. Riedrich-Möller *et al.*, “Nanoimplantation and Purcell enhancement of single nitrogen-vacancy centers in photonic crystal cavities in diamond,” *Applied Physics Letters*, vol. 106, no. 22, p. 221103, Jun. 2015, doi: 10.1063/1.4922117.

[130] N. Abe, Y. Mitsumori, M. Sadgrove, and K. Edamatsu, “Dynamically unpolarized single-photon source in diamond with intrinsic randomness,” *Scientific Reports*, vol. 7, no. 1, Art. no. 1, Apr. 2017, doi: 10.1038/srep46722.

[131] N. B. Manson and J. P. Harrison, “Photo-ionization of the nitrogen-vacancy center in diamond,” *Diamond and Related Materials*, vol. 14, no. 10, pp. 1705–1710, Oct. 2005, doi: 10.1016/j.diamond.2005.06.027.

[132] R. Albrecht, A. Bommer, C. Deutsch, J. Reichel, and C. Becher, “Coupling of a Single Nitrogen-Vacancy Center in Diamond to a Fiber-Based Microcavity,” *Physical Review Letters*, vol. 110, no. 24, p. 243602, Jun. 2013, doi: 10.1103/PhysRevLett.110.243602.

[133] D. Sell, J. Yang, S. Doshay, R. Yang, and J. A. Fan, “Large-Angle, Multifunctional Metagratings Based on Freeform Multimode Geometries,” *Nano Letters*, vol. 17, no. 6, pp. 3752–3757, Jun. 2017, doi: 10.1021/acs.nanolett.7b01082.

[134] E. W. Wang, D. Sell, T. Phan, and J. A. Fan, “Robust design of topology-optimized metasurfaces,” 2019, doi: 10.1364/OME.9.000469.

[135] P. R. Dolan, X. Li, J. Storteboom, and M. Gu, “Complete determination of the orientation of NV centers with radially polarized beams,” *Opt. Express, OE*, vol. 22, no. 4, pp. 4379–4387, Feb. 2014, doi: 10.1364/OE.22.004379.

[136] “Lumerical Inc.” .

[137] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, “Adjoint shape optimization applied to electromagnetic design,” *Optics Express*, vol. 21, no. 18, p. 21693, Sep. 2013, doi: 10.1364/OE.21.021693.

[138] D. E. Aspnes and A. A. Studna, “Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV,” *Phys. Rev. B*, vol. 27, no. 2, pp. 985–1009, Jan. 1983, doi: 10.1103/PhysRevB.27.985.

[139] T. W. Hughes, M. Minkov, I. A. D. Williamson, and S. Fan, “Adjoint Method and Inverse Design for Nonlinear Nanophotonic Devices,” *ACS Photonics*, vol. 5, no. 12, pp. 4781–4787, Dec. 2018, doi: 10.1021/acsphotonics.8b01522.

[140] D. Sell, J. Yang, S. Doshay, R. Yang, and J. A. Fan, “Large-Angle, Multifunctional Metagratings Based on Freeform Multimode Geometries,” *Nano Letters*, vol. 17, no. 6, pp. 3752–3757, Jun. 2017, doi: 10.1021/acs.nanolett.7b01082.

[141] J. Lu and J. Vučković, “Nanophotonic computational design,” *Optics Express*, vol. 21, no. 11, p. 13351, Jun. 2013, doi: 10.1364/oe.21.013351.

[142] A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vucković, “Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer,” *Nature Photonics*, vol. 9, no. 6, pp. 374–377, Jun. 2015, doi: 10.1038/nphoton.2015.69.

[143] D. X. Xu *et al.*, “Silicon photonic integration platform-Have we found the sweet spot?,” *IEEE Journal on Selected Topics in Quantum Electronics*, vol. 20, no. 4, Jul. 2014, doi: 10.1109/JSTQE.2014.2299634.

[144] H. Yang *et al.*, “Transfer-printed stacked nanomembrane lasers on silicon,” *Nature Photonics*, vol. 6, no. 9, pp. 615–620, Sep. 2012, doi: 10.1038/nphoton.2012.160.

[145] P. Goy, J. M. Raimond, M. Gross, and S. Haroche, “Observation of cavity-enhanced single-atom spontaneous emission,” *Physical Review Letters*, vol. 50, no. 24, pp. 1903–1906, Jun. 1983, doi: 10.1103/PhysRevLett.50.1903.

[146] M. Parikh, “Corrections to proximity effects in electron beam lithography. I. Theory,” *Journal of Applied Physics*, vol. 50, no. 6, pp. 4371–4377, Jun. 1979, doi: 10.1063/1.326423.

[147] T. M. Babinec *et al.*, “A diamond nanowire single-photon source,” *Nature Nanotechnology*, vol. 5, no. 3, pp. 195–199, Feb. 2010, doi: 10.1038/nnano.2010.6.

[148] I. Bulu, T. Babinec, B. Hausmann, J. T. Choy, and M. Loncar, “Plasmonic resonators for enhanced diamond NV- center single photon sources,” *Optics Express*, vol. 19, no. 6, p. 5268, Mar. 2011, doi: 10.1364/oe.19.005268.

[149] J. T. Choy, I. Bulu, B. J. M. Hausmann, E. Janitz, I.-C. Huang, and M. Lončar, “Spontaneous emission and collection efficiency enhancement of single emitters in diamond via plasmonic cavities and gratings,” *Applied Physics Letters*, vol. 103, no. 16, p. 161101, Oct. 2013, doi: 10.1063/1.4817397.

[150] A. Karamlou, M. E. Trusheim, and D. Englund, “Metal-dielectric antennas for efficient photon collection from diamond color centers,” *Optics Express*, vol. 26, no. 3, p. 3341, Feb. 2018, doi: 10.1364/OE.26.003341.

[151] C. W. Cheng, K. T. Shiu, N. Li, S. J. Han, L. Shi, and D. K. Sadana, “Epitaxial lift-off process for gallium arsenide substrate reuse and flexible electronics,” *Nature Communications*, vol. 4, no. 1, pp. 1–7, Mar. 2013, doi: 10.1038/ncomms2583.

[152] S. Chakravarthi *et al.*, “Inverse-designed photon extractors for optically addressable defect qubits,” *Optica, OPTICA*, vol. 7, no. 12, pp. 1805–1811, Dec. 2020, doi: 10.1364/OPTICA.408611.

[153] M. Atatüre, D. Englund, N. Vamivakas, S. Y. Lee, and J. Wrachtrup, “Material platforms for spin-based photonic quantum technologies,” *Nature Reviews Materials*, vol. 3, no. 5. Nature Publishing Group, pp. 38–51, May 01, 2018, doi: 10.1038/s41578-018-0008-9.

[154] M. K. Boll, I. P. Radko, A. Huck, and U. L. Andersen, “Photophysics of quantum emitters in hexagonal boron-nitride nano-flakes,” *Optics Express*, vol. 28, no. 5, p. 7475, Mar. 2020, doi: 10.1364/oe.386629.

[155] M. Radtke, E. Bernardi, A. Slablab, R. Nelz, and E. Neu, “Nanoscale sensing based on nitrogen vacancy centers in single crystal diamond and nanodiamonds: achievements and challenges,” *Nano Futures*, vol. 3, no. 4, p. 042004, Dec. 2019, doi: 10.1088/2399-1984/ab5f9b.

[156] G. Davies, M. F. Hamer, and W. C. Price, “Optical studies of the 1.945 eV vibronic band in diamond,” *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 348, no. 1653, pp. 285–298, Feb. 1976, doi: 10.1098/rspa.1976.0039.

[157] A. M. Edmonds *et al.*, “Production of oriented nitrogen-vacancy color centers in synthetic diamond,” *Phys. Rev. B*, vol. 86, no. 3, p. 035201, Jul. 2012, doi: 10.1103/PhysRevB.86.035201.

[158] A. Tallaire *et al.*, “High quality thick CVD diamond films homoepitaxially grown on (111)-oriented substrates,” *Diamond and Related Materials*, vol. 41, pp. 34–40, Jan. 2014, doi: 10.1016/j.diamond.2013.11.002.

[159] J. Michl *et al.*, “Perfect alignment and preferential orientation of nitrogen-vacancy centers during chemical vapor deposition diamond growth on (111) surfaces,” *Appl. Phys. Lett.*, vol. 104, no. 10, p. 102407, Mar. 2014, doi: 10.1063/1.4868128.

[160] H. Ishiwata, M. Nakajima, K. Tahara, H. Ozawa, T. Iwasaki, and M. Hatano, “Perfectly aligned shallow ensemble nitrogen-vacancy centers in (111) diamond,” *Appl. Phys. Lett.*, vol. 111, no. 4, p. 043103, Jul. 2017, doi: 10.1063/1.4993160.

[161] M. Lesik *et al.*, “Perfect preferential orientation of nitrogen-vacancy defects in a synthetic diamond sample,” *Appl. Phys. Lett.*, vol. 104, no. 11, p. 113107, Mar. 2014, doi: 10.1063/1.4869103.

[162] M. Kasu, T. Makimoto, W. Ebert, and E. Kohn, “Formation of stacking faults containing microtwins in (111) chemical-vapor-deposited diamond homoepitaxial layers,” *Appl. Phys. Lett.*, vol. 83, no. 17, pp. 3465–3467, Oct. 2003, doi: 10.1063/1.1622105.

[163] C.-S. Yan and Y. K. Vohra, “Multiple twinning and nitrogen defect center in chemical vapor deposited homoepitaxial diamond,” *Diamond and Related Materials*, vol. 8, no. 11, pp. 2022–2031, Nov. 1999, doi: 10.1016/S0925-9635(99)00148-X.

[164] J. Achard, V. Jacques, and A. Tallaire, “Chemical vapour deposition diamond single crystals with nitrogen-vacancy centres: a review of material synthesis and technology for quantum sensing applications,” *J. Phys. D: Appl. Phys.*, vol. 53, no. 31, p. 313001, May 2020, doi: 10.1088/1361-6463/ab81d1.

[165] C. Osterkamp *et al.*, “Engineering preferentially-aligned nitrogen-vacancy centre ensembles in CVD grown diamond,” *Scientific Reports*, vol. 9, no. 1, Art. no. 1, Apr. 2019, doi: 10.1038/s41598-019-42314-7.

[166] R. A. Wambold *et al.*, “Adjoint-optimized nanoscale light extractor for nitrogen-vacancy centers in diamond,” *Nanophotonics*, vol. 10, no. 1, pp. 393–401, Nov. 2020, doi: 10.1515/nanoph-2020-0387.

[167] L. Li *et al.*, “Efficient Photon Collection from a Nitrogen Vacancy Center in a Circular Bullseye Grating,” *Nano Letters*, vol. 15, no. 3, pp. 1493–1497, Mar. 2015, doi: 10.1021/nl503451j.

[168] H. Yang *et al.*, “Transfer-printed stacked nanomembrane lasers on silicon,” *Nature Photonics*, vol. 6, no. 9, pp. 615–620, Sep. 2012, doi: 10.1038/nphoton.2012.160.

[169] T.-Y. Huang *et al.*, “A monolithic immersion metalens for imaging solid-state quantum emitters,” *Nature Communications*, vol. 10, no. 1, p. 2392, Dec. 2019, doi: 10.1038/s41467-019-10238-5.

[170] R. N. Patel *et al.*, “Efficient photon coupling from a diamond nitrogen vacancy center by integration with silica fiber,” *Light: Science and Applications*, vol. 5, no. 2, pp. e16032–e16032, Feb. 2016, doi: 10.1038/lsa.2016.32.

[171] J. T. Choy *et al.*, “Enhanced single-photon emission from a diamond–silver aperture,” *Nature Photonics*, vol. 5, no. 12, pp. 738–743, Dec. 2011, doi: 10.1038/nphoton.2011.249.

[172] I. Bulu, T. Babinec, B. Hausmann, J. T. Choy, and M. Loncar, “Plasmonic resonators for enhanced diamond NV- center single photon sources,” *Optics Express*, vol. 19, no. 6, p. 5268, Mar. 2011, doi: 10.1364/oe.19.005268.

[173] J. T. Choy, I. Bulu, B. J. M. Hausmann, E. Janitz, I.-C. Huang, and M. Lončar, “Spontaneous emission and collection efficiency enhancement of single emitters in diamond via plasmonic cavities and gratings,” *Applied Physics Letters*, vol. 103, no. 16, p. 161101, Oct. 2013, doi: 10.1063/1.4817397.

[174] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, “Adjoint shape optimization applied to electromagnetic design,” *Optics Express*, vol. 21, no. 18, p. 21693, Sep. 2013, doi: 10.1364/OE.21.021693.

[175] D. L. Moehring *et al.*, “Entanglement of single-atom quantum bits at a distance,” *Nature*, vol. 449, no. 7158, Art. no. 7158, Sep. 2007, doi: 10.1038/nature06118.

[176] W. Pfaff *et al.*, “Demonstration of entanglement-by-measurement of solid-state qubits,” *Nature Physics*, vol. 9, no. 1, Art. no. 1, Jan. 2013, doi: 10.1038/nphys2444.

[177] L. Robledo, L. Childress, H. Bernien, B. Hensen, P. F. A. Alkemade, and R. Hanson, "High-fidelity projective read-out of a solid-state spin quantum register," *Nature*, vol. 477, no. 7366, Art. no. 7366, Sep. 2011, doi: 10.1038/nature10401.

[178] P. Neumann *et al.*, "Quantum register based on coupled electron spins in a room-temperature solid," *Nature Physics*, vol. 6, no. 4, Art. no. 4, Apr. 2010, doi: 10.1038/nphys1536.

[179] P. Tamarat *et al.*, "Spin-flip and spin-conserving optical transitions of the nitrogen-vacancy centre in diamond," *New J. Phys.*, vol. 10, no. 4, p. 045004, Apr. 2008, doi: 10.1088/1367-2630/10/4/045004.

[180] N. B. Manson, J. P. Harrison, and M. J. Sellars, "Nitrogen-vacancy center in diamond: Model of the electronic structure and associated dynamics," *Phys. Rev. B*, vol. 74, no. 10, p. 104303, Sep. 2006, doi: 10.1103/PhysRevB.74.104303.

[181] K.-M. C. Fu, C. Santori, P. E. Barclay, L. J. Rogers, N. B. Manson, and R. G. Beausoleil, "Observation of the Dynamic Jahn-Teller Effect in the Excited States of Nitrogen-Vacancy Centers in Diamond," *Phys. Rev. Lett.*, vol. 103, no. 25, p. 256404, Dec. 2009, doi: 10.1103/PhysRevLett.103.256404.

[182] D. Le Sage *et al.*, "Optical magnetic imaging of living cells," *Nature*, vol. 496, no. 7446, Art. no. 7446, Apr. 2013, doi: 10.1038/nature12072.

[183] D. A. Hopper, H. J. Shulevitz, and L. C. Bassett, "Spin Readout Techniques of the Nitrogen-Vacancy Center in Diamond," *Micromachines*, vol. 9, no. 9, Art. no. 9, Sep. 2018, doi: 10.3390/mi9090437.

[184] H. C. Davis *et al.*, "Mapping the microscale origins of magnetic resonance image contrast with subcellular diamond magnetometry," *Nature Communications*, vol. 9, no. 1, Art. no. 1, Jan. 2018, doi: 10.1038/s41467-017-02471-7.

[185] E. Togan, “Optical Control of Individual Nitrogen-Vacancy Centers in Diamond,” Feb. 2013, Accessed: Apr. 21, 2021. [Online]. Available: https://dash.harvard.edu/handle/1/10336914.

[186] H.-H. Hsiao, C. H. Chu, and D. P. Tsai, “Fundamentals and Applications of Metasurfaces,” *Small Methods*, vol. 1, no. 4, p. 1600064, 2017, doi: https://doi.org/10.1002/smtd.201600064.

[187] A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, “Planar Photonics with Metasurfaces,” *Science*, vol. 339, no. 6125, Mar. 2013, doi: 10.1126/science.1232009.

[188] N. I. Zheludev and Y. S. Kivshar, “From metamaterials to metadevices,” *Nature Materials*, vol. 11, no. 11, Art. no. 11, Nov. 2012, doi: 10.1038/nmat3431.

[189] M. Farmahini-Farahani and H. Mosallaei, “Birefringent reflectarray metasurface for beam engineering in infrared,” 2013.

[190] S. Zhang *et al.*, “High efficiency near diffraction-limited mid-infrared flat lenses based on metasurface reflectarrays,” *Opt. Express, OE*, vol. 24, no. 16, pp. 18024–18034, Aug. 2016, doi: 10.1364/OE.24.018024.

[191] A. Forouzmand and H. Mosallaei, “Tunable two dimensional optical beam steering with reconfigurable indium tin oxide plasmonic reflectarray metasurface,” *J. Opt.*, vol. 18, no. 12, p. 125003, Dec. 2016, doi: 10.1088/2040-8978/18/12/125003.

[192] D. M. Pozar, S. D. Targonski, and H. D. Syrigos, “Design of millimeter wave microstrip reflectarrays,” *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 2, pp. 287–296, 1997, doi: 10.1109/8.560348.

[193] A. Moreau *et al.*, “Controlled-reflectance surfaces with film-coupled colloidal nanoantennas,” *Nature*, vol. 492, no. 7427, Art. no. 7427, Dec. 2012, doi: 10.1038/nature11615.

[194] N. Mohammadi Estakhri and A. Alù, “Manipulating optical reflections using engineered nanoscale metasurfaces,” *Phys. Rev. B*, vol. 89, no. 23, p. 235419, Jun. 2014, doi: 10.1103/PhysRevB.89.235419.

[195] P. Thureja, G. K. Shirmanesh, K. T. Fountaine, R. Sokhoyan, M. Grajower, and H. A. Atwater, “Array-Level Inverse Design of Beam Steering Active Metasurfaces,” *ACS Nano*, vol. 14, no. 11, pp. 15042–15055, Nov. 2020, doi: 10.1021/acsnano.0c05026.

[196] Q. Tan, Z. Xu, D. H. Zhang, T. Yu, S. Zhang, and Y. Luo, “Polarization-Controlled Plasmonic Structured Illumination,” *Nano Lett.*, vol. 20, no. 4, pp. 2602–2608, Apr. 2020, doi: 10.1021/acs.nanolett.0c00091.

[197] G. K. Shirmanesh, R. Sokhoyan, P. C. Wu, and H. A. Atwater, “Electro-optically Tunable Multifunctional Metasurfaces,” *ACS Nano*, vol. 14, no. 6, pp. 6912–6920, Jun. 2020, doi: 10.1021/acsnano.0c01269.

[198] J. Park, J.-H. Kang, S. J. Kim, X. Liu, and M. L. Brongersma, “Dynamic Reflection Phase and Polarization Control in Metasurfaces,” *Nano Lett*, vol. 17, p. 21, 2017, doi: 10.1021/acs.nanolett.6b04378.

[199] Y. Yang, W. Wang, P. Moitra, I. I. Kravchenko, D. P. Briggs, and J. Valentine, “Dielectric Meta-Reflectarray for Broadband Linear Polarization Conversion and Optical Vortex Generation,” *Nano Lett*, vol. 14, p. 42, 2014, doi: 10.1021/nl4044482.

[200] S. Sun *et al.*, “High-Efficiency Broadband Anomalous Reflection by Gradient Meta-Surfaces,” *Nano Lett*, vol. 12, p. 8, 2012, doi: 10.1021/nl3032668.

[201] Z. H. Jiang, S. Yun, L. Lin, J. A. Bossard, D. H. Werner, and T. S. Mayer, “Tailoring Dispersion for Broadband Low-loss Optical Metamaterials Using Deep-subwavelength Inclusions,” *Scientific Reports*, vol. 3, no. 1, Art. no. 1, Mar. 2013, doi: 10.1038/srep01571.

[202] X. Ni, N. K. Emani, A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Broadband Light Bending with Plasmonic Nanoantennas," *Science*, vol. 335, no. 6067, pp. 427–427, Jan. 2012, doi: 10.1126/science.1214686.

[203] M. Kang, T. Feng, H.-T. Wang, and J. Li, "Wave front engineering from an array of thin aperture antennas," *Opt. Express, OE*, vol. 20, no. 14, pp. 15882–15890, Jul. 2012, doi: 10.1364/OE.20.015882.

[204] L. Huang *et al.*, "Dispersionless Phase Discontinuities for Controlling Light Propagation," *Nano Lett.*, vol. 12, no. 11, pp. 5750–5755, Nov. 2012, doi: 10.1021/nl303031j.

[205] M. Kang, J. Chen, X.-L. Wang, and H.-T. Wang, "Twisted vector field from an inhomogeneous and anisotropic metamaterial," *J. Opt. Soc. Am. B, JOSAB*, vol. 29, no. 4, pp. 572–576, Apr. 2012, doi: 10.1364/JOSAB.29.000572.

[206] T. Søndergaard, J. Jung, S. I. Bozhevolnyi, and G. D. Valle, "Theoretical analysis of gold nano-strip gap plasmon resonators," *New J. Phys.*, vol. 10, no. 10, p. 105008, Oct. 2008, doi: 10.1088/1367-2630/10/10/105008.

[207] J. Zuloaga and P. Nordlander, "On the Energy Shift between Near-Field and Far-Field Peak Intensities in Localized Plasmon Systems," *Nano Lett.*, vol. 11, no. 3, pp. 1280–1283, Mar. 2011, doi: 10.1021/nl1043242.

[208] F. Ding, A. Pors, and S. I. Bozhevolnyi, "Gradient metasurfaces: a review of fundamentals and applications," *Rep. Prog. Phys.*, vol. 81, no. 2, p. 026401, Dec. 2017, doi: 10.1088/1361-6633/aa8732.

[209] A. Pors, O. Albrektsen, I. P. Radko, and S. I. Bozhevolnyi, "Gap plasmon-based metasurfaces for total control of reflected light," *Scientific Reports*, vol. 3, no. 1, p. 2155, Dec. 2013, doi: 10.1038/srep02155.

[210] C. Qu *et al.*, “Tailor the Functionalities of Metasurfaces Based on a Complete Phase Diagram,” *Phys. Rev. Lett.*, vol. 115, no. 23, p. 235503, Dec. 2015, doi: 10.1103/PhysRevLett.115.235503.

[211] S. Fan, W. Suh, and J. D. Joannopoulos, “Temporal coupled-mode theory for the Fano resonance in optical resonators,” *J. Opt. Soc. Am. A, JOSAA*, vol. 20, no. 3, pp. 569–572, Mar. 2003, doi: 10.1364/JOSAA.20.000569.

[212] R. Holland and J. W. Williams, “Total-Field versus Scattered-Field Finite-Difference Codes: A Comparative Assessment,” *IEEE Transactions on Nuclear Science*, vol. 30, no. 6, pp. 4583–4588, Dec. 1983, doi: 10.1109/TNS.1983.4333175.

[213] J. B. Schneider, “Plane waves in FDTD simulations and a nearly perfect total-field/scattered-field boundary,” *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 12, pp. 3280–3287, Dec. 2004, doi: 10.1109/TAP.2004.836403.

[214] I. R. Capoglu and G. S. Smith, “A Total-Field/Scattered-Field Plane-Wave Source for the FDTD Analysis of Layered Media,” *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 1, pp. 158–169, Jan. 2008, doi: 10.1109/TAP.2007.913088.

[215] M. J. Dodge, “Refractive properties of magnesium fluoride,” *Appl. Opt., AO*, vol. 23, no. 12, pp. 1980–1985, Jun. 1984, doi: 10.1364/AO.23.001980.

[216] P. B. Johnson and R. W. Christy, “Optical Constants of the Noble Metals,” *Phys. Rev. B*, vol. 6, no. 12, pp. 4370–4379, Dec. 1972, doi: 10.1103/PhysRevB.6.4370.

[217] N. I. Landy, S. Sajuyigbe, J. J. Mock, D. R. Smith, and W. J. Padilla, “Perfect Metamaterial Absorber,” *Phys. Rev. Lett.*, vol. 100, no. 20, p. 207402, May 2008, doi: 10.1103/PhysRevLett.100.207402.

[218] K. Aydin, V. E. Ferry, R. M. Briggs, and H. A. Atwater, “Broadband polarization-independent resonant light absorption using ultrathin plasmonic super absorbers,” *Nature Communications*, vol. 2, no. 1, Art. no. 1, Nov. 2011, doi: 10.1038/ncomms1528.

[219] Y. Avitzour, Y. A. Urzhumov, and G. Shvets, “Wide-angle infrared absorber based on a negative-index plasmonic metamaterial,” *Phys. Rev. B*, vol. 79, no. 4, p. 045131, Jan. 2009, doi: 10.1103/PhysRevB.79.045131.

[220] G. K. Shirmanesh, R. Sokhoyan, R. A. Pala, H. A. Atwater, and T. J. Watson, “Dual-Gated Active Metasurface at 1550 nm with Wide (>300°) Phase Tunability,” vol. 18, p. 25, 2018, doi: 10.1021/acs.nanolett.8b00351.

[221] M. C. Sherrott *et al.*, “Experimental Demonstration of >230° Phase Modulation in Gate-Tunable Graphene−Gold Reconfigurable Mid-Infrared Metasurfaces,” 2017, doi: 10.1021/acs.nanolett.7b00359.

[222] J. Park, J.-H. Kang, X. Liu, and M. L. Brongersma, “Electrically Tunable Epsilon-Near-Zero (ENZ) Metafilm Absorbers,” *Scientific Reports*, vol. 5, no. 1, Art. no. 1, Nov. 2015, doi: 10.1038/srep15754.

[223] Y. Shen *et al.*, “Optical vortices 30 years on: OAM manipulation from topological charge to multiple singularities,” *Light: Science & Applications*, vol. 8, no. 1, Art. no. 1, Oct. 2019, doi: 10.1038/s41377-019-0194-2.

[224] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, “Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes,” *Phys. Rev. A*, vol. 45, no. 11, pp. 8185–8189, Jun. 1992, doi: 10.1103/PhysRevA.45.8185.

[225] R. M. Gomes, A. Salles, F. Toscano, P. H. S. Ribeiro, and S. P. Walborn, “Observation of a Nonlocal Optical Vortex,” *Phys. Rev. Lett.*, vol. 103, no. 3, p. 033602, Jul. 2009, doi: 10.1103/PhysRevLett.103.033602.